Université de Montréal

Using Whole-Exome Sequencing Data in an Exome-Wide Association Study Approach to Identify Genetic Risk Factors Influencing Acute Lymphoblastic Leukemia Response: A Focus on Asparaginase Complications & Vincristine-Induced Peripheral Neuropathy

par

RACHID ABAJI

Département de Pharmacologie et de Physiologie

Faculté de médicine

Thèse présentée en vue de l'obtention du grade de

Doctorat en Pharmacologie

(option Phamacogénomique)

Décembre 2018

© RACHID ABAJI, 2018

Université de Montréal

Département de Pharmacologie et de Physiologie

Faculté de médicine

Cette thèse intitulée :

Using Whole-Exome Sequencing Data in an Exome-Wide Association Study Approach to Identify Genetic Risk Factors Influencing Acute Lymphoblastic Leukemia Response: A Focus on Asparaginase Complications & Vincristine-Induced Peripheral Neuropathy

Présentée par :

Rachid Abaji

A été évaluée par un jury composé des personnes suivantes

Dr. Noël J-M Raynal
Président-rapporteur

Dre. Maja Krajinovic
Directeur de recherche

Dr. Simon De Denus
Membre du jury

Dre. Isabelle Laverdière
Examinateur externe

Dr. Réné Cardinal
Représentant du doyen

I. Resumé

Le traitement de la leucémie lymphoblastique aiguë (LLA) de l'enfant, une affection d'origine maligne des cellules progénitrices lymphoïdes, s'est considérablement amélioré au cours des dernières décennies. En effet, le taux de succès du traitement a dépassé 90% dans des conditions favorables. Cependant, des toxicités liées au traitement peuvent être fatales et entrainer l'interruption ou la cessation du traitement. L'allergie, la pancréatite et la thrombose sont des complications fréquentes du traitement de la LLA et sont associées à l'utilisation de l'asparaginase (ASNase), tandis qu'une toxicité fréquente due à la vincristine (VCR) induit la neuropathie périphérique (VIPN). Étant donné que l'ajustement du schéma posologique afin d'augmenter l'efficacité et diminuer la toxicité est un processus sensible, ceci demeure un défi majeur dans plusieurs protocoles de traitement. La pharmacogénétique étudie comment des altérations de la composante génétique peuvent influer sur la variabilité interindividuelle observée dans la réponse au traitement. Une meilleure compréhension de la base moléculaire de cette variabilité pourrait améliorer considérablement les résultats du traitement, en permettant la personnalisation de ce dernier en fonction du profil génétique du patient.

Des études récentes suggèrent l'avantage d'appliquer l'analyse de l'exome à la découverte de variants associés à des traits humains complexes ainsi qu'à des phénotypes de réactions médicamenteuses. L'objectif de notre travail était d'utiliser les données de séquençage pour réaliser des études d'association à l'échelle de l'exome, y compris des étapes de filtrage et de validation, afin d'identifier de nouveaux variants génétiques susceptibles de moduler le risque de développer des complications associées à ASNase et à VIPN.

Douze SNP étaient associés à des complications due à l'ASNase dans la cohorte initiale, dont 3 étaient associés à une allergie, 3 à une pancréatite et 6 à une thrombose. Parmi ceux-ci, les variants rs3809849, rs11556218 et rs34708521 des gènes *MYBBP1A*, *IL16* et *SPEF2* respectivement ont été associés à des complications multiples et leur association à une pancréatite a été répliquée dans une cohorte de validation indépendante. En ce qui concerne la VCR, trois variantes ont été associées à la modulation du risque de VIPN: rs2781377 dans *SYNE2*, rs10513762 dans *MRPL47* et rs3803357 dans *BAHD1*. Nous démontrons également le puissant effet combiné de la présence de plusieurs variants de risque pour chacune des toxicités étudiées et fournissons des modèles de prédiction du risque pour la pancréatite et le VIPN basés sur la méthode d'évaluation du risque génétique pondérée et qui ont été validés à l'interne.

De plus, étant donné une association du polymorphisme du gène *MYBBP1A* avec de multiples issus de traitement, nous avons cherché à comprendre comment cette altération génétique se traduit par des variabilités de réponse aux traitements à l'ASNase. En utilisant la technique CRISPR-CAS9 pour induire l'inactivation de gènes dans des lignées cellulaires cancéreuses PANC1 (pancréatiques) nous avons testé la différence de viabilité entre les cellules inactivées et les cellules du type sauvage à la suite de la suppression du gène et du traitement par ASNase. Nos résultats suggèrent un rôle fonctionnel de ce gène dans la modulation de la viabilité, de la capacité de prolifération et de la morphologie des cellules knock-out, ainsi que dans leur sensibilité à l'ASNase, et plaident en outre pour que le gène influence l'issus du traitement de la LLA par ASNase.

Le présent travail démontre que l'utilisation de l'approche de séquençage de l'exome entier dans le contexte d'une étude d'association à l'échelle de l'exome est une stratégie valide « sans hypothèse » pour identifier de nouveaux marqueurs génétiques modulant l'effet du traitement de la LLA de l'enfant, et souligne l'importance de l'effet synergique de la combinaison des locus à risque.

Mots clés:

Leucémie lymphoblastique aiguë; effets indésirables des médicaments; asparaginase; étude d'association; pan-exomique; neuropathie; pancréatite; pharmacogénétique; vincristine; séquençage de l'exome entier.

II. Abstract

Treatment of childhood acute lymphoblastic leukemia (ALL), a malignant disorder of lymphoid progenitor cells has improved significantly over the past decades and treatment success rates have surpassed 90% in favorable settings. However, treatment-related toxicities can be life-threatening and cause treatment interruption or cessation. Allergy, pancreatitis and thrombosis are common complications of ALL treatment associated with the use of asparaginase (ASNase), while vincristine-induced peripheral neuropathy (VIPN) is a frequent toxicity of vincristine (VCR). It is a sensitive process and a constant struggle to adjust the dosing regimen to ensure maximum efficacy and minimum toxicity. Pharmacogenetics studies show alterations in the genetic component between individuals can influence the observed variability in treatment response. A better understanding of the molecular basis of this variability in drug effect could significantly improve treatment outcome by allowing the personalization of ALL treatment based on the genetic profile of the patient.

Emerging reports suggest the benefit of applying exome analysis to uncover variants associated with complex human traits as well as drug response phenotypes. Our objective in this work was to use available whole-exome sequencing data to perform exome-wide association studies followed by stepwise filtering and validation processes to identify novel variants with a potential to modulate the risk of developing ASNase complications and VIPN.

Twelve SNPs were associated with ASNase complications in the discovery cohort including 3 associated with allergy, 3 with pancreatitis and 6 with thrombosis. Of those, rs3809849 in MYBBP1A, rs11556218 in IL16 and rs34708521 in SPEF2 genes were associated with multiple complications and their association with pancreatitis was replicated in an independent validation cohort. As for VCR, three variants were associated with modulating the risk of VIPN: rs2781377 in SYNE2, rs10513762 in MRPL47 and rs3803357 in BAHD1. We also demonstrate a strong combined effect of harbouring multiple risk variants for each of the studied toxicities, and provide internally-validated risk-prediction models based on the weighted genetic risk score method for pancreatitis and VIPN.

Furthermore, given the association of the polymorphism in MYBBP1A gene with multiple treatment outcomes, we aimed at understanding how this genetic alteration translates into differences in ASNase treatment response through cell-based functional analysis. Using CRISPR-CAS9 technology we produced gene knockout of PANC1 (pancreatic) cancer cell-lines and tested the difference in viability between the knockouts and wild-type cells following gene deletion and ASNase treatment. Our results suggest a functional role of this gene in modulating the viability, proliferation capacity and the morphology of the knockout cells as well as their sensitivity to ASNase and further advocates the implication of the gene in influencing the outcome of ALL treatment with ASNase.

The present work demonstrates that using whole-exome sequencing data in the context of exome-wide association study is a successful "hypothesis-free" strategy for identifying novel genetic markers modulating the effect of childhood ALL treatment and highlights the importance of the synergistic effect of combining risk loci.

Keywords:

Acute lymphoblastic leukemia; adverse drug reactions; asparaginase; association study; exome-wide; neuropathy; pancreatitis; pharmacogenetics; vincristine; whole-exome sequencing.

III. Table of Contents

I. Resumé .... **i**

II. Abstract .... **iv**

III. Table of Contents .... **vii**

IV. List of Figures .... **xv**

V. List of Tables .... **xviii**

VI. Abbreviations and Acronyms .... **xx**

VII. Dedication .... **xxvi**

VIII. Acknowledgement .... **xxvii**

IX. Preface .... **xxix**

Section A 1

Chapter 1 1

1. General Introduction 2

1.1. Definition & Statistics 2

1.2. Prognostic Factors 5

1.2.1. Age & WBC count 6

1.2.2. Immunophenotype 6

1.2.3. Cytogenetics & Molecular Genetics 7

1.2.4. Early Treatment Response: 11

1.3. ALL Treatment and Outcome 12

1.3.1. Improvement in Treatment Outcome 12

1.3.2. Contemporary Therapy ... 16
1.3.3. Treatment of Refractory/Relapsed ALL ... 24
1.3.4. Targeted Therapy and Precision Medicine ... 25
1.3.5. Short and Long Term Toxic Effects of Treatment ... 26
1.3.6. Overview on selected ALL treatment-related toxicities important for this thesis. ... 28
1.3.6.1. Chemotherapy-Induced Acute Pancreatitis ... 28
1.3.6.2. Chemotherapy-Induced Peripheral Neuropathy ... 29
1.4. Pharmacogenomics and Pharmacogenetics ... 31
1.4.1. Genetic-association approaches ... 32
1.4.2. Genotyping vs. Sequencing methods ... 34
1.4.2.1. Genotyping ... 34
1.4.2.2. Sequencing ... 36
1.4.2.3. Whole-Exome Sequencing ... 41
1.4.3. Genome Editing Techniques ... 42
1.5. Study Hypotheses ... 44
1.6. Research Objectives ... 45
1.7. References ... 46

Chapter 2 ... 56

2.1. Abstract: ... 59
2.2. Introduction ... 60
2.3. TPMT Pharmacogenetics ... 62
2.3.1. Acute Lymphoblastic Leukemia (ALL) ... 66
2.3.2. Inflammatory bowel disease (IBD) ... 71
2.3.3. Autoimmune Disorders ... 74

2.4. Cost-Effectiveness 76

2.5. Impact on Clinical Practice 77

2.6. TPMT in the New Era of Sequencing 79

2.7. Conclusion 80

2.8. Perspective 82

2.9. Acknowledgments 83

2.10. Disclosure 83

2.11. References 84

Section B 98

Chapter 3 98

3.1. Abstract 101

3.2. General Introduction 102

3.2.1. Asparaginase and Acute Lymphoblastic Leukemia 102

3.2.2. Mechanism of Action, Resistance & Formulations 104

3.2.3. Pharmacogenetics of Asparaginase 108

3.2.3.1. Hypersensitivity Reactions, Pancreatitis & Thrombosis 108

3.2.3.2. Other Less Common Toxicities 117

3.2.3.3. Relapse 118

3.2.4. MicroRNA 120

3.3. General Conclusion 121

3.4. Declarations 123

3.5. Rreferences 124

3.6. Tables & Figures 130

Chapter 4 134

4.1. Abstract 138

4.2. Introduction 139

4.3. Results 142

4.3.1. Asparaginase-related complications 142

4.3.2. Association Study 143

4.3.3. Replication Analysis 145

4.3.4. Combined Effect Model 146

4.3.5. Risk Prediction 147

4.4. Discussion 149

4.5. Patients and methods 154

4.5.1. Study population and endpoints in the analysis 154

4.5.2. Whole Exome Sequencing (WES) 155

4.5.3. Validation of top-ranking EWAS signals by Genotyping 156

4.5.4. Risk Prediction 157

4.6. Authorship Contributions 158

4.7. Acknowledgements 158

4.8. Disclosure of Conflicts of Interest 158

4.9. Funding 158

4.10. References 159

4.11. Tables 164

4.12. Table and Figure Legends 168

4.13. Figures 172

4.14. Supplemental Material 178
4.15. Unpublished Data 192
4.15.1. rs3809849 in *MYBBP1A* gene and risk of osteonecrosis 193
4.15.2. rs11556218 in IL16 gene and event-free survival 196
4.15.3. rs11556218 in IL16 gene and drug sensitivity 198

Chapter 5 200

5.1. Abstract 203
5.2. Introduction 205
5.3. Materials and Methods 206
5.3.1. PANC1 cell line acquisition and maintenance 206
5.3.2. MYBBP1A Knock-out PANC1 cell line production 207
5.3.2.1. Cas9 cloning into the PANC1 cell genome 207
5.3.2.2. Knocking-out of MYBBP1A 207
5.3.3. Proliferation capacity assay 208
5.3.4. IC_{50} determination assay 209
5.3.5. Colony formation assay 210
5.3.6. Flow Cytometry analysis 211
5.3.7. Epithelial-Mesenchymal Transition characterization 212
5.4. Results 213
5.4.1. MYBBP1A Knock-out PANC1 cell line production & characterization 213
5.4.2. Proliferation capacity 214
5.4.3. IC_{50} determination and in-vitro sensitivity 214
5.4.4. Colony formation 215
5.4.5. Flow Cytometry analysis 215

5.4.6. Epithelial-Mesenchymal Transition ... 217
5.5. Discussion ... 218
5.6. Authorship Contributions ... 225
5.7. Disclosure of Conflicts of Interest ... 225
5.8. Figure Legends ... 226
5.9. Figures ... 229
5.10. References ... 234
5.11. Supplemental Material ... 237
5.11.1. Supplemental Figures ... 238
5.11.2. Supplemental Methods ... 242
Section C ... 244
Chapter 6 ... 244
6.1. Abstract ... 247
6.2. Introduction ... 248
6.3. Patients and Methods ... 251
6.3.1. Study population and endpoints in the analysis ... 251
6.3.2. Methods ... 253
6.3.2.1. Risk Prediction using weighted Genetic Risk Score ... 254
6.4. Results ... 254
6.4.1. Vincristine-induced peripheral neuropathy ... 254
6.4.2. Association Study ... 255
6.4.3. Combined Effect Model ... 256
6.4.4. Replication Analysis ... 257

6.4.5. Risk Prediction 257
6.5. Discussion 259
6.6. Conclusion 264
6.7. Summary Points 265
6.8. Acknowledgements 266
6.9. Authorship Contributions 267
6.10. Disclosure of Conflicts of Interest 267
6.11. Ethical conduct of research 267
6.12. References 268
6.13. Reference Annotations 273
6.14. Tables 275
6.15. Figure Legends 277
6.16. Figures 279
6.17. Supplemental Material 282

Section D 291

Chapter 7 291

7. General Discussion 292
7.1. Discussion of Section A 292
7.2. Discussion of Section B 294
7.2.1. MYBBP1A & Pancreatitis 295
7.2.2. IL16 & Pancreatitis 299
7.2.3. rs3809849, rs11556218 & Pancreatitis 301
7.2.4. Other associations worth discussing 303

7.2.4.1. rs3809849 and chemotherapy-induced osteonecrosis 303

7.2.4.2. rs11556218 and drug sensitivity, and ALL prognosis 306

7.3. Discussion of Section C 308

7.4. Limitations 311

7.5. Prospective Studies 316

7.5.1. Pancreatitis 316

7.5.2. Vincristine-induced peripheral neuropathy 317

7.6. Conclusion 319

7.7. References 321

Annex I I

Copyright Authorisations I

Annex II X

List of other scientific contributions X

Annex III XII

List of Scholarships & Prizes XII

IV. List of Figures

Chapter 1

Figure 1. Differentiation of hematopoietic stem cells. 3
Figure 2. The relative frequency of major B-ALLs and T-lineage subtypes of ALL. 9
....**Erreur ! Signet non défini.**
Figure 3. Visual illustration of Polymerase Chain Reaction (PCR)-based genotyping techniques. 35
Figure 4. Visual illustration of the workflow of the array-based exome-enrichment and whole-exome Sequencing approach. 40

Chapter 2

Figure 1. Metabolic pathways involved in the mechanism of action of thiopurines.....................96

Chapter 3

Figure 1. Mechanism of action of Asparginase. 133

Chapter 4

Figure 1. The selection process following the exome-wide association study............................172
Figure 2. Top-ranking EWAS signals common for several asparaginase-related toxicities.......173
Figure 3. Association of rs3809849 in MYBBP1A gene with ASNase-related toxicities and with event free- and overall survival. 174
Figure 4. Combined-effect model. 175
Figure 5. Replication analysis in the independent validation cohort. 176

Figure 6. Distribution of patients with pancreatitis among risk groups established using wGRS from the comprehensive combined-effect model in QcALL & DFCI cohort. 177

Supplemental figure S1. Comprehensive combined-effect model of all SNPs significantly associated with pancreatitis. 188

Supplemental figure S2. Combined-effect analysis in Combined-cohort for the 3 SNPs and 5 SNPs significantly associated with pancreatitis. 190

Unpublished Data Figure U1. Association of the genotype of rs3809849 polymorphism in the *MYBBP1A* gene with the risk of osteonecrosis during ALL treatment. 193

Unpublished Data Figure U2. Association of rs11556218 in IL16 gene with white blood cell count at presentation & event-free survival. 196

Unpublished Data Figure U3. Cell viability assay in lymphoblastoid cell-lines in relation to rs11556218 *IL16* gene polymorphism..198

Chapter 5

Figure 1. Production of Cas9 expressing, & MYBBP1A gene knock-out PANC1 cell lines and evaluation of cell proliferation capacity. 229

Figure 2. *In-vitro* sensitivity to asparaginase and vincristine in relation to *MYBBP1A gene Knock-out*. 230

Figure 3. Effect of MYBBP1A gene knock-out from PANC1 cells on their colony formation capacity and response to treatment with asparaginase. 231

Figure 4. Impact of MYBBP1A gene deletion on PANC1 cellular functions: a) Cell-Cycle and b) Apoptosis/Necrosis, and response to asparaginase exposure. 232

Figure 5. Relative expression of markers associated with epithelial-mesenchymal transition in PANC1 cells a) in response to MYBBP1A gene deletion and b) asparaginase exposure. 233

Supplemental Figure S1. Result of the mismatch cleavage assay performed on the cell population used for clonal selection. 238

Supplemental Figure S2. Relative protein expression levels of MYBBP1A protein. 239

Supplemental Figure S3. *In-vitro* sensitivity to asparaginase and vincristine in relation to *MYBBP1A gene Knock-out 96 hours post incubation with* a) asparaginase (ASNase) or b) vincristine (VCR). 240
Supplemental Figure S4. Association of MYBBP1A gene expression levels with survival probability in different types of cancer. 241

Chapter 6

Figure 1. Confirmatory step following the exome-wide association study. 279
Figure 2. Performance of the wGRS based combined-effect model in the discovery cohort and classification efficiency in both the discovery and replication cohorts. 280
Figure 3. Performance of the comprehensive combined-effect model in predicting the risk of VIPN and classifying patients into risk groups. 281
Supplemental Figure S1. Meta-analysis of the top-ranking associations combining both cohorts. 286
Supplemental Figure S2. Association between the top-ranking hits and the expression of their respective genes based on the genotype. 287

Chapter 7

Discussion Figure 1. Additive effect of carrying the minor alleles of rs11556218 in *IL16* gene and rs3809849 in *MYBBP1A* gene on the risk of pancreatitis. 302
Discussion Figure 2. Association of rs3803357 in *BAHD1* gene with Overall Survival in the a) QcALL and b) DFCI cohorts. 310

V. List of Tables

Chapter 1

Table 1. Summary of outcomes derived from most recent front-line trials for children and adolescents newly diagnosed with acute lymphoblastic leukemia. 13

Table 2. Asparaginase exposure in Dana Farber Cancer Institute protocols. 15

Table 3. Therapy on DFCI ALL Consortium Protocols: 1981–2000. 19

Table 4. Therapy on AIEOP-BFM ALL 2000 protocol. 22

Table 5. Class-specific and universal side-effects of chemotherapeutic agents. 27

Chapter 2

Table 1. Summary of selected studies which investigated the influence of TPMT pharmacogenetics on thiopurine treatment response in childhood acute lymphoblastic leukemia. 91

Table 2. Summary of selected studies which investigated the influence of TPMT pharmacogenetics on thiopurine treatment response in inflammatory bowel disease. 95

Chapter 3

Table 1. This table summarizes the prominent studies in the literature which investigated the pharmacogenetics of asparaginase and highlights the most important finding. 130

Chapter 4

Table 1. Characteristics of the discovery and the replication cohort. 164

Table 2. Top-ranking signals from the exome-wide association study confirmed by genotyping. 165
Table 3. Performance of the comprehensive genetic model in predicting the risk of pancreatitis. 167
Supplemental Table S1. Association of WES data with asparaginase-related complications...179
Supplemental Table S2. Function and differential protein expression of genes selected for further investigation. 183
Supplemental Table S3. Multivariate analysis for primary and secondary associations confirmed by genotyping 186
Supplemental Table S4. Combined-cohort analysis performed for SNPs with reproducible associations with pancreatitis. 187
Unpublished Data Table U1. Stratification analysis of the association of rs3809849 polymorphism in the MYBBP1A gene with the risk of osteonecrosis based in clinical subgroups in the combined-DFCI cohort. 194
Unpublished Data Table U2. Multi-variant analysis for the risk of osteonecrosis associated adjusting for rs3809849_MYBBP1A polymorphism genotype along with clinical factors. 195

Chapter 6

Table 1. Characteristics of the discovery and the replication cohorts. 275
Table 2. Top-ranking signals from the exome-wide association study confirmed by genotyping. 276
Supplemental Table S1. Association of WES data with vincristine-induced peripheral neuropathy. 283
Supplemental Table S2. Association of the risk alleles with the reduction in the VCR dose administered. 284
Supplemental Table S3. Association of the risk alleles with the number of episodes of high-grade VIPN. 285
Supplemental Table S4. Analysis of the association between the risk alleles and lower-grade (I/II), higher-grade (III/IV) and all-grades (I-IV) VIPN. 285

VI. Abbreviations and Acronyms

6-MP: 6-mercaptopurine

6-TGN: 6-Thioguanine Nucleotides

***ACTG1**: Actin Gamma 1*

***ADAMTS17**: A Disintegrin-Like And Metalloprotease (Reprolysin Type) With Thrombospondin Type 1 Motif, 17*

ADRs: Adverse Drug Reactions

***AhR**: Aryl Hydrocarbon Receptor*

AIEOP: Associazione Italiana di Ematologiae Oncologia Pediatrica (Italian Association of Pediatric Haematology and Oncology)

***AKT**: AKT Serine/Threonine Kinase 1*

ALL: Acute Lymphoblastic Leukemia

ASN: Asparagine

ASNase: Asparaginase

ASO: Allele-Specific Oligonucleotides

***ASS1**: ArginoSuccinate Synthase 1*

***ATF5**: Activating Transcription Factor 5*

AUC: Area Under the Curve

AZA: Azathioprine

***BAHD1**: Bromo adjacent homology domain containing protein 1*

***BCR-ABL1**: BCR (breakpoint cluster region) and ABL1 (Abelson) gene fusion*

BFM: Berlin-Frankfurt-Münster

BSA: Body Surface Area

Cas9: CRISPR Associated Protein 9

CI: Confidence Interval

CIPN: chemotherapy-induced peripheral neuropathy

CMT: Charcot-Marie-Tooth

CNS: Central Nervous System

COG: Children's Oncology Group

CPIC: Clinical Pharmacogenetics Implementation Consortium

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

CRLF2*: Cytokine Receptor Like Factor 2*

CTCAE: Common Terminology Criteria for Adverse Events

CYP*: Cytochrome P450*

DCOG: Dutch Childhood Oncology Group

DFCI: Dana-Farber Cancer Institute

DHFR*: Dihydrofolate Reductase*

E.coli*: Escherichia coli*

EFS: Event-free survival

EMT: Epithelial-Mesenchymal Transition

EORTC CLG: European Organization for Research and Treatment of Cancer-Children's Leukemia Group

eQTL: Expression Quantitative Trait Loci

Erwinia*: Erwinia chrysanthemi*

ETV6-RUNX1* (*TEL-AML1*): ETV6 (ETS Variant 6) & RUNX1 (Runt-Related Rranscription Factor 1) gene fusion*

EWAS: Exome-Wide Association Study

FDA: U.S. Food and Drug Administration

FDR: False Discovery Rate

GC: Glucocorticoid

GRIA1*: Glutamate Ionotropic Receptor AMPA Type Subunit 1*

HLA*: Human Leukocyte Antigen*

HR: High-Risk

HSCT: Hematopoietic Stem Cell Transplant

HSR: HyperSensitivity Reactions

IBD: Inflammatory Bowel Disease

IC50: Minimum Inhibitory Concentration 50

IKZF1*: IKAROS Family Zinc Finger 1*

IL16*: Interleukin 16*

iPSC: Induced Pluripotent Stem Cells

IR: Intermediate:Risk

ITPA*: Inosine Triphosphate Pyrophosphatase*

LCLs: Lymphoblastoid Cell Lines

MAF: Minor Allele Frequency

miRNAs: micro inhibitory RNAs

MLL*: mixed-lineage leukemia*

MPEG1*: Macrophage Expressed 1*

MRD: minimum residual disease

MRPL47*: Mitochondrial Ribosomal Protein L47*

MTHFR*: 5,10-Methylenetetrahydrofolate reductase*

MTX: methotrexate

MYBBP1*: MYB Binding Protein 1a*

NES: Normalized Effect Size

NFATC2*: Nuclear Factor of Activated T Cells 2*

NF-kB*: Nuclear Factor kappaB*

NGS: Next Generation Sequencing

NOPHO: Nordic Society of Paediatric Haematology and Oncology

NSAA: Nadir serum asparaginase activity.

ON: Osteonecrosis

OR: Odd-Ratio

OS: Overall Survival

PANC1: Pancreatic Cell Line 1

PAX5*: Paired Box 5 (B-Cell Lineage Specific Activator Protein)*

PBX1*: Pre-B-Cell Leukemia Homeobox 1*

PCR: Polymerase Chain Reaction

PEG ASNase: Pegylated Asparaginase

PGC-1α*: PPAR-gamma coactivator-1α*

PGx: Pharmacogenomics and Pharmacogenetics

Ph-like: Philadelphia Chromosome Like

Ph-positive: Philadelphia Chromosome Positive

PKD2L1*: Polycystin 2 Like 1*

PREP1*: Pbx Regulating Protein-1*

PRSS1/PRSS2*: Protease, Serine 1/2*

QcALL: Quebec Childhood Acute Lymphoblastic Leukemia

RCTs: Randomized Clinical Trials

RIN3*: Ras Interaction/Interference Protein 3*

ROC: Receiver Operator Characteristic

SEM: Standard Error of the Mean

SJCRH: St Jude Children's Research Hospital

SJUHC: Sainte-Justine University Hospital Centre

SLC39A12*: Solute Carrier Family 39 Member 12*

SLC7A1*: Solute Carrier Family 7 Member 13*

SMN: Second Myeloid Neoplasms

SNP: Single Nucleotide Polymorphism

SPECC1*: Sperm Antigen With Calponin Homology And Coiled-Coil Domains 1*

SPEF2*: Sperm Flagellar 2*

SR: Standard-Risk

SYNE2*: Spectrin repeats containing nuclear envelope 2*

TCF3-PBX1*: TCF3 (Transcription Factor 3) & PBX1 (PBX Homeobox 1) gene fusion*

***TPMT**: Thiopurine S-Methyl Transferase*

***TS**: Thymidylate Synthase*

UKALL: United Kingdom Acute Lymphoblastic Leukaemia.

VCR: Vincristine

VIPN: Vincristine-Induced Peripheral Neuropathy

WBC: White Blood Cell

WES: Whole-Exome Sequencing

wGRS: weighted Genetic Risk Score

WGS: Whole-Genome Sequencing

WT: Wild-Type

***YTHDC2**: YTH Domain Containing 2*

VII. Dedication

I dedicate this thesis to my wife, May.
Your love and support helped me go all the way.

& to my baby girl, Mila, the joy of my life.

Also, to my parents and my family, who, by being so untypical, always provoked me to give it all and to become the best version of myself.

VIII. Acknowledgement

It is my great pleasure to acknowledge all of those who helped me through this exciting journey and made this thesis possible.

I would like to start by sincerely thanking **Dr. Maja Krajinovic** for offering me this great opportunity and guiding me through all of it with patience, understanding and support. Always seeing my potential and pushing me to give my best, and for which, I will remain forever grateful.

Also, I wish to acknowledge the members of my thesis committee, **Dr. Isabelle Laverdière**, **Dr. Simon De Denu**s, and **Dr. R-Noël J-M Raynal**, for accepting their roles and taking the time to go through my work.

I would like to also deeply thank **Dr. Réné Cardinal** and **Dr. Yves Théôret** for being there for me throughout the entire journey, always believing in me and providing encouragement, guidance and extremely valuable recommendations.

I would also like to express my sincere gratitude to **Dr. France Varin**, who was the first faculty member to see my potential and who paved the way for the start of this journey, along with **Dr. Catherine Litalien**, who I also deeply thank.

It goes without saying that I thank all previous and current members of the Krajinovic's lab for the help, support and valuable friendship they always provided. I specifically thank **Vincent Gagné** for always being so kind and patient; **Aziz Rezgui** for his constant encouragement and support; and **Kateryna Petrykey** for offering me her friendship and for sharing all the stressful moments throughout our common journey.

I even extend my gratitude to all members of the Research Centre of the Sainte-Justine hospital, and particularly to members of the 6th and 7th floor, for always being kind and friendly, and providing essential help when I was struggling with my experiments or with learning new techniques.

Finally, I thank all the patients who accepted to participate in the studies and made this work possible, as well as the funding agencies and organizations that provided the financial means for doing it. I especially thank the Faculty of Medicine and the Faculty of Higher Education of the University of Montreal, the Cole Foundation and the Network of Applied Medical Genetics for their generous scholarships.

IX. Preface

The present thesis titled "Using Whole-Exome Sequencing Data in an Exome-Wide Association Study Approach to Identify Genetic Risk Factors Influencing Acute Lymphoblastic Leukemia Response: A Focus on Asparaginase Complications & Vincristine-Induced Peripheral Neuropathy" has been carried out by me under the guidance and supervision of Dr. Maja Krajinovic, and is submitted to the faculty of higher education at the University of Montreal in partial fulfillment of the requirements for the degree Doctor of Philosophy in Pharmacology (Pharmacogenomics option). This work is presented in the by-article format.

Being a practicing pharmacist, it has always been intriguing to me how the same drug administered in the same dose to different patients would result in a spectrum of effects that can range from complete absence of response all the way to severe life-threatening toxicities. This observation, combined with my passion about genetics, ignited my interest in conducting pharmacogenetics research that would help to advance our understanding of the genetic basis of variability in drug response. Therefore, during my four years of doctoral studies, I tried to get involved in different aspects of pharmacogenetics research ranging from reviewing and summarizing the available literature, to discovering and validating novel genetic markers, passing by fundamental and translational research to determine their usefulness and applicability, and ending by assessing the need for implementation of pharmacogenes in clinical practice.

It is worth mentioning here that I used childhood acute lymphoblastic leukemia (ALL) as a disease model to learn and apply pharmacogenetics techniques and to investigate the role of genetic variability in altering the drug response.

In the body of this thesis, in the first chapter of Section-A, I will provide a brief, but detailed, introduction covering the basic information essential for the understanding of the context of this work and the different notions and definitions that are discussed through it. In the second chapter, I present a review paper titled "*Thiopurine S-methyltransferase polymorphisms in acute lymphoblastic leukemia, inflammatory bowel disease and autoimmune disorders: influence on treatment response* ". This paper was published in 2017 in the *Pharmacogenomics and Personalized Medicine* journal and provides an overview of the history and temporal evolution of *TPMT* towards becoming one of the most important pharmacogenes in clinical practice. I discuss the results, conclusions and recommendations of selected studies that investigated the pharmacogenetics influence of *TPMT* gene on thiopurine treatment in ALL, inflammatory bowel disease and autoimmune disorders, and also briefly address the cost-effectiveness of this pharmacogenetics approach and its impact on clinical practice

In Section B, I present three articles that targeted different aspects of the pharmacogenetics of asparaginase (ASNase) as a key component of ALL treatment along with a special chapter containing results not presented in a paper format. The first article in this section is presented in Chapter-3 and is a review article titled "*Pharmacogenetics of Asparaginase in Acute Lymphoblastic Leukemia".* It was published in 2019, in the special issue of the *Cancer Drug Resistance* journal, titled "Pharmacogenetics of Cancer" and it

highlights the most important findings reported in studies of the pharmacogenetics of ASNase related complications and treatment outcome.

The second article of this section is an original research paper presented in Chapter-4 and titled "*Whole-exome sequencing identified genetic risk factors for asparaginase-related complications in childhood ALL patients".* It was published in *Oncotarget* journal in 2017 and describes the results obtained from using whole-exome sequencing (WES) data to perform exome-wide association studies (EWAS) with ASNase-related toxicities and highlights their interactions and pertinence to the studied outcome, with a special focus on acute pancreatitis. This work suggests that *MYBBP1A* gene as an important candidate in modulating ASNase response that is associated with increasing risk of developing all of the studied complications.

The third article of Section-B, presented in Chapter-5 and titled "*Characterization of the functional impact of MYBBP1A gene on asparaginase sensitivity and risk of pancreatitis following exome-wide association study results*" is an original research work currently in preparation. In this EWAS follow-up study, I aimed at confirming and characterizing the involvement of *MYBBP1A* gene in modulating the cellular response to ASNase by studying the effect of gene deletion in PANC1 pancreatic cells, using CRISPR-CAS9 technology, on cellular behaviour and biological functions before and after treatment with ASNase.

The next section, Section-C, only has one chapter, Chapter-6, represents an original research paper that was published in 2018 in *Pharmacogenomics* journal and is titled "*Genetic risk factors for VIPN in childhood acute lymphoblastic leukemia patients identified*

using whole-exome sequencing". This work was performed in a similar manner of the one described in Chapter-4, but was focused on identifying genetic variants involved in modulating the risk of vincristine-induced peripheral neuropathy; a common side effect to the administration of vincristine as an important chemotherapeutic agent in childhood ALL treatment.

The last chapter of this thesis, Chapter 7, is presented in Section D and it provides a summary of the major findings, as well as detailed discussion on the two most prominent genes in this work, MYBBP1A and IL16, and trying to address the different possible mechanisms that these gene could be exerting their effect on modulating the response to ASNase. It also discusses the limitations of the work and suggests future studies that can help to better understand the role of the identified genes in the respective toxicities.

Section A

Chapter 1

General Introduction

This chapter, as indicated in the title, is meant to prepare the readers to navigate through the following chapters of the thesis by providing the essential information relative to the diverse topics discussed in this work. It also outlines the working hypotheses that formed the basis of the research design, and defines the objectives that the conducted studies were aiming to achieve.

1. General Introduction

1.1. Definition & Statistics

Leukemia is a type of cancer that affects the hematopoietic precursors of the lymphoid lineage. Acute Lymphoblastic Leukemia (ALL) is a fast progressing leukemic malignancy which results from an abnormal transformation and proliferation of lymphoid progenitor cells in the bone marrow and the blood.[1-4] It is a result of the deregulated control of the blood stem cells which affects their ability to differentiate into healthy mature blood cells, thus affecting the number and functions of different blood components (i.e. red blood cells, white blood cells, and platelets) and consequently provoking a wide range of complications.[2,4] Figure.1 provides a quick outlook on blood cells development showing the differentiation of diverse lineages of blood and immune cells from a common blood stem cell, including T and B lymphocytes.[4]

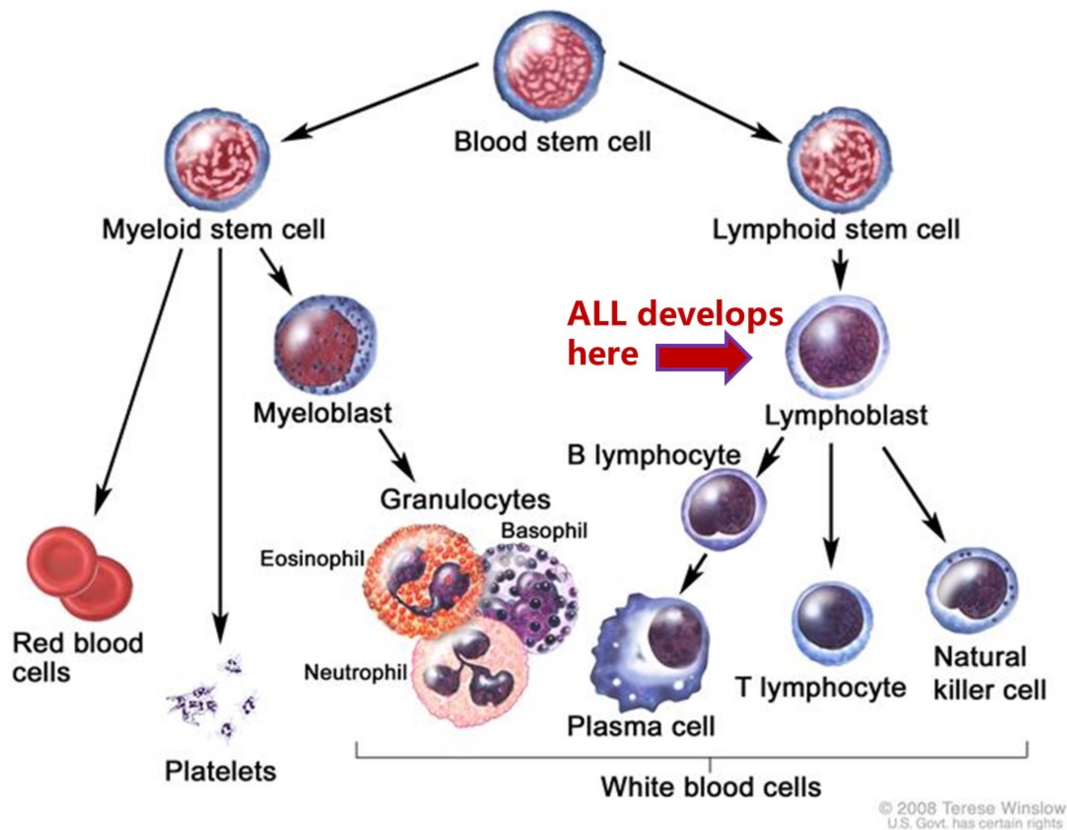


Figure 1. Differentiation of hematopoietic stem cells.

In the bone marrow, blood stem cells differentiate into either the myeloid or the lymphoid progenitor lines. Acute lymphoblastic leukemia is a result of aberrant differentiation of the lymphoid cells (B or T cell), leading to overproduction and accumulation in the blood, bone marrow, spleen and liver.

For the National Cancer Institute © 2008 Terese Winslow LLC, U.S. Govt. has certain rights. Reproduced with permission.

Generally, the presence of 20% lymphoblasts in the bone marrow or the blood is used as a cut-off to establish the ALL diagnosis.[1,5] Almost 80% of ALL cases occur in pediatric population and is referred to as childhood acute lymphoblastic leukemia.[3] The incidence can start as early as before birth, but there is a marked peak in between 1-5 years of age.[2,6] However, another peak can also be observed at around the age of 50, giving rise to adult acute lymphoblastic leukemia, which is usually associated with less favourable outcomes.[2,6] In fact, survival probability decreases with increasing patient's age at diagnosis, and sadly, the long-term survival rate among patients over 60 years of age is only about 10-15%.[1,7] Childhood ALL is the most common subtype of leukemia, accounting for approximately 25% of all childhood cancers and about 75-80% of leukemia cases in children.[2,8,9] Furthermore, it is the most frequent cause of death from cancer before 20 years of age.[6]

Genetics can play an important role in the incidence of ALL as it was shown that ethnicity is significantly associated with the risk of developing ALL; with black race individuals being the least affected, followed by those of the white race and then Hispanics having the highest incidence.[6] Moreover, in the same genetic context, male gender was found to be associated with a slightly higher, but significantly different, risk of childhood ALL than female gender (55% to 45%, respectively).[6] This inherent vulnerability of male gender is not surprising since it has been previously pointed out that the variability in epigenetic signature between genders, and the differential ability of the Y vs. X chromosomes in repairing damage to their genes, can render boys at increased risk of

developing various types of health conditions including different cancers.[6,8] For example, a recent study reported that the gene coding for the histone demethylase Ubiquitously Transcribed X-chromosome (UTX) tetratricopeptide repeat protein was found to be recurrently affected by somatic loss-of-function mutations in male T-cell acute lymphoblastic leukemia (T-ALL) patients and that *UTX* is capable of escaping X-inactivation in female T-ALL blasts as well as in normal T cells; thus adding to the growing body of evidence suggesting that *UTX* has a gender-specific tumor suppressor role in the context of T-ALL, among other cancers.[10]

1.2. Prognostic Factors

Classically, childhood ALL was majorly stratified into risk groups based on two important clinical factors, age and white blood cell counts at presentation.[1] However, it is largely recognized nowadays that in addition to clinical features at diagnosis, immunophenotype, pathophysiology and cytogenetic changes of cancer cells, genetics of the host, as well as response to initial treatment (also known as early response), can all interact together to affect the risk and prognosis of childhood ALL and should be used collectively to guide treatment regimens.[1]

1.2.1. Age & WBC count

Briefly, older age and higher WBC count are associated with a worsening prognosis and two groups of risk can be defined based on these parameters according to the Consensus criteria of the Rome/National Cancer Institute Workshop:[11] "standard risk" (1 > age < 10 years and initial WBC count of <50,000 per cubic millimeter) representing around two thirds of patients, and "high risk" (age ≥10 years, initial WBC count ≥50,000 per cubic millimeter, or both)[12] which roughly makes one third of patients. It must be noted that ALL in children < 1 year of age at diagnosis is usually associated with a worse outcome and is considered a special subgroup.[6,13,14]

1.2.2. Immunophenotype

Immunophenotyping based on the expression of the surface markers of lineage can distinguish between two subtypes of childhood ALL known as precursor B-cell and T-cell, making reference to the otherwise healthy mature lymphocytes expressing these markers, and representing around 85% and 15% of childhood ALL cases, respectively. This is important to understand the distinction between immunophenotypes since it was shown that age and WBC count at diagnosis have limited prognostic importance in T-cell ALL.[6,13]

1.2.3. *Cytogenetics & Molecular Genetics*

While several factors have been reported to predispose to an increased risk of developing childhood ALL including exposure to ionizing radiation, chemicals such as pesticides & certain solvents, viral infections like Epstein-Barr virus or human immunodeficiency virus, these factors can only explain a minor percentage of cases.[1,6,15,16]

Differences in the genetic make-up between patients have recently driven considerable attention as genetic variability and chromosomal aberrations have been described as early, probably initiating events, in developing ALL, and were shown to play an important role in disease detection, prognosis and treatment response.[1,6,17,18] Common genetic alterations include single nucleotide polymorphisms (SNPs), genomic insertions and deletions, as well as copy number variation.[19] These variants can be divided into disease-causing variants -with high penetrance and a large pathogenic effect- which are usually rare and mostly seen in single-gene Mendelian disorders, or can have lower penetrance and smaller pathogenic effect -typically present in higher frequency in cases compared to controls in association studies.[20] For instance, genes governing B-lymphoid development have been associated with ALL, most notably *PAX5* gene, which was estimated to be mutated in 35% of childhood ALL patients [21] followed by *IKZF1* gene reportedly mutated in 15% of cases.[22] Several association studies identified polymorphic variants in various other genes to be linked to an increased risk of ALL or to specific subtypes of it such as variants in *CEBPE*, *GATA3* and *ARID5B* genes.[23-25] Likewise, copy number variation within

genes involved in B cell proliferation and differentiation is a very frequent event observed in B-cell ALL patients.[18,21]

Tumor-specific genetic alterations can include inter-chromosomal translocations, uniparental disomy, and loss of heterozygosity. For example, loss of heterozygosity in an allele of tumor suppressor gene can results in tumorigenesis and may also influence drug effects thus modulating the evolution of the disease and its progression.[18,19,26] Several genetic translocations were extensively described in childhood ALL such as: t(12;21) [*ETV6-RUNX1*] gene fusion reported in around one quarter of cases; t(9;22) [*BCR-ABL1*] that results in the formation of an activated tyrosine-kinase and is also known as the Philadelphia chromosome (Ph-positive) ALL; and the translocation of t(1;19) [*TCF3-PBX1*] whose protein product alters cell differentiation arrest mechanisms among others. Additionally, multiple genomic rearrangement of the *CRLF2* gene [6,27] as well as more than 70 different chromosomal rearrangements involving the chromosome 11q23 mixed-lineage leukemia (*MLL*) gene,[1,21,28] have been described in ALL literature. Recently, a new subtype of ALL, characterized by exhibiting a gene expression profile similar to that of the Philadelphia chromosome but lacking the *BCR-ABL1* rearrangement, has been identified, and is also known as Philadelphia (Ph)-like ALL (or previously as *BCR-ABL*-like ALL). Interestingly, 90% of Ph-like cases seem to harbor a plethora of genetic alterations lading to kinase-activation.[1,6] The relative frequency of genetic alterations found in major B-ALLs and T-lineage subtypes of ALL as derived from front-line studies of childhood ALL are shown in Figure.2.

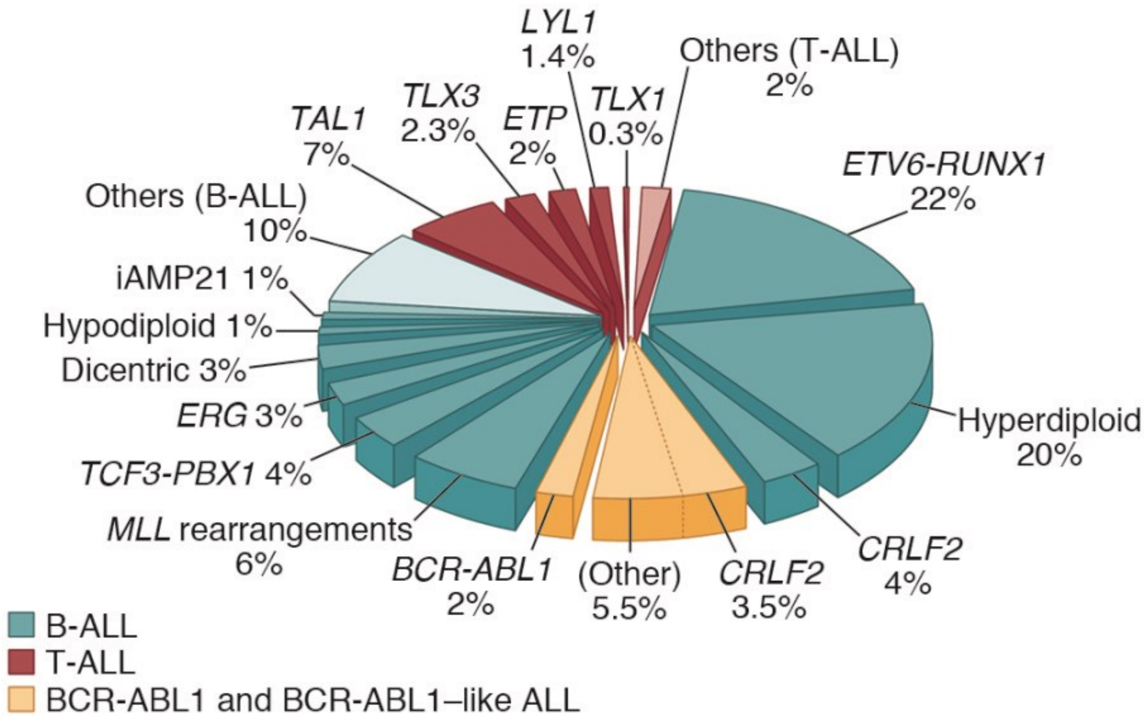


Figure 2. The relative frequency of major B-ALLs and T-lineage subtypes of ALL.

BCR-ABL1–like subtype and BCR-ABL1–positive ALL are shown in yellow to illustrate the high frequency of childhood B-ALL cases with genetic alterations activating tyrosine kinase and cytokine receptor signaling. Data are derived from front-line studies of childhood ALL.

Reproduced with permission from (Mullighan CG. Molecular genetics of B-precursor acute lymphoblastic leukemia. J Clin Invest 122(10), 3407-3415 (2012)). American Society for Clinical Investigation.

Interestingly, it was shown that genetic background variability related to race can be associated with differential risk of developing particular subtypes of ALL such as *TCF3-PBX1* ALL in Blacks [12] and *CRLF2*-rearrangement ALL in Hispanics.[27] Moreover, numerous genetic syndromes have also been associated with a higher risk of developing ALL in children, most notably being Down syndrome and Fanconi anemia, but ataxia telangiectasia Neurofibromatosis, Bloom syndrome, Li-Fraumeni syndrome and Nijmegen breakdown syndrome were also reported.[1,4,6,29-32]

It is highly important to have a detailed characterization of the patient's ALL subtype as certain genetic alterations can have prognostic utility since they were shown to be associated with treatment outcome of childhood ALL.[1,6,33] For example, high-risk of a poor outcome has been consistently reported for patient with intra-chromosomal amplification of chromosome 21,[34] BCR-ABL1 gene fusion,[35] Ph-like subtype of ALL,[1,36] MLL rearrangement,[37] and alterations of IKZF1;[38,39] as well as for patients showing hypodiploidy with less than 44 chromosomes,[40] and those with T-cell precursor ALL subtype.[41,42] On the other hand, ETV6-RUNX1 translocation and high hyperdiploidy are associated with favourable outcome.[6]

Moreover, variability in epigenetic signature, such as an aberrant acetylation or methylation profile, can modulate genetic expression, thereby influencing drug effect, and is a common feature of cancer cells.[19,43,44] Furthermore, it is increasingly recognized that even genomic regions that do not codify proteins such as micro inhibitory RNAs (miRNAs), which are RNA sequences that are around 22 nucleotides in size, can be strongly implicated in

regulatory functions as they can modulate the expression of over 60% of known genes, thus influencing sensitivity to drugs and treatment outcome.[45] Indeed, some miRNA-related polymorphisms have been shown to affect miRNA levels and function, and the expression of some of those miRNAs has been associated with drug response in ALL treatment.[45-49]

1.2.4. Early Treatment Response:

Recently, early response to treatment (also referred to as response to the initial therapy) has received a lot of attention and emerged as an important and independent prognostic tool in ALL treatment. The determination of the phenotype (i.e. type of response) is based on the evaluation of the time required to bring down the initial leukemic-cell population to undetectable levels, known as minimum residual disease (MRD).[1,6,13] This method uses molecular techniques such as the polymerase chain reaction or flow cytometry to monitor the disease at submicroscopic levels, which helps further refining the risk-stratification process at different stages of therapy, consequently improving the treatment outcome.[50-54]

1.3. ALL Treatment and Outcome

1.3.1. Improvement in Treatment Outcome

The first temporary remission of leukemia induced by chemotherapy was reported around 7 decades ago, in 1948.[55] In the 1960s, the survival rate of childhood ALL was estimated to be less than 10%.[6,56] Nowadays, the 5 years event-free survival (EFS) and overall survival (OS) rates are reported to surpass 85% and 90%, respectively, for most international treatment protocols;[57-63] thus making childhood ALL an exemplary model for progressive improvement.[1,6,8,56] Table.1 provides a short summary of outcomes derived from most recent front-line trials for children and adolescents newly diagnosed with ALL. Similar improvement was also reported for 10-year survival which has witnessed an increase of more than 20 percentage points in the last three decades in patients aged 0–14 years, which is being considered recently as a new cut-off value for age-based risk stratification in childhood ALL. [8]

Table 1. Summary of outcomes derived from most recent front-line trials for children and adolescents newly diagnosed with acute lymphoblastic leukemia.

Reference	Study Protocol	Years	No. Of Patients	EFS§ (%)	OS (%)
Vrooman et al.	DFCI Protocol 00-01	2000-2004	492	80.0	91
Conter et al. Schrappe et al.	AIEOP-BFM ALL2000	2000-2006	4480	80.3	91.1
Hunger et al.	COG	2000-2005	6994	N/A	91.3
Pui et al.	SJCRH Total Therapy Study XV	2000-2007	498	85.6	93.5
Vora et al	UK ALL 2003	2003-2011	3126	87.2	95.1
Veerman et al.	DCOG ALL-9	1997-2004	859	81	86
Domenech et al.	EORTC CLG 58591	1998-2008	1940	82.6	89.7
Schmiegelow et al.	NOPHO ALL-2000	2000-2007	1023	79	89

AIEOP-BFM, denotes Italian Association of Pediatric Haematology and Oncology and Berlin-Frankfurt-Münster; ALL, Acute Lymphoblastic Leukemia; COG, Children's Oncology Group; DCOG, Dutch Childhood Oncology Group; DFCI, Dana Farber Cancer Institute Consortium; EORTC CLG, European Organization for Research and Treatment of Cancer-Children's Leukemia Group; EFS, Event-free survival; OS, Overall Survival; NOPHO Nordic Society of Paediatric Haematology and Oncology; SJCRH, St Jude Children's Research Hospital; UKALL, United Kingdom Acute Lymphoblastic Leukaemia.

§ Survival percentages shown are the rates at 5 years except for the rates for the AIEOP-BFM trial, which were reported at 7 years.

Reproduced with permission from (Hunger SP, Mullighan CG. Acute Lymphoblastic Leukemia in Children. N Engl J Med 373(16), 1541-1552 (2015)), Copyright Massachusetts Medical Society.[6]

This was achieved through the introduction and continuous refining of multi-agent chemotherapeutic regimens, paired with the progressive advancement in risk-stratification based on clinical features of the patients, a better understanding of the biological mechanisms underlying the disease, the ability to exploit genetic differences between cancer-cells and host-cells, as well as the incorporation of the initial treatment response as a dynamic parameter into the risk-calculation equation and the adoption of precision-medicine treatment strategies.[6,13] Table 2 provides a brief comparison of ASNase exposure between four consequent treatment protocols of the Dana Farber Cancer Institute Consortium to highlight the evolution of the use of ASNase which is a main focus of this thesis and will be discussed in further details throughout different chapters.

Table 2. Asparaginase exposure in Dana Farber Cancer Institute protocols.

Asparaginase Dose Evolution	
Induction (4 weeks)	
Protocol 87-01	*E. coli*, *Erwinia* or PEG ASNase × 1 dose (randomized; investigational window; 5 days pre-day 0)
Protocol 91-01	None
Protocol 95-01	*E. coli* or *Erwinia* ASNase 25 000 IU/m^2 × 1 dose (randomized; day 4)
Protocol 00-01	*E. coli* ASNase 25,000 IU/m2 IM × 1 dose
Intensification (20–30 weeks) every 3-week cycle	
Protocol 87-01	*E. coli* ASNase 25 000 IU/m^2 weekly
Protocol 91-01	Randomized to *E. coli* ASNase 25 000 IU/m^2 weekly or PEG ASNase 2500 IU/m^2 every 2 weeks
Protocol 95-01	Randomized to *E. coli* ASNase 25 000 IU/m^2 weekly or *Erwinia* ASNase 25 000 IU/m^2 weekly
Protocol 00-01	Randomized to fixed dosing of *E. coli* ASNase (based upon BSA) or individualized dosing (based upon NSAA every 3 weeks)

Abbreviations: ASNase, asparaginase; PEG, pegylated; BSA, body surface area; NSAA, nadir serum asparaginase activity.

Reproduced from author's own article (Wolthers BO, Frandsen TL, Patel CJ et al. Trypsin encoding PRSS1-PRSS2 variation influence the risk of asparaginase-associated pancreatitis in children with acute lymphoblastic leukemia: a Ponte di Legno toxicity working group report. Haematologica doi:10.3324/haematol.2018.199356 (2018)), Copyright European Hematology Association.[64]

Nevertheless, while the landscape looks promising for childhood ALL, it is important to note that the prognosis for adulthood ALL is still dismal, with almost half of the patients failing to achieve long-term remission, up until recently.[1,65,66] Encouragingly, it has been suggested lately that using pediatric-inspired protocols may be helpful in increasing survival of the adolescent and young adults population (i.e. 15-39 years), with some preliminary results showing a 5-year EFS of as high as 72%.[67-69] Unfortunately, however, for infants that are less than 1 year of age, the survival remains low despite the ongoing efforts aiming at improving it. One possible contributing factor to this poor prognosis is the fact that infant ALL is usually associated with *MLL* gene rearrangement, which, on its own, is associated with unfavourable outcomes, and any further intensification of chemotherapy can cause significant long-term and short-term toxicities in this vulnerable population.[8,70]

1.3.2. Contemporary Therapy

A major milestone in anti-leukemia treatment was the introduction of an intensive regimen that employed sets of combinations of 8 drugs administered over two phases (induction and consolidation) for a period of 8 weeks. This treatment strategy was later referred to as protocol-I and became the backbone of most contemporary protocols for ALL.[6] Indeed, modern treatment strategies for childhood ALL last 2–2.5 years in total and include distinct phases each of them having a specific objective. Table 3 summarizes the evolutionary history of the Dana Farber Cancer Institute Consortium (DFCI) protocols and Table 4 provides details on the Associazione Italiana Ematologia Oncologia Pediatrica and

the Berlin-Frankfurt-Munster study protocol AIEOP-BFM ALL 2000. Both of these protocols will be discussed extensively throughout this thesis.

The initial phase is essentially a remission induction therapy that usually lasts 4 to 6 weeks and includes an L-asparaginase (ASNase) formulation, a glucocorticoid (e.g. prednisone or dexamethasone; GCs) and vincristine (VCR), as wells as the optional use of an anthracycline. By the end of this phase, remission is successfully induced in most patients (85-95%), but relapse is still possible due to the submicroscopic residual disease. To reduce this risk and prevent the development of overt CNS leukemia, patients undergo a remission consolidation phase which includes 6 to 9 months of intensive combination chemotherapy. In general, high dose methotrexate (MTX) along with 6-mercaptopurine (6-MP) are commonly used in this phase, accompanied by frequent pulses of VCR, GCs and ASNase for 20–30 weeks. Basically, the drug combinations in this phase tend to include chemotherapeutic agents that have different mechanisms of action from those applied in the induction phase and might also include cytarabine, etoposide, and cyclophosphamide. This concept is important in order to minimize drug resistance and assure the elimination of submicroscopic residual disease by taking advantage of the synergistic effects obtained by combining the different molecules. The last phase is primarily a maintenance therapy and can last between 18 and 30 months depending on the protocol and the risk group. This is a low-intensity antimetabolite-based treatment comprising a daily oral 6-MP or thioguanine and a weekly oral MTX administered along with optional periodic pulses of glucocorticoids and vincristine every 5 to 7 days in certain protocols.[6,13]

Historically, cranial radiation was routinely employed in many protocols to further prevent CNS relapse, but its use was gradually abandoned (or reserved only for patients with the highest risk) due to its associated toxicities such as the risk of developing a second malignant neoplasms and the concerns about its long-term effects on cognitive skills leading to intellectual disability, especially in young adults. Instead, it was replaced by intrathecal therapy that was incorporated into the induction remission phase of most protocols and which includes the administration of intrathecal methotrexate, either alone, or in combination with cytarabine and hydrocortisone (referred to as triple intrathecal treatment). However, the administration of this therapy in other phases is variable across the different protocols, with some of them also administering it during the remission consolidation phase while others throughout the entire course of treatment.[6,13,56]

One of the hallmarks of childhood ALL treatment is the stratification of patients into risk groups. While the definition and treatment of high-risk childhood ALL remains controversial, the use of prognostic factors affecting the treatment outcome can allow the classification of patients into groups based on their risk of experiencing treatment failure. Protocols offer different blocks of chemotherapy with varying intensities and patients are then assigned to one of these blocks depending on their risk-stratified group. Accordingly, patients with favorable prognostic features can be treated with less toxic regimens while those at high-risk of failure or relapse can be assigned to receive more intense regimens to help eradicating the highly aggressive disease.[6,13]

Table 3. Therapy on DFCI ALL Consortium Protocols: 1981–2000.

Phase	*Treatment*
Induction (4 weeks)	**IT cytarabine*** × 1 dose (day 0), **IT chemotherapy** day 14
	Vincristine 1.5 mg/m^2 q week (maximum=2 mg) (days 0, 7, 14, 21)
	Prednisone 40 mg/m^2/day (days 0–28)
	Doxorubicin 30 mg/m^2/dose (days 0 and 1)
	Protocol 81-01: 45 mg/m^2/dose x 1 dose
	Protocol 95-01: randomized +/− **dexrazoxane** 300 mg/m^2 (HR only)
	Protocol 00-01: + **dexrazoxane** 300 mg/m^2 (HR only)
	Methotrexate × 1 dose (day 2): dose per protocol
	Protocol 81-01: None
	Protocol 85-01: 40 mg/m^2
	Protocol 87-01: 40 mg/m2 or 4 g/m^2 with **leucovorin** (randomized)
	Protocols 91-01 + 95-01 + 00-01: 4 g/m^2 with leucovorin
	Asparaginase
	Protocol 81-01: None
	Protocol 85-01: E.coli ASNase × 1 dose (investigational window; 5 days pre-day 0)
	Protocol 87-01: E. coli, Erwinia or PEG ASNase × 1 dose (randomized; investigational window; 5days pre-day 0)
	Protocol 91-01: None
	Protocols 95-01: E.coli or Erwinia ASNase 25,000 IU/m2 × 1 dose (randomized; day 4)
	Protocols 00-01: E.coli ASNase 25,000 IU/m^2 × 1 dose (randomized; day 4)
CNS therapy (3 weeks)	**IT chemotherapy** twice weekly × 2 weeks (4 doses)
	Vincristine 2.0 mg/m^2 IV day 1 (maximum=2 mg)
	6-MP 50 mg/m^2/day orally (days 1–15)
	HR only: **doxorubicin** 30 mg/m^2 on day 1
	Protocol 95-01: randomized +/− **dexrazoxane** 300 mg/m^2
	Cranial Radiation per protocol (beginning day 1)
	Protocol 81-01: SR-18Gy; HR-28 Gy
	Protocol 85-01: SR-18Gy; HR-24 Gy
	Protocol 87-01: SR-No XRT; HR-18 Gy
	Protocol 91-01: SR girls-No XRT; SR boys and HR-18 Gy.
	Protocol 95-01: SR: randomized-No XRT versus 18 Gy; HR-18 Gy
	Protocol 00-01: SR: No XRT; HR without CNS 12 Gy; HR with CNS 18 Gy

Intensification (20–30 weeks)	**Every 3 week cycles:**	
	SR	**IT chemotherapy**
		Vincristine 2.0 mg/m^2 IV day 1 (maximum = 2 mg)
		Prednisone 40 mg/m^2/day orally (days 1–5)
		Protocol 91-01: **dexamethasone** 6 mg/m^2/day instead of prednisone
		Protocol 00-01: randomized to **dexamethasone** or **prednisone**
		Methotrexate 30 mg/m^2 IV or IM days 1, 8, 15
		6-MP 50 mg/m^2/day orally days 1–15
		Protocol 91-01: randomized oral 6-MP vs IV 6-MP 1000 mg/m^2 on days 1 and 8 of each cycle for first 12 months of treatment
		Asparaginase IM according to protocol:
		Protocols 81-01 + 85-01 + 87-01: E.coli ASNase 25,000 IU/m^2 weekly
		Protocol 91-01: randomized E.coli ASNase 25,000 IU/m^2 weekly or PEG ASNase 2500 IU/m^2 every 2-weeks
		Protocol 95-01 : randomized E.coli ASNase 25,000 IU/m^2 weekly or Erwinia ASNase 25000 IU/m^2 weekly
		Protocol 00-01: randomized E.coli ASNase fixed dosing 25,000 IU/m2 weekly or individualized dosing starting at 12,500 IU/m2 weekly and adjusted every 3 weeks based on the nadir serum asparaginase activity levels
	HR *(same as SR patients except)*	**Prednisone** higher dose (120 mg/m^2/day orally days 1–5)
		Protocol 91-01: **dexamethasone** 18 mg/m^2/day instead of prednisone
		Methotrexate None
		Doxorubicin 30 mg/m^2 at day 1 of each cycle (maximum cumulative dose per cycle: 345 mg/m^2 in *81-01*, 360 mg/m^2 in *91-01,* 300 mg/m^2 in *95-01 + 00-01*)
		Protocol 95-01 (randomized): +/− **dexrazoxane** 300 mg/m^2
		Protocol 00-01: + **dexrazoxane** 300 mg/m^2
Continuation (until 24 months CCR)	**Every 3 week cycles:**	
	IT chemotherapy	
	SR: same as intensification, except **no asparaginase**	
	HR: same as SR patients	

Abbreviations: 6MP, 6-mercaptopurine; ASNase, asparaginase; CCR, continuous complete remission; CNS, central nervous system; DFCI, Dana-Farber Cancer Institute; HR, high risk; IM, intramuscular; IT intrathecal; IV, intravenous; SR, standard risk; XRT, radiotherapy.

**IT cytarabine was dosed according to age. Patients with CNS leukemia at diagnoses (CNS-2 and CNS-3) received twice weekly doses of IT cytarabine until the cerebrospinal fluid was clear of blast cells on three consecutive examinations.*

Reproduced with permission from (L B Silverman, K E Stevenson, J E O'Brien, B L Asselin, R D Barr et al. Long-term results of Dana-Farber Cancer Institute ALL Consortium protocols for children with newly diagnosed acute lymphoblastic leukemia (1985–2000). Leukemia. 2010 Feb; 24(2): 320–334. Copyright Springer Nature.[61]

Table 4. Therapy on AIEOP-BFM ALL 2000 protocol.

Drug/Administration route	mg/m²/day	Day
Prephase		
Prednisone/po-iv	60	1-7
Methotrexate/it	By age**	1
Induction: Protocol IA		
Vincristine/iv	1.5 (max: 2 mg)	8, 15, 22, 29
Prednisone/po-iv (®*)	60	8-28 then tapered
or Dexamethasone/po or iv (®*)	10	8-28 then tapered
Daunorubicin/iv	30	8, 15, 22, 29
l-Asparaginase/iv§	5,000 IU/m²	12, 15, 18, 21, 24, 27, 30, 33
Methotrexate/i.t.§	By age**	1, 12, 33°
Consolidation: Protocol IB		
Cyclophosphamide/iv	1,000	36, 64
Mercaptopurine/po	60	36-63
Cytarabine/iv-sc	75	38-41, 45-48, 52-55, 59-62
Methotrexate/i.t.§	By age**	45, 59
Protocol M		
Mercaptopurine/po	25	1-56
Methotrexate/iv	5,000§	8, 22, 36, 50
Methotrexate/i.t.	By age**	8, 22, 36, 50
Leucovorin rescue§	15 mg/m²/dose	42, 48, 54 h after start HD-MTX
Reinductioninduction (Protocol II)		
Dexamethasone/po-iv§	10	1-21 then tapered
Vincristine/iv	1.5 (max: 2 mg)	8, 15, 22, 29
Doxorubicin/iv§	30	8, 15, 22, 29
l-Asparaginase/iv§	10,000 IU/m²	8, 11, 15, 18
6-Thioguanine/po	60	36-49
Cyclophosphamide/iv	1,000	36
Cytarabine/iv-sc	75	38-41, 45-47
Methotrexate/i.t.	By age**	38, 45°
Cranial irradiation^	By age	
Reinductioninduction (Protocol III)		
Dexamethasone/po-iv	10	1-14 then tapered
Vincristine/iv	1.5(max: 2 mg)	1, 8
Doxorubicin/iv	30	1, 8
l-Asparaginase/iv§	10,000 IU/m²	1, 4, 8, 11
6-Thioguanine/po	60	15-28
Cyclophosphamide/iv	500	15
Cytarabine/iv-sc	75	17-20, 24-27
Methotrexate/i.t.	By age**	17, 24°
Cranial irradiation^	By age	
Interim Maintenance		
Mercaptopurine/po	50	Daily
Methotrexate/po	20	Weekly
HR Block 1		
Dexamethasone/po-iv	20	1-5
Vincristine/iv	1.5 (max: 2 mg)	1, 6
HD Cytarabine/iv	2,000 x 2	5
Methotrexate/iv	5,000	1
Leucovorin rescue§	15 mg/m²/dose	42, 48, 54h after start HD-MTX
Cyclophosphamide/iv	200 (q12h x 5)	2-4
l-Asparaginase/iv§	25,000 IU/m² §	6, 11
MTX/Ara-C/Pred/i.t.§	By age***	1
HR Block 2		
Dexamethasone/po-iv	20	1-5
Vindesine/iv	3	1, 6
Daunorubicin/iv	30	5
Methotrexate/iv	5,000	1
Leucovorin rescue§	15 mg/m²/dose	42, 48, 54h after start HD-MTX
Ifosphamide/iv	800 (q12h x 5)	2-4
l-Asparaginase/iv§	25,000 IU/m² §	6, 11
MTX/Ara-C/Pred/i.t.§	By age***	1°
HR Block 3		
Dexamethasone/po-iv	20	1-5
HD Cytarabine/iv	2,000 (q12h x 4)	1, 2
Etoposide/iv	100 (q12 h x 5)	3-5
l-Asparaginase/iv-im	25,000 IU/m²§	6, 11
MTX/Ara-C/Pred/i.t.§	By age***	5
Continuation phase§		
Mercaptopurine/po	50†	Daily
Methotrexate/po	20†	Weekly

Abbreviations: im, intramuscular; iv, intravenous; po, per os; i.t., intrathecal; sc: subcutaneous
®* According to first randomization
**Age-adjusted doses of intrathecal methotrexate: ≥1 and < 2 years: 8 mg; ≥2 and < 3 years: 10 mg; ≥3 years: 12 mg.
*** Age-adjusted doses of triple intrathecal MTX, ARA-C and methylprednisolone, respectively: ≥ 1 and < 2 years: 8, 20 and 6 mg; ≥ 2 and < 3 years: 10, 26 and 8 mg; ≥ 3 years: 12, 30 and 10 mg.
° Patients with initial CNS involvement receive additional i.t. therapy: on day 18 and 27 during Protocol IA, on day 1 and 18 during Reinductioninduction Protocol II, on day 1 during Reinductioninduction Protocol III, on day 5 during HR Block 2.
^ Cranial irradiation (CRT) was given at the following dosage: 1-2 years: 12 Gy (also in case of CNS involvement at diagnosis); age ≥2 years: 12 Gy (preventive) or 18 Gy (therapeutic for CNS involvement at diagnosis).
†Doses were adjusted to WBC (target range 2000-3000/µl).

§Differences of AIEOP protocol:

i.t. MTX during Induction Protocol IA: day 1, 15, 29

i.t. MTX during Consolidation Protocol IB: day 38, 52

Patients with initial CNS involvement receive additional i.t. therapy: on day 8 and 22 during Induction Protocol IA

l-Asparaginase given i.m.

High Dose Methotrexate: 5,000 mg/m2 only in patients with T-ALL or CNS/testicular involvement at diagnosis; 2,000 mg/m2 in all other patients.

Leucovorin Rescue: 7.5 mg/m2/dose for levorotatory compound (instead of 15 mg/m2 of the racemic compound used in BFM group) given at hours 42 and 48 for HD-MTX 2 g/m2; at hours 42, 48 and 54 for HD-MTX 5 g/m2.

Cranial radiotherapy was administered at the following dosage: age 1-2 years: 12 Gy (preventive) or 18 Gy (therapeutic for CNS involvement at diagnosis); age ≥2 years: 18 Gy (preventive) or 24 Gy (therapeutic for CNS involvement at diagnosis).

HR blocks: l-Asparaginase given at 10,000 IU/m2 im at day 6 only

HR blocks: i.t. MTX at day 1

Doxorubicin in Protocol II in HR patients: 25 mg/m2

Dexamethasone in Protocol II in HR patients, age ≥10 years: 10 mg/m2 days 1-7 and 15-21

Reproduced with permission from (Conter V, Bartram CR, Valsecchi MG, et al. Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study. Blood. 2010;115(16):3206-3214). American Society of Hematology.[72]

1.3.3. Treatment of Refractory/Relapsed ALL

While most children with ALL are cured, certain subsets are at high risk of relapse. It is generally known that cure rates drop significantly following relapse, which can affect between 15 and 20% of patients with childhood ALL.[1,6,13,73] Even with the introduction of intensified cytotoxic chemotherapy and allogeneic hematopoietic stem cell transplant (HSCT), overall survival from relapsed ALL is barely approaching 40%.[1,6,13]

The length of the time-period a patient spends in first complete remission (i.e. time to relapse), the site of relapse, and the immunophenotype of leukemic cells have all been linked to the prognosis of the disease; with shorter times to relapse, T-cell ALL phenotype, and bone marrow disease pondering a worsening prognosis. Moreover, cells from a relapsed disease tend to have a more resistant profile compared to cells from the original disease. This can be partly explained by the fact that ALL is frequently a polyclonal disease and that genetic alterations in sub-clones might allow them to escape the initial treatment and repopulate the host with more aggressive and highly resistant leukemic cells that already survived the selective pressure.[1,6,13,74]

Allogeneic HSCT is considered for patients at a very high risk of relapse and/or treatment failure and studies have shown that it is best to undergo the transplant after achieving MRD-negative disease status. Candidate patients include those showing hypodiploidy or those who already experienced an induction failure.[1,6,13]

1.3.4. Targeted Therapy and Precision Medicine

It can be argued that the ground-breaking improvement that childhood ALL treatment witnessed over the past few decades partially stemmed from the constant enrichment of knowledge on the effects of existing therapies at the molecular level, leading to their more effective use through better dosing and scheduling of drug combinations, rather than the introduction of new chemotherapeutics. Another milestone that marks the advancement is the breakthrough discoveries of the genetic basis of ALL which offered the possibility of applying a personalised treatment approach tailored to the genetic make-up of individual patients. This has paved the way for the incorporation of pharmacogenetics as a powerful tool for the application of precision-medicine.[75-77] One notable example showing the importance of understanding the differences among drugs and ALL subtypes at a molecular level is the use of tyrosine kinase inhibitors such as imatinib in patients with Ph-positive ALL expressing the BCR-ABL1 fusion protein. This chimeric protein can be seen in 2 to 5% of children with ALL and in as high as 25% in patients with adulthood ALL. Its presence was shown to be associated with poor prognosis and high risk of relapse in multiple studies. In fact, before the incorporation of imatinib into treatment protocols of Ph-positive childhood ALL patients, less than half of the children survived.[6,35] The use of imatinib in this subgroup of childhood ALL patients, in combination with other chemotherapeutic agents, revolutionized the treatment of this high-risk group and significantly increased the 3-years event-free survival rates from 35% to 80% , while also reducing the number of patients requiring HSCT in the first remission.[70,78]

1.3.5. Short and Long Term Toxic Effects of Treatment

Given the number and the relatively highly toxic-profiles of agents used in combinations of chemotherapy against ALL, treatment related toxicity can theoretically arise in any system or organ and can endanger the lives of affected patients and subsequently alter the treatment outcome. To be more precise, around 1 to 2% of childhood ALL patients who successfully attain remission may still die during remission due to these toxic effects.[6,79] This risk is influenced by patients' clinical features and host genetics, which can modulate drug metabolism and activity. It is especially elevated for children with Down's syndrome, infants and older teenagers, as well as for patients receiving more intensive therapy.[6] Statistically speaking, as survival rates of childhood ALL improve, toxicity-related death would account for a greater percentage of all-causes mortality.

Opportunistic infections are considered to be the leading cause of death related to side-effects of ALL treatment. Other frequent toxicities include cardiotoxicity, hepatotoxicity, nephrotoxicity, bone toxicities, metabolic syndrome and obesity, hypersensitivity reactions, pancreatitis, thromboembolism, central or peripheral neuropathy, among others, each being caused by one or more highly effective antileukemic agents as illustrated in Table 5. As a matter of fact, it has been estimated that 50 to 70% of all patients will, at one point, experience at least one of these treatment-related toxicity.[6,56,80,81] Furthermore, survivors of childhood ALL are at risk of higher treatment-associated morbidity with a cumulative incidence of chronic health conditions reaching almost 75%.[6,75,82]

Nonetheless, precision medicine strategies offer the potential to mitigate these risks by allowing the individualization of therapy, thus tailoring the drug exposure based on predicted risk/benefit equilibrium between relapse risk and possible toxic effects.[6,75]

Table 5. Class-specific and universal side-effects of chemotherapeutic agents.

Drug	***Principle Side-Effects***
Asparaginase	Anaphylaxia, allergies and hypersensitivity reactions, pancreatitis, hepatiti, venous thrombosis and other coagulopathies
Vincristine	Peripheral neurotoxicity, constipation, jaw pain, neuropathic pain
Steroids	Mood alteration, sleep disorder, metabolic disturbances, diabetes mellitus, hyperphagia and weight gain, avascular necrosis and osteotoxicity, psychosis (in rare cases)
Methotrexate	Hepatitis, convulsive seizures (especially upon intrathecal administration), severe mucositis
Anthracyclines	Cardiotoxicity, myocardial failure and dysrhythmias
Cytarabine	Cerebellar ataxia, chemical conjunctivitis
Universal side-effects common to most cytotoxic agents *(incidence varies with differing drugs)*	
Temporary hair-loss, mucositis, myelosuppression, immunosuppression, nausea and vomiting, impaired fertility	

Reproduced freely based on data from (Bomken SN Vormoor HJ, Childhood leukaemia, Paediatrics and Child Health Volume 19, Issue 8, August 2009, Pages 345-350.)[83]

1.3.6. Overview on selected ALL treatment-related toxicities important for this thesis.

1.3.6.1. Chemotherapy-Induced Acute Pancreatitis

Acute pancreatitis is defined as the histological inflammation of the pancreatic parenchyma. It is often associated with clinical symptoms such as abdominal pain, along with serum amylase and/or lipase elevation reaching more than three times upper-normal limits and abnormal finding in imaging results. The mechanisms underlying the etiology of this condition are considerably different between the pediatric and adult populations but the treatment, like diagnostics, is based on current adult therapy strategies. Supportive care remains the most important intervention and can include oxygen supplementation and fluid resuscitation, as well as pain management.[81,84] The symptoms of acute pancreatitis can range from mild inflammation leading to minor elevation of pancreatic-enzyme levels adequately managed by supportive care and dose adjustment, to severe pancreatitis that often forces the interruption of the potentially causative agent, consequently compromising the efficacy of treatment. Moreover, it could lead to an extended hospitalization time, and in more severe cases, it might require intensive care and surgical intervention.[80,84] While its occurrence in children is generally rare compared to adults, it is a common complication of childhood cancer treatment, especially in ALL. Many chemotherapeutic agents used in ALL treatment have been associated with acute pancreatitis such as l-asparaginase,

methotrexate, 6-mercaptopurine, doxorubicin, arabinoside and steroids. Also, acute pancreatitis is well recognized as a complication of HSCT and has been associated with the use of certain antibiotics to control infections in children treated with chemotherapeutics such as trimethoprim/sulfomethoxazole, erythromycin and voriconazole.[64,81,84]

Pancreatitis, along with thrombosis, hypersensitivity reactions and allergies, are among the most common side-effects attributed to the administration of l-asparaginase during childhood ALL treatment, and will be discussed in details in the second review article presented in Chapter 3 of this thesis.

1.3.6.2. Chemotherapy-Induced Peripheral Neuropathy

One adverse-reaction of particular interest to this work is the chemotherapy-induced peripheral neuropathy (CIPN). It can be attributed to several chemotherapeutic agents including bortezomib, thalidomide, taxanes (paclitaxel, docetaxel), platinum compounds (cisplatin, carboplatin, oxaliplatin) and vinca alkaloids (vincristine, vinblastine).[85] These agents belong to various chemotherapeutic groups and exert their toxic effects via different mechanisms that are not yet fully understood, but generally imply, damage to the peripheral nerve endings usually attributed to DNA damage, oxidative stress, mitochondria toxicity, or ion channel remodeling.[86]

CIPN typically develops in a symmetrical, length-dependent distribution primarily affecting the large sensory nerve fibers in the extremities of the upper and lower limbs.

Symptoms can manifest as paresthesias ("pins and needles"), dysesthesias and sensory loss that is typically distributed in a "glove-and-stocking" fashion along with defects in deep tendon reflexes.[85] It is associated with debilitating symptoms that may continue to worsen even after stopping the treatment (coasting effect) and can be long-term or permanent for up to 40% of the survivors,[86] subjecting them to other comorbidities and affecting their quality of life.[85-88] Moreover, such symptoms, depending on their severity, could be dose-limiting leading to dose reduction, treatment interruption or even stopping the treatment earlier than scheduled; thus affecting the efficacy of the treatment and influence patient survival.[85,89]

Given that there are no preventative, neuroprotective or curative measures with confirmed efficacy for CIPN at the present,[85,90] it would be very valuable to establish early-detection strategies based on a comprehensive understanding of the clinical and genetic risk factors, as well as the pathophysiological mechanisms implicated in the development of this toxicity. Such understanding could allow clinicians to design patient-tailored treatments that would minimize the risk of CIPN while maintaining the efficacy of treatment, especially for patients who are already considered to be at higher risk of peripheral neuropathy like patients with certain genetic syndromes or diabetes mellitus.[85]

Genetic studies have identified potential CIPN predisposing polymorphisms in genes associated with regulating pharmacokinetic-pathways such as drug efflux and metabolism (*ABCB1*, *CYP2C8*, *GSTP1,* and *GSTM1,* among others).[85,86,89,91,92] In this context, the most widely studied gene-superfamily is that of cytochrome P450, particularly the family of genes encoding the CYP3A enzymes such as *CYP3A4 and CYP3A5* genes in relation with

vincristine-induced peripheral neuropathy.[85,86,89,91,92] *S*everal variants in these genes were linked to a less efficient clearance of vincristine and more severe neuropathy, suggesting their use as predictors of the severity of VIPN. For example, the predicted intrinsic clearance of vincristine is 5-fold greater in CYP3A5 expressers versus non-expressers, leading scientists to hypothesize that carriers of one or two copies of the active *CYP3A5*1* allele may experience less vincristine-induced peripheral neuropathy than subjects with no dysfunctional alleles.[66,85,89,91,92] However, results reported in literature are inconsistent for most CIPN associated drugs and further studies are needed to understand the effect of the genetic component on the risk of developing this toxicity.[85,93]

1.4. Pharmacogenomics and Pharmacogenetics

Pharmacogenomics and Pharmacogenetics (PGx) aim at understanding the effect of inter-individual genetic variability on the outcome of a given treatment and to apply this acquired knowledge towards maximizing the efficacy and minimizing the toxicity of therapy.[75,76] This is especially important in the pediatric population which is considered to be more vulnerable to the impact of toxic effects of medications and is subject to age-dependent pharmacokinetics. The premise is that understanding the influence of genetic variability on the therapeutic response would allow clinicians to devise safer and more effective individual dosing regimens tailored to the specific genetic profile of each patient.[75-77] Of noteworthy, pharmacogenetics and pharmacogenomics are usually used interchangeably.[19]

A lot of noteworthy examples of the usefulness of pharmacogenetics in pediatric oncology can be found in studies of the treatment outcome of acute lymphoblastic leukemia. For example, polymorphisms in the *dihydrofolate reductase* (*DHFR*), *thymidylate synthase* (*TS*) and *5,10-Methylenetetrahydrofolate reductase* (*MTHFR*) genes were linked to treatment outcome of antifolates such as MTX, described earlier as a key component of maintenance therapy in childhood ALL.[66,94-96] Many studies have found that these genetic alterations were associated with a lower treatment efficacy, a higher probability of childhood ALL relapse, an increased frequency of adverse drug events and a greater risk of discontinuing MTX treatment.[66,94-98] Furthermore, the role of some of these variants, particularly in the *DHFR* gene, in predicting treatment outcome of childhood ALL was tested and validated across different treatment protocols and thus can be considered a useful biomarker for treatment personalization.[66,96-98] Another important example showing the elegant history and the dynamic evolution of PGx-guided interventions in childhood ALL is that of the *thiopurine S-methyltransferase* (*TPMT*) gene. It highlights the benefits of a progressive understanding of the role of genetics in influencing the clinical outcome of childhood ALL treatment.[66,75,77] This will be the focus of the first review article presented in this thesis and will be discussed in details in Chapter 2.

1.4.1. Genetic-association approaches

Generally speaking, PGx of ALL treatment focuses on studying the genetic alterations in genes that can affect the pharmacokinetics or the pharmacodynamics of chemotherapeutic agents such as those involved in regulating the activity of metabolizing

enzymes, drug transporters and molecular targets, as well as how they interact together to produce a specific drug-related response or phenotype. There are many approaches currently being used in PGx and the choice of following one approach or the other depends on the type of information available and the goal of the research. Some approaches, such as candidate-gene studies, aim at investigating the role of mutations in a specific site of interest preselected based on prior knowledge or hypothesis suggesting its involvement in modulating the activity of a given drug. Others apply non-biased strategies, such as association studies across the entire genome, in order to detect signals coming from any regions with potential involvement, regardless of whether the association with functional modifications was known before, and then try to understand their implication in influencing treatment response.[18,19,95,99,100] It is important to note that the possibility of obtaining false positive findings when querying a large number of variants is considerably high and is proportionally related to the number of tests performed; thus requiring the implementation of a suitable method to correct for the inherent error in multiple testing. This can be addressed by either applying the Bonferroni adjustment method to recalculate the significance threshold based on the number of associations tested, or by obtaining the false discovery rate (FDR) associated with the findings.

1.4.2. *Genotyping vs. Sequencing methods*

1.4.2.1. *Genotyping*

Genotyping involves the targeted sampling of specific sites of interest in the genome and is used to determine which genetic DNA variants, from a predetermined list, an individual possesses. Depending on the type of variants of interest, their number, location, and available resources, a variety of methods can be applied to perform genotyping.

Polymerase chain reaction, coupled with restriction fragment length polymorphism analysis, can be used if the polymorphism under investigation falls within the active digestion site of a restriction enzyme. The concept being that this polymorphism would result in a differential endonuclease activity depending on the presence or absence of a particular nucleotide, thus producing fragments of different lengths that can be distinguished though gel electrophoresis as illustrated in Figure.3a.[101]

Allele-specific oligonucleotide hybridization technique can also be paired to polymerase chain reaction and is particularly useful if a large number of samples are to be screened for one or a few variants. This technique exploits the influence of the SNP of interest (presence or absence of a specific nucleotide) on the DNA-binding affinity of oligonucleotide probes tagged with radioactivity. This would result in an on/off signal reflecting the success or failure of hybridization, respectively, and thus the specific genotype of the sample at this precise site as illustrated in Figure.3b.[102]

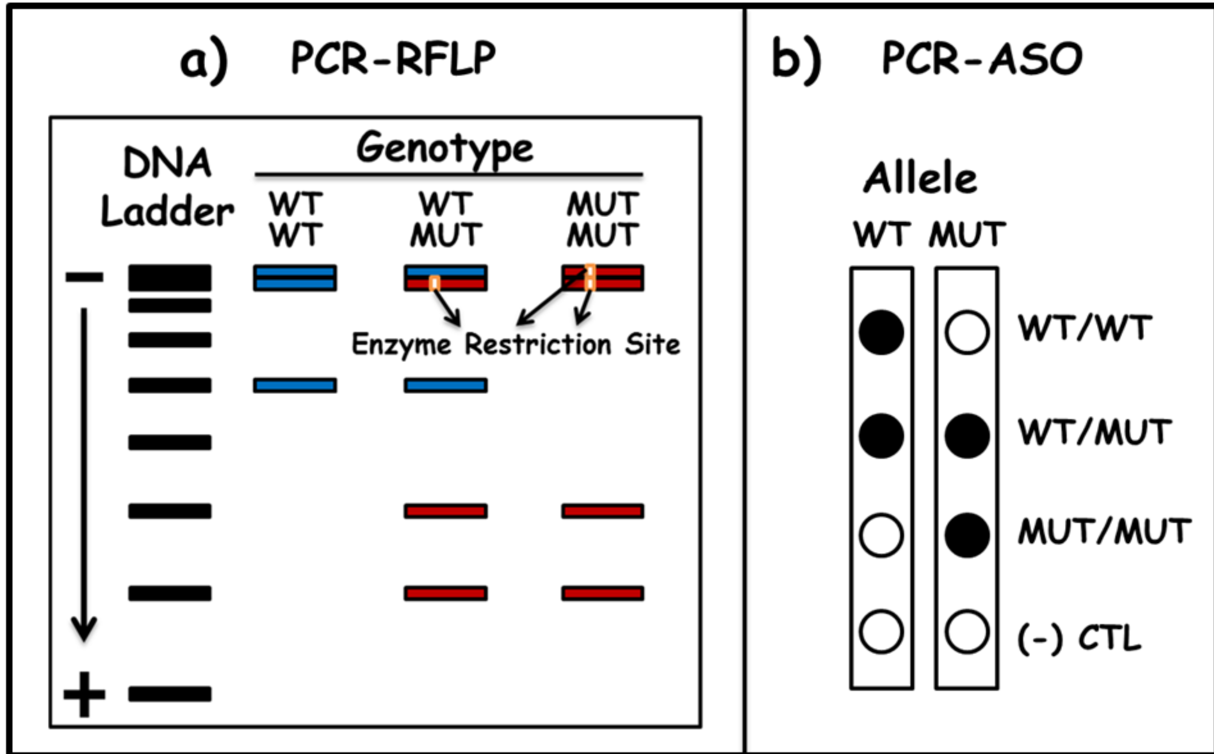


Figure 3. Visual illustration of Polymerase Chain Reaction (PCR)-based genotyping techniques.

Following PCR amplification of the DNA sequence of interest, the allelic combination of the sample, and thus the genotype, can be obtained using **a)** Restriction Fragment Length Polymorphism (PCR-RFLP) approach, or **b)** Allele-Specific Oligonucleotide hybridization (PCR-ASO) technique. WT: wild-type; MUT: mutant; (-)CTL: negative control.

However, for quick screening aimed at investigating large numbers of variants at once, especially common variants, a more efficient and accurate method is the use of genotyping arrays. Illumina's Human1M BeadChip® gene array-based technology and Affymetrix GeneChip® Genome-Wide Human SNP Array, as well as arrays detecting microRNAs have all been used to detect SNPs and copy number variations in ALL.[103] The application of microarray genotyping approaches has significantly advanced our understanding of the molecular biology of ALL.[20,103] However, their usefulness is restricted by their limited capacity when it comes to detecting chromosomal rearrangements, structural variations and focal aberrations such as small insertions and deletions.[103]

1.4.2.2. Sequencing

Sequencing is a method used to determine the exact sequence of a certain length of DNA or RNA that can vary from a short stretch of nucleotides to the entire genome. Many techniques can be applied for genome sequencing in childhood ALL and each has its own advantages and disadvantages.

Sanger sequencing (also known as dideoxy or capillary electrophoresis sequencing), is the traditional sequencing method whereby DNA polymerase adds fluorescent nucleotides one by one onto a growing DNA template strand. Each of the incorporated nucleotide is tagged by a unique fluorescent colour. Accordingly, the sequential addition of new a nucleotide with each cycle gives rise to a sequence of colours that can be translated into the corresponding nucleotide sequence.[20,104]

Another type of sequencing is based on massively parallel DNA-sequencing systems, referred to as next-generation sequencing (NGS) approaches.[20,103] In principle, the concepts behind Sanger and NGS technologies are similar. However, the major difference lies in the fact that NGS approaches sequence millions of fragments simultaneously, hence providing a high-throughput platform that allows for sequencing a huge number of different DNA strands at once, thus producing data on hundreds to thousands of genes per run.[20,40]

While Sanger sequencing is can be regarded as a fast, cost-effective sequencing method for low numbers of targets, it is often viewed as having a lower sensitivity and a limited discovery power compared to NGS, and is generally not considered as cost-effective for sequencing high numbers of targets. On the other hand, NGS approaches provide a comprehensive genomic coverage, more accurate data with deep sequencing, and a higher sensitivity to detect novel or low frequency variants; therefore enabling population-scale genome research. Nonetheless they are rather time-consuming and less cost-effective when it comes to sequencing low numbers of targets.[104]

Several NGS approaches exist, and to each its advantages and disadvantages. Targeted sequencing can detect specific alterations in the sequence of a particular gene of interest, but it cannot necessarily recognize changes of copy number variation or genetic rearrangements.[103] Another method employs mRNA sequencing (RNA-seq) technique in order to perform a transcriptome analysis that would detect changes in protein-coding transcripts, and to a lesser extent, alterations in regulatory regions that can influence the process of leukemogenesis or the response to treatment. It is particularly useful in

detecting chimeric proteins resulting from gene fusions frequently reported in ALL, as well as new genetic isoforms of a particular RNA sequence.[103]

Whole-exome sequencing (WES) and whole-genome sequencing (WGS) are well known examples of NGS technologies with many genes and causal variants discovered by their application.[20,103] They emerged recently as powerful tools to detect sequence mutations and are being widely used for spotting genetic alterations associated with the development of ALL and its outcome. By comparing the host and tumor sequences, somatic and germline mutations can be distinguished and the comparison of the DNA sequence of cancer cells with that of normal cells could offer a deeper understanding of cancer. WES is used to sequence the coding regions of the genome since it is able to capture the exons along with the promoter and non-coding domains as illustrated in Figure.4. It is useful for detecting and studying low-frequency mutations, and to a lesser extent, copy number variations of the sequenced genes, but is not particularly efficient for detecting insertions, deletion, or structural rearrangement mutations. However, it provides a less challenging lower-cost alternative to WGS for spotting genetic alterations with high penetrance. WGS on the other hand, is capable of detecting all types of genetic alteration and is more suitable for dealing with high-frequency variants compared to WES. Nonetheless, it is considerably more costly and demanding, as it requires superior sequencing procedures and the comparison is performed against the entire human genome. Thus, it is important to understand the points of strength, as well as the limitations, of each technique since both uncommon and frequent genetic variants were shown to influence phenotypes in childhood ALL.[20,99,103]

Of note, in the new era for pharmacoeconomics, it is important to evaluate the incremental costs and health outcomes associated with these technologies in order argue for their added-benefit compared with those used in current practice and to support their utility. In an analysis that evaluated a final number of 36 studies published in generally used databases between 2005 and 2016 and which investigated the use of WES and WGS in a variety of genetic conditions in clinical practice, the authors reported that the cost estimates for a single test ranged from $555 to $5,169 for WES and from $1,906 to $24,810 for WGS.[105] Interestingly, both of the highest estimates were reported in Canadian studies. The analysis also reported that the WES cost estimates for a trio ranged from $3,825 to $9,304.[105] However, while these ranges seem wide, the authors noted that this could be in part because many publications did not state which components were included in cost estimates, and thus limiting the debate for a health economic evidence base to support a more widespread use of WES and WGS in clinical practice. The authors also reported a tendency for the cost to decline over the course of time but also noted that this observation was based on limited evidence and a small sample size.[105]

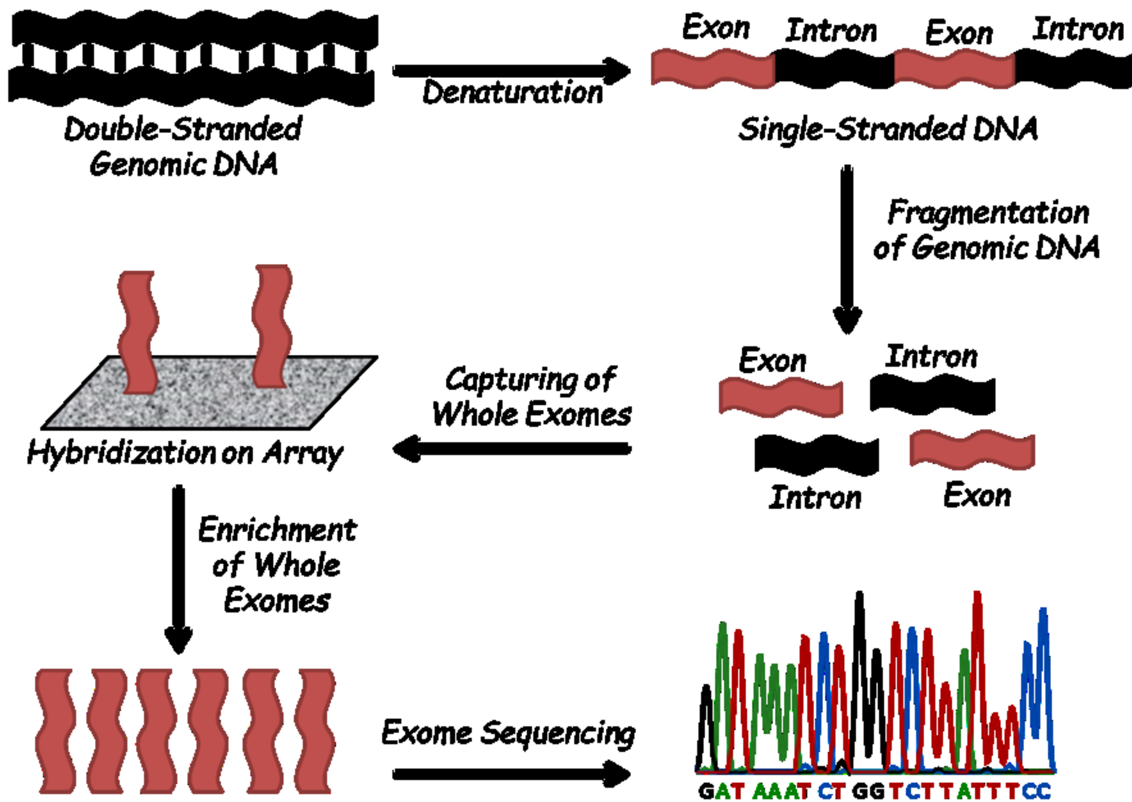


Figure 4. Visual illustration of the workflow of the array-based exome-enrichment and whole-exome Sequencing approach.

Following the denaturation of the double-stranded genomic DNA and its fragmentation, whole-exomes are captured and enriched through binding to complementary probes fixed on a high-density microarray while all other parts of the genome are eliminated. The enriched whole-exomes fragments are then sequenced through a process that gives rise to a sequence of colors each corresponding to a particular nucleotide, which is then translated into the nucleotide sequence of the fragment.

1.4.2.3. *Whole-Exome Sequencing*

The human genome contains approximately 3 billion base pairs distributed across coding and noncoding sequences. WES is the application of the next-generation sequencing technology to identify variations in the coding regions and splice-site variants of the genome, also known as exons, which only make up around 1% of the entire human genome (3×10^7 base pairs). Nonetheless, it has been estimated that around 85% of disease-causing mutations, as well as many disease-predisposing ones, are located in the exonic regions of the human genome, which provides a strong argument for the utility of sequencing the complete coding regions in order to detect such variants.[20,106,107] Current WES technologies allow coverage of over 95% of exons. Since the appearance of the first report on selective sequencing of whole exome in 2009, WES has improved our understanding of the genetic pathology of many heterogeneous monogenic phenotypes such as hearing loss, intellectual disabilities and movement disorders, as well as common diseases and complex disorders including cardiovascular disease, hypertension obesity and diabetes. Furthermore, many examples exist in literature that support the diagnostic and preclinical application of WES for the characterization of mutations in genes leading to phenotypically similar disorders, as well as its application for therapeutic purposes such as the identification of pharmacogenetics variants and gene–disease–drug interactions.[20,108,109]

Particularly, since sequence variations may modulate the predisposition to cancer development and the response to treatment, WES is widely used to identify germline and somatic mutations and studying their influence on outcome in cancer, as well as to build and reconstruct cancer mutation networks. For example, in a study that compared 10 non-familial pancreatic neuroendocrine tumors using data derived from WES, causative genetic mutations in *DAXX/ATRX*, *MEN1*, genes and the mTOR pathway were identified and were then used to support disease prognosis.[110]

1.4.3. Genome Editing Techniques

Recent breakthroughs in molecular biotechnology techniques allowed to scientists to decipher the genetic code and unlocked the gates to the possibility of altering the DNA through editing genes or altering pathways as well as changing the fate of mRNA through post-transcriptional modifications. This pushed the once "traditional" health-care to step into the era of molecular and precision medicine. While earlier techniques based on protein-based nuclease systems such as meganuclease, transcription activator-like effector nucleases (TALENs), and zinc-finger nucleases (ZFNs) suffered from lower specificity due to their off-targets side effects, the more recent discovery of CRISPR/Cas9 technology holds the promise to take genomics to the next level by providing better efficiency, feasibility, and clinical application.[111]

ZFNs and TALENs are both based on the concept of exploiting nuclease proteins for DNA sequence editing. While ZFNs are a combination of proteins that exhibit a zinc-finger-binding domain (that would reach and recognise the desirable splice site) coupled with restriction endonucleases which would then cut at a specific codon. TALENS use a similar principle to ZFNs whereby a restriction nuclease is fused to a DNA-binding protein domain called TAL effector, but the components of the array have the advantage of being able to recognize individual nucleotides rather than codons (triplets of nucleotides) as is the case with ZFNs, thus slightly reducing the risk of producing an off-target effect by making TALENs a little more site specific.[111]

The simplicity and specificity of CRISPR/Cas9 technology allowed its widespread popularity over the other techniques. The concept of this technique is based on RNA-DNA systems and was first adopted from the ancient natural immune system of some prokaryotic cells like Archea and some bacteria. The artificial CRISPR/Cas9 system could be programmed to target any DNA sequence for cleavage. Briefly, CRISPR RNA, also termed guide RNA (gRNA), is specifically engineered to recognize DNA target-site by manipulating the nucleotide sequence of this guide RNA. It is then coupled with the activity of a Cas9 enzyme that has a nuclease function, which, with the guidance of gRNA, reaches the selected sites and creates double-stranded DNA nicks, causing a desired site-specific cleavage and the destruction of the DNA fragment in question. Moreover, the created nick can then be specifically filled by inserting a chosen sequence of nucleotides, thus allowing scientists to alter the genetic code in a very specific manner including the creation of a single nucleotide polymorphism.[111]

1.5. *Study Hypotheses*

- The observed inter-individual variability of patients' susceptibility to treatment-induced complications during childhood acute lymphoblastic leukemia treatment is influenced by the genetic background of the individual and is associated with specific genetic variations.

- Using whole-exome sequencing data in the context of a hypothesis-free exome-wide association study approach can allow detecting novel genetic markers associated with modulation of patients' predisposition to particular response phenotypes.

- Functional analyses and molecular understanding of the individual contribution of each of these genetic variants, as well as the overall contribution of the patient's genetic signature, can provide a valuable insight on the mechanisms predisposing patients differentially to the studied therapeutic responses; which can then be applied towards reducing the frequency and/or severity of adverse drug reactions, and improving patient care and treatment outcomes.

1.6. Research Objectives

- To analyze available whole-exome sequencing data through an exome-wide association study approach in order to identify common germline variants influencing the susceptibility of childhood acute lymphoblastic leukemia patients to:
  - Asparaginase-related complications (allergies, pancreatitis and thrombosis).
  - Vincrisitine-induced peripheral neuropathy.

- To validate the top-ranking association signals and assess the impact of the identified common germline variants on patients' susceptibility to adverse drug reactions and their influence on treatment outcome.

- To explore the clinical utility of the detected variants in predicting the patient's risk of developing a particular complication.

- To investigate the functional impact of the associated variants using in vitro cell-based assays.

1.7. References

1. Terwilliger T, Abdul-Hay M. Acute lymphoblastic leukemia: a comprehensive review and 2017 update. *Blood Cancer J.* 2017;7(6):e577.

2. Katz AJ, Chia VM, Schoonen WM, Kelsh MA. Acute lymphoblastic leukemia: an assessment of international incidence, survival, and disease burden. *Cancer Causes Control.* 2015;26(11):1627-1642.

3. Zuckerman T, Rowe JM. Pathogenesis and prognostication in acute lymphoblastic leukemia. *F1000Prime Rep.* 2014;6:59.

4. National-Cancer-Institute. Childhood Acute Lymphoblastic Leukemia Treatment (PDQ®)–Health Professional Version. 2018; https://www.cancer.gov/types/leukemia/hp/child-all-treatment-pdq. Accessed 09-12, 2018.

5. Alvarnas JC, Brown PA, Aoun P, et al. Acute Lymphoblastic Leukemia, Version 2.2015. *J Natl Compr Canc Netw.* 2015;13(10):1240-1279.

6. Hunger SP, Mullighan CG. Acute Lymphoblastic Leukemia in Children. *N Engl J Med.* 2015;373(16):1541-1552.

7. Rowe JM. Prognostic factors in adult acute lymphoblastic leukaemia. *Br J Haematol.* 2010;150(4):389-405.

8. Ma H, Sun H, Sun X. Survival improvement by decade of patients aged 0-14 years with acute lymphoblastic leukemia: a SEER analysis. *Sci Rep.* 2014;4:4227.

9. Schuz J, Erdmann F. Environmental Exposure and Risk of Childhood Leukemia: An Overview. *Arch Med Res.* 2016;47(8):607-614.

10. Van der Meulen J, Sanghvi V, Mavrakis K, et al. The H3K27me3 demethylase UTX is a gender-specific tumor suppressor in T-cell acute lymphoblastic leukemia. *Blood.* 2015;125(1):13-21.

11. Smith M, Bleyer A, Crist W, Murphy S, Sallan SE. Uniform criteria for childhood acute lymphoblastic leukemia risk classification. *J Clin Oncol.* 1996;14(2):680-681.

12. Pui CH, Sandlund JT, Pei D, et al. Results of therapy for acute lymphoblastic leukemia in black and white children. *JAMA.* 2003;290(15):2001-2007.

13. Cooper SL, Brown PA. Treatment of pediatric acute lymphoblastic leukemia. *Pediatr Clin North Am.* 2015;62(1):61-73.

14. Arico M, Valsecchi MG, Conter V, et al. Improved outcome in high-risk childhood acute lymphoblastic leukemia defined by prednisone-poor response treated with double Berlin-Frankfurt-Muenster protocol II. *Blood.* 2002;100(2):420-426.

15. Geriniere L, Bastion Y, Dumontet C, Salles G, Espinouse D, Coiffier B. Heterogeneity of acute lymphoblastic leukemia in HIV-seropositive patients. *Ann Oncol.* 1994;5(5):437-440.

16. Sehgal S, Mujtaba S, Gupta D, Aggarwal R, Marwaha RK. High incidence of Epstein Barr virus infection in childhood acute lymphocytic leukemia: a preliminary study. *Indian J Pathol Microbiol.* 2010;53(1):63-67.

17. Ferrando AA, Look AT. Clinical implications of recurring chromosomal and associated molecular abnormalities in acute lymphoblastic leukemia. *Semin Hematol.* 2000;37(4):381-395.

18. Wu C, Li W. Genomics and pharmacogenomics of pediatric acute lymphoblastic leukemia. *Crit Rev Oncol Hematol.* 2018;126:100-111.

19. Cheok MH, Pottier N, Kager L, Evans WE. Pharmacogenetics in acute lymphoblastic leukemia. *Semin Hematol.* 2009;46(1):39-51.

20. Rabbani B, Tekin M, Mahdieh N. The promise of whole-exome sequencing in medical genetics. *J Hum Genet.* 2014;59(1):5-15.

21. Mullighan CG, Goorha S, Radtke I, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature.* 2007;446(7137):758-764.

22. Kuiper RP, Schoenmakers EF, van Reijmersdal SV, et al. High-resolution genomic profiling of childhood ALL reveals novel recurrent genetic lesions affecting pathways involved in lymphocyte differentiation and cell cycle progression. *Leukemia.* 2007;21(6):1258-1266.

23. Perez-Andreu V, Roberts KG, Harvey RC, et al. Inherited GATA3 variants are associated with Ph-like childhood acute lymphoblastic leukemia and risk of relapse. *Nat Genet.* 2013;45(12):1494-1498.

24. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. *Nat Genet.* 2009;41(9):1006-1010.

25. Trevino LR, Yang W, French D, et al. Germline genomic variants associated with childhood acute lymphoblastic leukemia. *Nat Genet.* 2009;41(9):1001-1005.

26. Kiechle-Schwarz M, Bauknecht T, Wienker T, Walz L, Pfleiderer A. Loss of constitutional heterozygosity on chromosome 11p in human ovarian cancer. Positive correlation with grade of differentiation. *Cancer.* 1993;72(8):2423-2432.

27. Harvey RC, Mullighan CG, Chen IM, et al. Rearrangement of CRLF2 is associated with mutation of JAK kinases, alteration of IKZF1, Hispanic/Latino ethnicity, and a poor outcome in pediatric B-progenitor acute lymphoblastic leukemia. *Blood.* 2010;115(26):5312-5321.

28. Meyer C, Hofmann J, Burmeister T, et al. The MLL recombinome of acute leukemias in 2013. *Leukemia.* 2013;27(11):2165-2176.

29. Bielorai B, Fisher T, Waldman D, et al. Acute lymphoblastic leukemia in early childhood as the presenting sign of ataxia-telangiectasia variant. *Pediatr Hematol Oncol.* 2013;30(6):574-582.

30. Chessells JM, Harrison G, Richards SM, et al. Down's syndrome and acute lymphoblastic leukaemia: clinical features and response to treatment. *Arch Dis Child.* 2001;85(4):321-325.

31. German J. Bloom's syndrome. XX. The first 100 cancers. *Cancer Genet Cytogenet.* 1997;93(1):100-106.

32. Shah A, John BM, Sondhi V. Acute lymphoblastic leukemia with treatment--naive Fanconi anemia. *Indian Pediatr.* 2013;50(5):508-510.

33. Kishi S, Cheng C, French D, et al. Ancestry and pharmacogenetics of antileukemic drug toxicity. *Blood.* 2007;109(10):4151-4157.

34. Harrison CJ, Moorman AV, Schwab C, et al. An international study of intrachromosomal amplification of chromosome 21 (iAMP21): cytogenetic characterization and outcome. *Leukemia.* 2014;28(5):1015-1021.

35. Arico M, Schrappe M, Hunger SP, et al. Clinical outcome of children with newly diagnosed Philadelphia chromosome-positive acute lymphoblastic leukemia treated between 1995 and 2005. *J Clin Oncol.* 2010;28(31):4755-4761.

36. Roberts KG, Li Y, Payne-Turner D, et al. Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. *N Engl J Med.* 2014;371(11):1005-1015.

37. Moorman AV, Ensor HM, Richards SM, et al. Prognostic effect of chromosomal abnormalities in childhood B-cell precursor acute lymphoblastic leukaemia: results from the UK Medical Research Council ALL97/99 randomised trial. *Lancet Oncol.* 2010;11(5):429-438.

38. Mullighan CG, Su X, Zhang J, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *N Engl J Med.* 2009;360(5):470-480.

39. Kuiper RP, Waanders E, van der Velden VH, et al. IKZF1 deletions predict relapse in uniformly treated pediatric precursor B-ALL. *Leukemia.* 2010;24(7):1258-1264.

40. Nachman JB, Heerema NA, Sather H, et al. Outcome of treatment in children with hypodiploid acute lymphoblastic leukemia. *Blood.* 2007;110(4):1112-1115.

41. Coustan-Smith E, Mullighan CG, Onciu M, et al. Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. *Lancet Oncol.* 2009;10(2):147-156.

42. Jain N, Lamb AV, O'Brien S, et al. Early T-cell precursor acute lymphoblastic leukemia/lymphoma (ETP-ALL/LBL) in adolescents and adults: a high-risk subtype. *Blood.* 2016;127(15):1863-1869.

43. Janczar K, Janczar S, Pastorczak A, et al. Preserved global histone H4 acetylation linked to ETV6-RUNX1 fusion and PAX5 deletions is associated with favorable outcome in pediatric B-cell progenitor acute lymphoblastic leukemia. *Leuk Res.* 2015;39(12):1455-1461.

44. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer.* 2011;11(10):726-734.

45. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 2009;19(1):92-105.

46. Lopez-Lopez E, Gutierrez-Camino A, Pinan MA, et al. Pharmacogenetics of microRNAs and microRNAs biogenesis machinery in pediatric acute lymphoblastic leukemia. *PLoS One.* 2014;9(3):e91261.

47. Gutierrez-Camino A, Umerez M, Martin-Guerrero I, et al. Mir-pharmacogenetics of Vincristine and peripheral neurotoxicity in childhood B-cell acute lymphoblastic leukemia. *Pharmacogenomics J.* 2017.

48. Mei Y, Gao C, Wang K, et al. Effect of microRNA-210 on prognosis and response to chemotherapeutic drugs in pediatric acute lymphoblastic leukemia. *Cancer Sci.* 2014;105(4):463-472.

49. Schotte D, De Menezes RX, Akbari Moqadam F, et al. MicroRNA characterize genetic diversity and drug resistance in pediatric acute lymphoblastic leukemia. *Haematologica.* 2011;96(5):703-711.

50. Bassan R, Spinelli O, Oldani E, et al. Improved risk classification for risk-specific therapy based on the molecular study of minimal residual disease (MRD) in adult acute lymphoblastic leukemia (ALL). *Blood.* 2009;113(18):4153-4162.

51. Bruggemann M, Raff T, Flohr T, et al. Clinical significance of minimal residual disease quantification in adult patients with standard-risk acute lymphoblastic leukemia. *Blood.* 2006;107(3):1116-1123.

52. Jacquy C, Delepaut B, Van Daele S, et al. A prospective study of minimal residual disease in childhood B-lineage acute lymphoblastic leukaemia: MRD level at the end of induction is a strong predictive factor of relapse. *Br J Haematol.* 1997;98(1):140-146.

53. Raff T, Gokbuget N, Luschen S, et al. Molecular relapse in adult standard-risk ALL patients detected by prospective MRD monitoring during and after maintenance treatment: data from the GMALL 06/99 and 07/03 trials. *Blood.* 2007;109(3):910-915.

54. Sutton R, Venn NC, Tolisano J, et al. Clinical significance of minimal residual disease at day 15 and at the end of therapy in childhood acute lymphoblastic leukaemia. *Br J Haematol.* 2009;146(3):292-299.

55. Farber S, Diamond LK. Temporary remissions in acute leukemia in children produced by folic acid antagonist, 4-aminopteroyl-glutamic acid. *N Engl J Med.* 1948;238(23):787-793.

56. Al-Mahayri ZN, Patrinos GP, Ali BR. Pharmacogenomics in pediatric acute lymphoblastic leukemia: promises and limitations. *Pharmacogenomics.* 2017;18(7):687-699.

57. Conter V, Arico M, Basso G, et al. Long-term results of the Italian Association of Pediatric Hematology and Oncology (AIEOP) Studies 82, 87, 88, 91 and 95 for childhood acute lymphoblastic leukemia. *Leukemia.* 2010;24(2):255-264.

58. Mitchell C, Richards S, Harrison CJ, Eden T. Long-term follow-up of the United Kingdom medical research council protocols for childhood acute lymphoblastic leukaemia, 1980-2001. *Leukemia.* 2010;24(2):406-418.

59. Moricke A, Zimmermann M, Reiter A, et al. Long-term results of five consecutive trials in childhood acute lymphoblastic leukemia performed by the ALL-BFM study group from 1981 to 2000. *Leukemia.* 2010;24(2):265-284.

60. Pui CH, Pei D, Sandlund JT, et al. Long-term results of St Jude Total Therapy Studies 11, 12, 13A, 13B, and 14 for childhood acute lymphoblastic leukemia. *Leukemia.* 2010;24(2):371-382.

61. Silverman LB, Stevenson KE, O'Brien JE, et al. Long-term results of Dana-Farber Cancer Institute ALL Consortium protocols for children with newly diagnosed acute lymphoblastic leukemia (1985-2000). *Leukemia.* 2010;24(2):320-334.

62. Kamps WA, van der Pal-de Bruin KM, Veerman AJ, Fiocco M, Bierings M, Pieters R. Long-term results of Dutch Childhood Oncology Group studies for children with acute lymphoblastic leukemia from 1984 to 2004. *Leukemia.* 2010;24(2):309-319.

63. Hunger SP, Lu X, Devidas M, et al. Improved survival for children and adolescents with acute lymphoblastic leukemia between 1990 and 2005: a report from the children's oncology group. *J Clin Oncol.* 2012;30(14):1663-1669.

64. Wolthers BO, Frandsen TL, Patel CJ, et al. Trypsin encoding PRSS1-PRSS2 variation influence the risk of asparaginase-associated pancreatitis in children with acute lymphoblastic leukemia: a Ponte di Legno toxicity working group report. *Haematologica.* 2018.

65. Jabbour E, O'Brien S, Konopleva M, Kantarjian H. New insights into the pathophysiology and therapy of adult acute lymphoblastic leukemia. *Cancer.* 2015;121(15):2517-2528.

66. Mei L, Ontiveros EP, Griffiths EA, Thompson JE, Wang ES, Wetzler M. Pharmacogenetics predictive of response and toxicity in acute lymphoblastic leukemia therapy. *Blood Rev.* 2015;29(4):243-249.

67. Nachman J, Siebel N, Sather H, et al. Outcome for Adolescent and Young Adults 16–21 years of age (AYA) with Acute Lymphoblastic Leukemia (ALL) Treated on the Children' s Cancer Group (CCG) 1961 Study. *Blood.* 2004;104:683-683.

68. Stock W, Luger SM, Advani AS, et al. Favorable Outcomes for Older Adolescents and Young Adults (AYA) with Acute Lymphoblastic Leukemia (ALL): Early Results of U.S. Intergroup Trial C10403. *Blood.* 2014;124:796-796.

69. Hough R, Rowntree C, Goulden N, et al. Efficacy and toxicity of a paediatric protocol in teenagers and young adults with Philadelphia chromosome negative acute lymphoblastic leukaemia: results from UKALL 2003. *Br J Haematol.* 2016;172(3):439-451.

70. Bhojwani D, Howard SC, Pui CH. High-risk childhood acute lymphoblastic leukemia. *Clin Lymphoma Myeloma.* 2009;9 Suppl 3:S222-230.

71. Vrooman LM, Stevenson KE, Supko JG, et al. Postinduction dexamethasone and individualized dosing of Escherichia Coli L-asparaginase each improve outcome of children and adolescents with newly diagnosed acute lymphoblastic leukemia: results from a randomized study--Dana-Farber Cancer Institute ALL Consortium Protocol 00-01. *J Clin Oncol.* 2013;31(9):1202-1210.

72. Conter V, Bartram CR, Valsecchi MG, et al. Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study. *Blood.* 2010;115(16):3206-3214.

73. Nguyen K, Devidas M, Cheng SC, et al. Factors influencing survival after relapse from acute lymphoblastic leukemia: a Children's Oncology Group study. *Leukemia.* 2008;22(12):2142-2150.

74. Mullighan CG. Molecular genetics of B-precursor acute lymphoblastic leukemia. *J Clin Invest.* 2012;122(10):3407-3415.

75. Abaji R, Krajinovic M. Current perspective on pediatric pharmacogenomics. *Expert Opin Drug Metab Toxicol.* 2016;12(4):363-365.

76. Wheeler HE, Maitland ML, Dolan ME, Cox NJ, Ratain MJ. Cancer pharmacogenomics: strategies and challenges. *Nat Rev Genet.* 2013;14(1):23-34.

77. Abaji R, Krajinovic M. Thiopurine S-methyltransferase polymorphisms in acute lymphoblastic leukemia, inflammatory bowel disease and autoimmune disorders: influence on treatment response. *Pharmgenomics Pers Med.* 2017;10:143-156.

78. Schultz KR, Carroll A, Heerema NA, et al. Long-term follow-up of imatinib in pediatric Philadelphia chromosome-positive acute lymphoblastic leukemia: Children's Oncology Group study AALL0031. *Leukemia.* 2014;28(7):1467-1471.

79. Blanco E, Beyene J, Maloney AM, et al. Non-relapse mortality in pediatric acute lymphoblastic leukemia: a systematic review and meta-analysis. *Leuk Lymphoma.* 2012;53(5):878-885.

80. Frandsen TL, Heyman M, Abrahamsson J, et al. Complying with the European Clinical Trials directive while surviving the administrative pressure - an alternative approach to toxicity registration in a cancer trial. *Eur J Cancer.* 2014;50(2):251-259.

81. Schmiegelow K, Muller K, Mogensen SS, et al. Non-infectious chemotherapy-associated acute toxicities during childhood acute lymphoblastic leukemia therapy. *F1000Res.* 2017;6:444.

82. Oeffinger KC, Mertens AC, Sklar CA, et al. Chronic health conditions in adult survivors of childhood cancer. *N Engl J Med.* 2006;355(15):1572-1582.

83. Bomken SN, Josef Vormoor H. Childhood leukaemia. *Paediatrics and Child Health.* 2009;19(8):345-350.

84. Stefanovic M, Jazbec J, Lindgren F, Bulajic M, Lohr M. Acute pancreatitis as a complication of childhood cancer treatment. *Cancer Med.* 2016;5(5):827-836.

85. Park SB, Goldstein D, Krishnan AV, et al. Chemotherapy-induced peripheral neurotoxicity: a critical analysis. *CA Cancer J Clin.* 2013;63(6):419-437.

86. Kerckhove N, Collin A, Conde S, Chaleteix C, Pezet D, Balayssac D. Long-Term Effects, Pathophysiological Mechanisms, and Risk Factors of Chemotherapy-Induced Peripheral Neuropathies: A Comprehensive Literature Review. *Front Pharmacol.* 2017;8:86.

87. Verstappen CC, Koeppen S, Heimans JJ, et al. Dose-related vincristine-induced peripheral neuropathy with unexpected off-therapy worsening. *Neurology.* 2005;64(6):1076-1077.

88. Kim BJ, Park HR, Roh HJ, et al. Chemotherapy-related polyneuropathy may deteriorate quality of life in patients with B-cell lymphoma. *Qual Life Res.* 2010;19(8):1097-1103.

89. Egbelakin A, Ferguson MJ, MacGill EA, et al. Increased risk of vincristine neurotoxicity associated with low CYP3A5 expression genotype in children with acute lymphoblastic leukemia. *Pediatr Blood Cancer.* 2011;56(3):361-367.

90. Albers J, Chaudhry V, Cavaletti G, Donehower R. Interventions for preventing neuropathy caused by cisplatin and related compounds. *Cochrane Database Syst Rev.* 2007(1):CD005228.

91. Dennison JB, Jones DR, Renbarger JL, Hall SD. Effect of CYP3A5 expression on vincristine metabolism with human liver microsomes. *J Pharmacol Exp Ther.* 2007;321(2):553-563.

92. Ceppi F, Langlois-Pelletier C, Gagne V, et al. Polymorphisms of the vincristine pathway and response to treatment in children with childhood acute lymphoblastic leukemia. *Pharmacogenomics.* 2014;15(8):1105-1116.

93. Abaji R, Ceppi F, Patel S, et al. Genetic risk factors for VIPN in childhood acute lymphoblastic leukemia patients identified using whole-exome sequencing. *Pharmacogenomics.* 2018;19(15):1181-1193.

94. Krajinovic M, Moghrabi A. Pharmacogenetics of methotrexate. *Pharmacogenomics.* 2004;5(7):819-834.

95. Gervasini G, Vagace JM. Impact of genetic polymorphisms on chemotherapy toxicity in childhood acute lymphoblastic leukemia. *Front Genet.* 2012;3:249.

96. Lopez-Lopez E, Gutierrez-Camino A, Bilbao-Aldaiturriaga N, Pombar-Gomez M, Martin-Guerrero I, Garcia-Orad A. Pharmacogenetics of childhood acute lymphoblastic leukemia. *Pharmacogenomics.* 2014;15(10):1383-1398.

97. Al-Shakfa F, Dulucq S, Brukner I, et al. DNA variants in region for noncoding interfering transcript of dihydrofolate reductase gene and outcome in childhood acute lymphoblastic leukemia. *Clin Cancer Res.* 2009;15(22):6931-6938.

98. Ceppi F, Gagne V, Douyon L, et al. DNA variants in DHFR gene and response to treatment in children with childhood B ALL: revisited in AIEOP-BFM protocol. *Pharmacogenomics.* 2018;19(2):105-112.

99. Relling MV, Ramsey LB. Pharmacogenomics of acute lymphoid leukemia: new insights into treatment toxicity and efficacy. *Hematology Am Soc Hematol Educ Program.* 2013;2013:126-130.

100. Patnala R, Clements J, Batra J. Candidate gene association studies: a comprehensive guide to useful in silico tools. *BMC Genet.* 2013;14:39.

101. Kennedy GC, Matsuzaki H, Dong S, et al. Large-scale genotyping of complex DNA. *Nat Biotechnol.* 2003;21(10):1233-1237.

102. Labuda D, Krajinovic M, Richer C, et al. Rapid detection of CYP1A1, CYP2D6, and NAT variants by multiplex polymerase chain reaction and allele-specific oligonucleotide assay. *Anal Biochem.* 1999;275(1):84-92.

103. Coccaro N, Anelli L, Zagaria A, Specchia G, Albano F. Next-Generation Sequencing in Acute Lymphoblastic Leukemia. *Int J Mol Sci.* 2019;20(12).

104. Illumina. Science and Education /Technology /Next-Generation Sequencing: NGS vs. Sanger Sequencing. 2019; https://www.illumina.com/science/technology/next-generation-sequencing/ngs-vs-sanger-sequencing.html. Accessed 01-12, 2019.

105. Schwarze K, Buchanan J, Taylor JC, Wordsworth S. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genet Med.* 2018;20(10):1122-1130.

106. Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet.* 2003;33 Suppl:228-237.

107. Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? *J Med Genet.* 2011;48(9):580-589.

108. Krauthammer M, Kong Y, Ha BH, et al. Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. *Nat Genet.* 2012;44(9):1006-1014.

109. Wang K, Kan J, Yuen ST, et al. Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat Genet.* 2011;43(12):1219-1223.

110. Jiao Y, Shi C, Edil BH, et al. DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science.* 2011;331(6021):1199-1203.

111. Khan SH. Genome-Editing Technologies: Concept, Pros, and Cons of Various Genome-Editing Techniques and Bioethical Concerns for Clinical Application. *Mol Ther Nucleic Acids.* 2019;16:326-334.

Section A

Chapter 2

Thiopurine S-methyltransferase polymorphisms in acute lymphoblastic leukemia, inflammatory bowel disease and autoimmune disorders: influence on treatment response

The following chapter presents, in a story-like mode, the evolution of the *TPMT* gene as an important pharmacogenomics marker and one of the most solid examples of the success of multi-generational pharmacogenetics research in demystifying the impact of the genetic component on the variability of response to the thiopurine group of drugs.

It presents from a clinical perspective, how clinical observations can lead to identifying patterns of associations, which in turn trigger investigations resulting in a better understanding of the mechanisms of involvement of different factors influencing the overall treatment response, as well as their interplay. It also highlights that even when such pharmacogenomics mechanisms are well understood, therefore allowing for personalised treatment initiatives, obstacles can often arise during the implementation of such interventions, necessitating further investigation and showing the importance of exploring more connections and considering the entirety of the pathway involved rather than focusing on patching the problem at one specific site. While many review articles discussed the pharmacogenetics of TPMT, the work presented in this chapter is distinguished by its emphasis on prospectively conducted clinical trials and the clinical impact of pharmacogenetics intervention on drug response and treatment outcome. Finally, and given the wide implementation of TPMT pharmacogenetics-guided protocols around the world to mitigate the toxicity of thiopurines while maintaining the maximum efficacy, this paper also tries to tackle the next important questions concerning the cost-effectiveness of these pharmacogenetics approaches and what impact have they had on clinical practice.

I particularly take pride in the work presented in this paper and published in the journal of Pharmacogenomics and Personalized Medicine, as it represents my first initiative to write a review paper from a clinical perspective and put together a dynamic display summarizing the evolving multi-generational research conducted by brilliant groups of investigators around the world. This entire work was performed by me (under the supervision and guidance of Dr. Maja Krajinovic) including the literature review, selecting candidate articles to be included, and drafting the manuscript.

REVIEW

Short running header: Influence of TPMT polymorphisms on treatment response

Rachid Abaji et al.

Thiopurine S-methyltransferase polymorphisms in acute lymphoblastic leukemia, inflammatory bowel disease and autoimmune disorders: influence on treatment response

Author name[1]: Rachid Abaji

Author name[2]: Maja Krajinovic

[1]Author affiliations: Research Center, CHU Sainte-Justine, University of Montreal, Montreal, Quebec, Canada; Department of Pharmacology, University of Montreal, Montreal, Quebec, Canada

[2]Author affiliations: Research Center, CHU Sainte-Justine, University of Montreal, Montreal, Quebec, Canada; Departments of Pediatrics and Pharmacology, University of Montreal, Montreal, Quebec,
Canada

Correspondance: Maja Krajinovic
Adresse : Centre de Recherche, CHU Sainte-Justine, 3175 chemin de la Côte-Ste-Catherine, Montréal, QC, Canada. H3T 1C5
Tel: +1(514)345-4931 #6259
E-mail: maja.krajinovic@umontreal.ca

2.1. Abstract:

The *thiopurine S-methyltransferase* (*TPMT*) gene encodes for the TPMT enzyme which plays a crucial role in the metabolism of thiopurine drugs. Genetic polymorphisms in this gene can affect the activity of the TPMT enzyme and have been correlated with variability in the response to treatment with thiopurines. Advances in the pharmacogenetics of *TPMT* allowed the development of dosing recommendations and treatment strategies to optimize and individualize thiopurine prescribing in attempt to enhance treatment efficacy while minimizing toxicity. The influence of genetic polymorphisms in the *TPMT* gene on clinical outcome has been well-documented and replicated in many studies. In this review we provide an overview of the evolution, results, conclusions and recommendations of selected studies which investigated the influence of *TPMT* pharmacogenetics on thiopurine treatment in acute lymphoblastic leukemia, inflammatory bowel disease and autoimmune disorders. We focus mainly on prospective studies that explored the impact of individualized TPMT-based dosing of thiopurines on clinical response. Together, these studies demonstrate the importance of pre-emptive *TPMT* genetic screening and subsequent dose adjustment in mitigating the toxicity associated with thiopurine treatment while maintaining treatment efficacy and favourable long-term outcomes. In addition, we briefly address the cost-effectiveness of this pharmacogenetics approach and its impact on clinical practice as well as the importance of recent breakthrough advances in sequencing and genotyping techniques in refining the *TPMT* genetic screening process.

Keywords: TPMT, Pharmacogentics, Thiopurine, 6-Mercaptopurine, Azathioprine, ADRs.

2.2. Introduction

Thiopurine S-methyltransferase (TPMT) is an important cytoplasmic enzyme which catalyses the rate-limiting step in the metabolism of thiopurine drugs. It is coded by the *TPMT* gene and exerts its effect via S-adenosyl-L-methionine as the S-methyl donor and S-adenosyl-L-homocysteine as a by-product.[1-3] Thiopurine drugs, mainly 6-mercaptopurine (6-MP), and its pro-drug azathioprine (AZA), are implicated as anti-metabolite cytotoxic and immunosuppressive agents in the treatment of malignancies such as acute lymphoblastic leukemia (ALL), inflammatory disorders like inflammatory bowel disease (IBD) and many autoimmune disorders including rheumatoid arthritis, systemic lupus erythematosus, autoimmune hepatitis and generalized eczematous disorders.[3-5] However, gastrointestinal disturbances (like nausea and vomiting), rashes, as well as more serious adverse drug reactions (ADRs) like bone marrow toxicity, hepatotoxicity and pancreatitis can lead to discontinuation of therapy in up to a third of patients[6] which limits the use of these drugs.[2]

AZA is an inactive compound that must be converted into 6-MP via a glutathione-dependent process and both drugs eventually produce 6-thioguanine nucleotides (6-TGNs), a mechanism through which thiopurines exert both their cytotoxic and therapeutic effects.[7,8] Numerous studies have demonstrated that the efficacy and toxicity of thiopurine drugs are correlated to the activity of the TPMT enzyme as this enzyme competes with xanthine-oxidase and hypoxanthine-guanine-phosphoribosyl-transferase to determine the amount of 6-MP metabolized to 6-TGNs.[1,7-10] 6-TGNs then either incorporate directly into DNA which triggers delayed cytotoxicity or they inhibit intracellular signalling pathways

which ultimately promote cell death via apoptosis.[11] Furthermore, 6-MP is also metabolized to methyl-thioinosine-monophosphate which provokes an additional cytotoxic effect by inhibiting *de novo* purine synthesis.[12]

Thioguanine (TG) is also a pro-drug that belongs to the thiopurines family (2-amino-6-mercaptopurine) and is also partly metabolized by TPMT. Like AZA and 6-MP, it exerts its effect through mechanisms that involve the production of 6-TGNs, but they differ in the pathways implicated as depicted in figure-1. However, due to its more pronounced toxicity profile and the lack of additional benefit, its use became somewhat limited to the intensification phase of some anti-leukemia protocols.[13]

TPMT deficiency has been described around 3 decades ago and it is currently well-established that homozygous or compound heterozygous carriers of TPMT-deficient alleles have a significantly higher-risk of early severe myelosuppression than patients homozygous for the wild-type.[14,15] Patients with absent or reduced TPMT activity accumulate high doses of 6-TGNs, resulting in thiopurine-induced myelotoxicity that is characterised by early-onset of severe neutropenia when such patients are treated with standard doses of thiopurine drugs. This toxicity is particularly evident in patients carrying two non-functional alleles and requires treatment cessation or dose adjustment.[1,5,11,16-21] Bone marrow suppression has been linked to higher cumulative incidence of infections, mortality, and death.[5,22,23] Conversely, myelosuppression can be induced by a number of factors independent of TPMT in individuals taking thiopurines (i.e. co-medications, viral infections, underlying disease and idiosyncratic reactions,[24] as well as genetic polymorphisms in genes other than *TPMT* encoding enzymes involved in thiopurines

metabolism like *inosine triphosphate pyrophosphatase (ITPA)*[25] and *Nudix Hydrolase 15 (NUD15)*[26] genes. On the other end of the spectrum, some studies indicated that high TPMT activity has been linked to poor treatment response and that an elevated dose requirement is needed in order to achieve the therapeutic effect.[27,28]

2.3. TPMT Pharmacogenetics

Enzymatic activity of TPMT can be indirectly assessed through red blood cell enzyme activity assay (phenotype) or can be inferred from the genetic profile of the white blood cells.[1,11,29] Genetic polymorphisms in the *TPMT* gene can affect the enzymatic activity of TPMT and have been studied extensively. To date, over 38 variant alleles have been identified.[2,15,30-32] They have been correlated with variability in response to thiopurine drugs which provides an important example of the clinical importance of pharmacogenetics. Nonetheless, only few of these polymorphisms are considered in clinical settings which can identify the most frequent reduced-activity TPMT alleles and account for ≥95% of variant *TPMT* alleles.[12,18,31-33] The wild-type allele is defined as *TPMT*1*. The mutant *TPMT*2* allele is defined by the G238C transversion whereas the *TPMT*3* family alleles are defined by the G460A and A719G transitions [i.e. *TPMT*3A* (G460A and A719G), *TPMT*3B* (G460A) and *TPMT*3C* (A719G)].[12,18,31-33] The prevalence of TPMT variants is much higher among Caucasians (8.1–10.1%) than in Asian populations (2.3–4.2%)[15] and it is well-established that *TPMT*3A* is the most prevalent mutant allele in Caucasians, making up to (85%) of all observed mutant alleles,[14,18] while *TPMT*3C* is the most frequently found allele in African and South-East Asian populations.[14,34]

Other than variants in the coding region of *TPMT*, it is being increasingly acknowledged that variants in the non-coding regions such as the *TPMT*-promoter and introns can also affect the activity of the TPMT enzyme, possibly by influencing the transcription of its gene.[3] One well-studied example of such polymorphisms is the VNTR region. VNTR stands for `variable number of tandem repeats` which is a rare microsatellite region in the *TPMT* gene promoter. Interestingly, studies have shown that the architecture of this region can modulate *TPMT* transcription and possibly enzyme activity. For example, higher *TPMT* promoter activity was shown to be associated to a region that contains five or seven GCC repeats rather than six. Thus, studies suggest the use of VNTR-architecture as a pharmacogenomic biomarker to refine the *TPMT* genetic screening process currently used prior to the introduction of thiopurine therapy to enhance the treatment outcome in ALL.[3] However, contrary to the results of ALL studies, the expression of the TPMT gene seems to rather decrease in IBD patients treated with thiopurine drugs and thus VNTR genotype cannot predict the TPMT activity which seems to be influenced by the treated condition, the protocol used and the concomitant administration of other drugs.[3]

Across all ethnic groups, approximately 1 in 300 individuals are homozygous (or compound heterozygous) for a mutant *TPMT* allele and have very low or absent TPMT activity while around 4%–11% of individuals are heterozygous and are generally considered to have intermediate enzymatic activity.[1,18,31] Nonetheless, such genotype-based classification is not always representative of the actual state of enzymatic activity. In literature, conflicting data were obtained by studies that addressed the concordance between the genetic and phenotypic tests as results ranged from 100% or almost perfect

match in the majority of studies, to as low as 77% concordance in few of them.[1,3,9-11,19,33,35-38] This discordance was particularly observed in patients with intermediate activity in which the estimated probability of obtaining matching results varied from 70% to 86%.[11,33,36,39]

Many factors influence TPMT enzyme activity and eventually affect this genotype-phenotype concordance such as the age and gender of the patient, co-administration of drugs that could potentially interfere with the disease condition or with TPMT activity (e.g., methotrexate),[40] levels of TPMT co-factor S-adenosyl-methionine,[41] recent blood transfusion,[42] and life span of red blood cells[43] as well as untested rare or novel variants in the coding and regulatory regions of the *TPMT* gene (e.g. *TPMT*38* and the VNTR-architecture).[3,37,44] Furthermore, interethnic variability in the TPMT enzymatic activity levels have been observed, with people of Afro-Caribbean decent having lower activity than Caucasians and South Asians.[1,45,46] Taken together, there is always a risk of misclassifying patients if the decision was based on only one of the two abovementioned methods, but it is also unreasonable to perform both tests for all patients. As a recent randomized clinical trial (RCT) concluded that there was no advantage or disadvantage of TPMT genotyping compared with phenotyping,[16] a more recent study (Hindroff *et al.* 2012) concluded that genotyping was superior to phenotyping and should be considered as the primary choice for pre-treatment evaluation of TPMT function.[35] Nonetheless, phenotype testing supplemented by genotyping can be a useful strategy in specific circumstances (e.g. after recent blood transfusion and for confirmation of intermediate activity in known high risk patients).[16]

Recent advances in the pharmacogenetics of *TPMT* have allowed for the development of dosing recommendations and treatment strategies to optimize and individualize therapy with thiopurines in order to obtain maximum treatment benefit with minimal toxicity.[47] However, the implementation of pharmacogenetic tests in clinical practice is still somewhat limited due to the lack of robust evidence stemming directly from large-scale RCTs and proving the clinical utility of such strategy.[1,16,22] Nonetheless, giving the undisputable influence of pharmacogenetics on TPMT activity and the seriousness of thiopurines-induced toxicities -particularly myelosuppression, several regulatory agencies and clinical guidelines such as the U.S. Food and Drug Administration (FDA), The British National Formulary and The Clinical Pharmacogenetics Implementation Consortium (CPIC) recommend pre-treatment TPMT activity testing either by genotyping or phenotyping.[11,47] In general, most guidelines suggest that the initial dose of thiopurines be reduced to 10% of the standard dose when administered to homozygous carriers of TPMT-deficient alleles, as well as a reduction in administration frequency. For heterozygous patients, the recommendation differs slightly depending on the type of thiopurine used as CPIC guidelines suggest an initial dose of AZA and 6-MP that is 30–70% of the standard protocol dose while the recommendation for TG is 30-50% of that dose.[47]

As new data is being continuously generated by RCTs and studies of the long-term outcome of previous treatment protocols, the strength of the clinical-evidence should be constantly revised and the recommendations of the guidelines should be re-evaluated and modified when deemed necessary. In this review we provide an overview of the evolution, results, conclusions and recommendations of studies that investigated the influence of TPMT pharmacogenetics on clinical response to thiopurines in ALL, IBD and autoimmune disorders.

2.3.1. *Acute Lymphoblastic Leukemia (ALL)*

Childhood ALL is the most frequent pediatric cancer. The survival rate currently exceeds 85% in favourable settings. 6-MP is co-administered with methotrexate as key components in the maintenance therapy for pediatric ALL and their use is associated with significant reduction in disease relapse.[33]

An early study (Lennard *et al.* 1987) reported that TPMT activity was significantly higher in blood samples from ALL patients on long-term 6-MP treatment when compared to controls.[48] They also noted a relationship between low TPMT activity and the risk of developing severe myelosuppression in patients treated with thiopurine drugs, plausibly due to elevated 6-TGN concentrations.[48,49] Others reported that higher TPMT activity was linked to an elevated risk of relapse.[50] These findings led to the suggestion that genetic screening of TPMT activity could play a role in influencing treatment response of childhood ALL.[48,50] Indeed, one study of childhood ALL suggested that prospectively screening for major *TPMT* coding region polymorphisms followed by selective administration of an initially reduced dose of 6-MP to heterozygous patients and a subsequent gradual increase to a target range of blood cell count allowed these patients to eventually achieve the full drug dose without experiencing any toxicity.[51] Many clinical trials have investigated the impact of *TPMT* gene polymorphisms on treatment outcome with most of them demonstrating the benefit of pre-emptive TPMT screening but results were somewhat inconsistent (Table 1).

The Total Therapy XII trial explored the impact of 6-MP dose reduction from the standard protocol dose to a maximum tolerable dose subsequently to the development of myelosuppression and investigate the association between the maximum tolerated doses and *TPMT* genotypes in a total of 188 patients. The results showed that *TPMT* genotype was an important predictor of 6-MP toxicity in ALL patients as the cumulative incidence of dose reduction or treatment interruption was significantly different across the 3 groups (P<0.001); with wild-type patients having the lowest incidence (7%) followed by heterozygous carriers (35%) and homozygous carriers of TPMT-deficient alleles (100%).[10,23] They also concluded that administering lower doses of 6-MP in these patients was successful in maintaining adequately high levels of 6-TGN while allowing the administration of other agents at full protocol doses.[10] Furthermore, the investigators pointed out that the reduced activity patients tended to have improved event free survival (EFS) compared to wild-type patients (P =0.096) and that higher dose intensity of 6-MP was the most significant predictor of that outcome (P =.020).[23] However, the authors also observed a non-significant trend for patients with low TPMT activity to have higher incidence and shorter onset of secondary acute myeloid leukemia as well as higher cumulative incidence of brain tumours when compared to patients with wild-type.[52,53] In their later trial, Total Therapy XIIIB, which included 247 patients and pioneered the implementation of pharmacogenetics in leukemia therapy, they continued to administer a standard initial dose of 6-MP at the start of the continuation therapy but then selectively decreased the dose when deemed necessary based on a strategy that involved up-front knowledge of TPMT status combined with clinical tolerance and measurement of thiopurine metabolites levels.[54,55] They eventually reported that *TPMT* genotype was not associated

with the risk of hematologic relapse and that the long-term outcome showed no association with TPMT status (5-year cumulative incidences of 13.2% ± 2.3% and 6.7% ± 6.7% for wild-type and low-activity genotypes, respectively; p= 0.46), further confirming that considering pharmacogenetics of *TPMT* for dose adjustment of 6-MP dosage in ALL can help to reduce treatment associated toxicity while not compromising its efficacy.[54-57]

In the NOPHO-ALL-92 study of The Nordic Society of Paediatric Haematology and Oncology, a higher risk of relapse was observed in patients homozygous for wild-type (P = 0.02) and/or high TPMT activity (P = 0.002).[36,58,59] However, the authors also observed that patients with low TPMT activity, although at lower risk of relapse, had a higher risk of developing second myeloid neoplasms (SMN) associated with high levels of 6-TGN and methylated metabolites probably leading to DNA damage and subsequent malignancies. The authors believe that this theory explains why low TPMT activity patients did not have a superior overall survival (OS) to those with wild-type activity (P = 0.82) despite their lower risk of relapse.[38,60] These observations, together with the ones from the Total Therapy XII study, led the NOPHO to adopt *TPMT*-genotype dependent initial dosing of 6-MP in their later protocols, ALL-2000 and ALL-2008.[36,58] Indeed, the long-term survival results from the ALL-2000 trial indicate that selecting the initial 6-MP dose based on *TPMT* genotype did reduce the risk of SMN in heterozygous patients but at the expense of an increased risk of relapse. This explains why although a slight non-significant improvement in EFS was achieved by the new protocol, it had no difference in overall EFS or OS from its predecessor (5-years results: NOPHO-ALL-92: EFS = 77.4 ± 1.0%, OS = 87.6 ± 0.8%, n=1654; and NOPHO-ALL-2000: EFS = 79.4 ± 1.5% and OS = 89.1±1.1%, n=1023).[58,61]

In the United Kingdom ALL97 and ALL97/99 trials, wild-type and heterozygous patients on the 6-MP arm initially received the full dose of the drug which was later adjusted to clinical hematologic toxicity whereas TPMT-deficient patients received 10% of the dose also adjusted for toxicity. While investigators observed a similar finding to above studies in that patients with the *TPMT*1/*3A* genotype (n = 99, EFS = 88%) had better outcome at 5-years compared to *TPMT* wild-type patients (n = 1206, EFS = 80%, P = 0·05), paradoxically, patients with the TPMT*1/*3C genotype also had lower EFS than those with **1/*3A* genotype (n = 17, EFS = 53%, P = 0·002). Furthermore, patients with heterozygous genotypes were found to experience more myelosuppression, accumulated higher 6-TGN concentrations and required dose reduction more frequently. However, no association between the risk of secondary malignancy and *TPMT* genotype was found.[13,33,62] In their subsequent trial, ALL2003, which used minimal residual disease (MRD) to guide risk-stratification and treatment intensity, the protocol prospectively observed the influence of *TPMT* genotype on treatment outcome by applying pre-treatment genetic screening of the most common *TPMT* polymorphisms to 2387 of the study patients. The dosing regimen for 6-MP was similar to that of ALL97 in the sense that TPMT-deficient patients received 10% of the dose while the others received a standard dose subsequently adjusted according to a target cell count. The results showed that overall EFS -all TPMT genotypes confounded- was significantly higher than that of the previous protocol which was attributed to the improved survival in the *TPMT* wild-type and *TPMT*1/*3C* genotype groups (EFS at 5-years = 88%, 88% and 94% for *TPMT* wild-type, **1/*3A* and **1/*3C*, respectively). However, within this protocol, no significant differences in OS, EFS or relapse-free survival were observed with respect to *TPMT* genotypes. Thus, it was concluded that the improved risk-adapted protocol

had reduced the influence of *TPMT* genotypes on treatment outcome and that the only factor that affected outcome was MRD. Furthermore, there was no difference in survival within each MRD risk groups with respect to *TPMT* genotypes.[33] Overall, the cumulative experience of the many UKALL trials led to mandating pre-emptive TPMT screening for all children and young adults who start the ALL2011 trial protocol.[11]

In the Berlin-Frankfurt-Münster-2000 (BFM2000) trial, the 6-MP dose was reduced by 10-fold from the standard starting dose for TPMT deficient patients but no dose adjustment was carried for heterozygous carriers who were given similar doses of the homozygous carriers of the wild-type. The investigators assessed the genotypes of 814 patients and used MRD for risk-stratification. The results showed no difference in the rate of hematopoietic toxicity between *TPMT* heterozygous variant carriers and homozygous wild-type carriers or between *TPMT* status and the risk of developing secondary cancers. Interestingly, it was observed that *TPMT* genotype had a significant impact on MRD during induction consolidation treatment as heterozygous patients had better MRD response (2.9-fold reduction) indicating an increased clearance of disease likely due to higher intensity of 6-MP effect (Relative risk = 0.34; 95% CI, 0.13-0.86; P = .02).[63,64]

In summary, altogether, these trials demonstrate the importance of pre-emptive *TPMT* genetic screening and subsequent dose adjustment in mitigating the toxicity associated with thiopurine treatment while maintaining, if not enhancing, treatment efficacy and favourable long-term outcome.

2.3.2. *Inflammatory bowel disease (IBD)*

IBD is a polygenic chronic, relapsing and remitting disease of the gastrointestinal tract that can be divided into two major clinical subtypes, Crohn's disease and ulcerative colitis.[2,65] Thiopurines, particularly AZA, are proven effective in inducing and maintaining long-term remission in IBD patients.[5] More than 20% of patients experience severe ADRs that lead to dose modification, treatment interruption or cessation.[66,67] Bone marrow toxicity represented by leukopenia is one of the most serious thiopurine-related ADRs. Many studies investigated the influence of *TPMT* genotype on the efficacy and toxicity of thiopurines and most suggested a significant impact on clinical response (Table 2).

A systematic review followed by a meta-analysis (Higgs *et al.* 2010) which eventually combined the results of 47 studies that investigated the risk of myelosuppression with respect to intermediate TPMT activity demonstrated a 4.19-fold increase in odd-ratio of leukopenia (95% CI: 3.20–5.48) in IBD patients with reduced TPMT activity compared to wild-type. One critic of this meta-analysis is that it combined rather smaller studies with sample sizes of less than 100 patients in most cases and the majority having retrospective cohort designs. However, in a sub-analysis of this study that combined 834 patients coming only from the 11 studies that had a prospective cohort design, the significant association of reduced TPMT activity with the risk of myelosuppression had an odd-ratio of 4.3 (95% CI: 2.53–7.29).[1] Among these prospective studies, an observational study (Ansari *et al.* 2008) with pre-emptive *TPMT* genetic testing for all patients and a relatively large sample size of 207 participants found that heterozygous *TPMT* genotype strongly predicted treatment

withdrawal due to early-onset of ADRs following a conventional fixed-dosing regimen (79% vs. 35% in heterozygous and wild-type, respectively; P < 0.001). They highlighted that gastric intolerance (GI) was the most frequent reason for withdrawal among this group of patients and that myelotoxicity and GI occurred significantly more frequently among heterozygous (26% and 37%, respectively) than with wild-type patients (0.5% and 7%, respectively). Interestingly, they had a 100% concordance of genotype to phenotype activity and found that TPMT activity was strongly predictive of clinical response as it was significantly higher in non-responders.[68] Other prospective studies with pre-emptive *TPMT* genetic screening component reported that overall thiopurine-related ADRs were significantly more common among patients with low to intermediate TPMT activity when doses were not adjusted;[67] particularly myelotoxicity which was more profound in TPMT-deficient genotype.[67,69]

Two independent meta-analysis (Liu *et al.* 2015 and Dong *et al.* 2010) further investigated the impact of pharmacogenetics on treatment response by exclusively combining studies (14 and 9 studies, respectively) that investigated the association between *TPMT* polymorphisms and ADRs in IBD patients, regardless of the study design (i.e. cross-sectional cohort, prospective cohort and case control studies).[2,65] They involved 2,206 and 1309 patients respectively, and both concluded that *TPMT* polymorphisms were significantly associated with thiopurine-induced overall ADRs and bone marrow toxicity (around 3-fold and 6-fold increase in the odd-ratios, respectively) but not with hepatotoxicity, pancreatitis, flu-like symptoms, gastric intolerance or skin reactions.[2,65]

The TARGET trial is a pragmatic RCT that prospectively investigated the impact of genotype-guided initial dosing of AZA followed by upward-titration to the maximum tolerable dose of the full protocol dose as compared to no genotyping and full standard dose administration to all participants. It included 333 patients with inflammatory diseases and the primary aim was to see if this strategy would result in a significant reduction in the rate of ADRs-induced treatment cessation. No differences were found between the conventional and pharmacogenetics arms with respect to the frequency of treatment interruption due to ADRs (frequency: 27.7% vs. 28.8%; OR: 1.1; 95% CI: 0.66–1.8; $p = 0.74$). On the other hand, the study did not find any difference in the rate of remission between the intervention and control groups indicating that the adjustment did not affect treatment efficacy. However, the investigators did not provide a stratified analysis addressing the differences in outcomes according to genotype groups within each study arm or between the two arms, probably due to small sample size. Moreover, it is worth mentioning that the single patient with homozygous variant genotype in the study who was on the no-genotyping arm and subsequently received the full dose of AZA developed severe neutropenia which underlines the importance of genetic testing to identify this group of patients.[16]

A larger and more recent prospective RCT which involved 783 IBD patients, the TOPIC trial, similarly showed no significant overall impact of *TPMT*-genotype guided dosing of thiopurines on treatment efficacy or on the risk of hematologic ADRs (i.e. leukopenia and thrombocytopenia) between the genotyped and non-genotyped arms (frequency: 7.4% vs. 7.9%; relative risk: 0.93; 95% CI: 0.57–1.52). The efficacy results of this study further advocate that a reduced thiopurine dose does not result in under-treatment. Moreover, a

subgroup analysis of this study -which compared only carriers of *TPMT* variants between the two arms- revealed that the pharmacogenetic-approach was able to significantly decrease the risk of hematologic ADRs by 10-fold in carriers of at least one genetic variant (frequency: 2.6% vs. 22.9%; relative risk: 0.11; 95% CI: 0.01-0.85).[22] Furthermore, the results of the secondary aim of this study excluded any significant association between *TPMT* genotypes and anemia, hepatotoxicity, pancreatitis, skin rash, GI and general malaise which is consistent with the results of the aforementioned meta-analysis and other results in the literature.[2,22,65] It also suggested that factors other than *TPMT* genotype play an important role in the development of thiopurine-induced ADRs.[24]

2.3.3. Autoimmune Disorders

Autoimmune diseases are a group of heterogeneous conditions which basically involve a destructive attack against the host's tissues launched by a deregulated immune system like in systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), autoimmune hepatitis (AIH) and generalized-eczematous disorders. Thus, treatment strategies are usually based on the use of immunosuppressants which act by modifying the activity of the immune system. AZA is widely used as an immunosuppressive agent in autoimmune diseases but again, its use is limited by its ADRs.[14] Similar to the meta-analysis that focused on IBD, another meta-analysis which included 651 patients with autoimmune diseases coming from 11 studies (Liu *et al.* 2015) demonstrated that overall ADRs and AZA-induced bone marrow toxicity are significantly associated with *TPMT* polymorphisms with OR of

3.12 (95% CI: 1.48–6.56) and 3.76 (95% CI: 1.97–7.17), respectively. The results remained significant both in the analysis that grouped the homozygous and heterozygous carriers into one reduced-activity group and in the analysis that focused on heterozygous carriers only. The study also showed a significant association with gastric intolerance with OR= 6.43 (95% CI: 2.04–20.25), but the authors suspect that the observed association might have been driven by a single study since after excluding this study the association was no longer significant with OR 2.1 (95% CI: 0.36–12.42).[14] The study also excludes the association of *TPMT* polymorphisms with hepatotoxicity. The sub-analysis that examined the association with myelosuppression according to type of disease found significant results in SLE, RA and AIH subgroups. They also concluded that the risk prediction of bone marrow toxicity and overall ADRs based on *TPMT* variant-positive genotypes has high specificity (94.10% and 92.93%, respectively) but at the expense of sensitivity (16.30% and 14.85%, respectively).[14] Furthermore, in a prospective study (Meggitt *et al.* 2006) that investigated the impact of TPMT-activity guided AZA dosing on the treatment response in patients with atopic eczema, the investigators concluded that TPMT-based dosing was able to maintain the drug efficacy while reducing the predicted toxicity.[70]

2.4. Cost-Effectiveness

Most of the above studies concluded that TPMT testing could lead to improved prescribing of thiopurines which would ultimately result in an increased treatment efficacy and a reduction in the rate and intensity of ADRs. Nonetheless, the cost-effectiveness of such an intervention is still open to debate. Only a few studies have addressed the cost-effectiveness of *TPMT* pharmacogenetics interventions. In an effort towards this evaluation, a case study examined the cost-effectiveness of prospective *TMPT* genotyping in children with ALL treated with thiopurines and suggested positive results manifested in financial savings and a gain in life-years in the most favorable settings of the sensitivity analysis.[71] Similarly, another study (Winter *et al.* 2004) established a model based on a theoretical IBD population treated with AZA and found that pre-treatment screening for *TPMT* genotype would be cost-effective in avoiding patient mortality due to myelouppression.[29] However, data coming from randomized clinical trials do not necessarily support this conclusion as demonstrated by (Sayani *et al.* 2005) who found that such a technique incurred excessive cost associated with genotyping but did not predict AZA-induced toxicity in IBD patients.[72]

Nonetheless, these studies were too small and not adequately powered to answer this question.[16] A systematic review by (Payne *et al.* 2008) came to the conclusion that screening for TPMT activity either by genotyping or phenotyping was a cost-effective strategy that can be used to reduce healthcare costs while improving clinical effectiveness.[73] Another study by (Marra *et al.* 2002) aimed at the evaluation of the added-value of genetic-screening of *TPMT* followed by dose adjustment of AZA prior to the

initiation of treatment found that the genetic-based dosing dominated the standard dosing strategy in patients with rheumatological disease by reducing the treatment cost and the frequency of AZA-induced side-effects.[74] In a more recent prospective economic evaluation that was conducted alongside the TARGET study, in which the study aim was to test the cost-effectiveness of the *TPMT* genotyping approach in autoimmune diseases, the researchers concluded that the genetic approach had up-to 71% probability of being cost-effective depending on the cost of the genetic test. The results, however, were not conclusive as the observed economic advantage in the intervention group -owing to lower use of resources- was accompanied by a slight (almost negligible) reduction in the quality of life.[75]

2.5. Impact on Clinical Practice

Over the past decade, TPMT enzyme testing gained a lot of acceptance as reflected by the rapid increase in the number of tests performed in clinical practice.[16,76] This sudden increase was the inevitable result of multiple factors supporting this approach which include the increase in the available knowledge about the role of *TPMT* in treatment outcome, the stronger recommendations coming from clinical guidelines like the CPIC and the wider accessibility to genetic-testing (i.e. larger availability, reduced cost, faster turnaround of results and shorter interpretation time).[1,16] This shift in clinical practice was evaluated in the TARGET study which observed that the physicians did follow the recommendations coming from British clinical guidelines (e.g. British Association of Dermatologists Therapy and British Society for Rheumatology) for *TPMT* heterozygous

patients and chose a lower initial dose of AZA for those patients but the investigators also noted that the physicians used overall lower starting doses for wild-type patients as well.[16] This « safe » practice reflects the physicians' reservation regarding the sensitivity and specificity of this test which stems from the fact that being homozygous carrier of *TPMT* wild-type, although predictive of a reduced risk of AZA-induced myelosuppression, it does not completely eliminate the possibility. Indeed, it was mentioned earlier that a fraction of TPMT wild-type patients can still have intermediate TPMT-activity and that other factors play a role in the development of this ADR.[16,33,36] Moreover, other side-effects such as hepatotoxicity, pancreatitis, nausea and vomiting cannot be predicted by *TPMT* testing.[1,14,65] The adoption of pre-treatment *TPMT* screening seems to vary according to discipline as shown in a survey by (Fargher *et al.* 2007) with 94% of dermatologists, 60% of gastroenterologists and only 47% rheumatologists requesting it.[76] This could be related to the level of evidence available in the domain of practice and the strength of the recommendations of the respective guidelines and protocols used by each specialist (e.g. UK guidelines in dermatology and gastroenterology recommend the genetic screening while ALL2011 protocol mandates it).[1,11,76] However, from an evidence-based perspective, and beside the universally accepted association with hematotoxicity, the recommendations for pre-emptive genetic testing still have some margin to evolve. Plus, even in well-established scenarios like in the case of myelosuppression in TPMT-deficient patients, strong evidence is still lacking to support that the pharmacogenetic-approach would result in a significantly better outcome.[14,65]

2.6. TPMT in the New Era of Sequencing

The influence of genetic polymorphisms in the *TPMT* gene on treatment outcome has been well-documented and replicated in many studies. However, studies have also concluded that the genetic-based screening for TPMT activity should be interpreted with caution as the activity of the TPMT enzyme can be co-influenced by other factors, and the development of thiopurine-induced ADRs is a multi-factorial event.[14] For instance, most of the presented studies inferred TPMT activity by genotyping the most common non-functional *TPMT* alleles while results of a recent study that explored the sequencing data suggest that in certain populations, the inferred activity can be refined by incorporating the genotypes of other alleles. The study also identified a new variant in the *TPMT* gene, *TPMT*38* (T514C), which had an allelic frequency of 0.11% and was predicted to be a damaging mutation.[15] Moreover, as increasingly reported by different studies, genetic variants in other genes involved in thiopurines metabolism like *ITPA, hypoxanthine guanine phosphoribosyl transferase (HGPRT),* and *methylene tetrahydrofolate reductase (MTHFR)* as well as variants in genes independent of TPMT can influence thiopurines treatment outcome.[14,20,25,41,70,77] For example, genome-wide association studies have identified variants in the *PACSIN2* gene which influence TPMT activity and were linked to 6-MP related gastrointestinal toxicity in children with ALL, whereas variants in the *NUDT15* gene were associated with thiopurine-induced leukopenia.[26,41] However, since a lot of genes have significant differences in the frequencies of polymorphisms across major ethnic groups, it is important to evaluate the genetic profiles of patients in a global frame that considers all of the genes involved in a specific pathway to better understand the impact of ethnic diversity

on drug response. One particularly interesting example of the role of pharmacoethnicity is the case of NUDT15 in Japanese population as polymorphisms of this gene that are associated with higher risk of toxicity were more frequent than TPMT-deficient variants.[26] Studies also suggest that combining the effects of such polymorphisms with variants in *TPMT* gene could strengthen the predictive power of the risk of developing thiopurines-related toxicity.[14,22] This should soon become feasible with the breakthrough advances in sequencing and genotyping techniques. Indeed, in a recent study that tested the sensitivity, specificity and predictive values of the imputation of *TPMT*-alleles, most values were over 90% indicating that imputation of *TPMT* alleles can be used as a screening method for individuals with high-risk of developing serious thiopurine-induced ADRs.[46] Furthermore, non-genetic factors should be taken into consideration before thiopurine initiation as they can have a big influence on the outcome and might interfere with the genotype-guided dosing.[14]

2.7. Conclusion

In conclusion, although it is currently well established that *TPMT* polymorphisms can explain a certain portion of thiopurine-induced ADRs, particularly hematotoxicity, it is surely not capable of predicting all of them. Indeed, many studies have found that certain ADRs were not associated with a reduced TPMT activity such as pancreatitis and hepatotoxicity. This holds true in the context of ALL, IBD and the different types of autoimmune disorders. What is clear so far is that TPMT-deficient genotypes (homozygous variant carriers and compound heterozygous), and to lesser extent, heterozygous patients

are predisposed to thiopurine-induced severe hematotoxicity.[1,47] However, other factors such as disease progression and co-medications can also modulate the risk of myelosuppression regardless of the genotype. While TPMT-deficient patients will definitely benefit from dose reduction of thiopurines, the validity of this approach for heterozygous carriers is still arguable since studies have shown that not all of these patients are intolerant to thiopurine, and in fact, 30-60% of heterozygous patients do tolerate it.[33,47] Moreover, depending on the treated condition and treatment protocol used, *TPMT* wild-type patients also exhibit higher risks of worse outcome such as hematologic relapse in ALL and treatment failure in IBD, which adds an extra layer of complexity to the already troublesome process of finding the best therapeutic regimen that would ensure maximum efficacy and minimum toxicity.[47] Consequently, regular clinical testing and hematologic assessment remain the mainstay in the monitoring of thiopurine treatment while genetic testing adds the advantage of refining the initial dosing and patient-stratification processes, as well as suggesting customized monitoring for certain patient groups. One nice example backed with strong clinical-evidence is the abovementioned scenario of myelosuppression. Pre-emptive *TPMT* genetic screening and tailored thiopurine initial dosing followed by upward/downward titration and hematological monitoring to a target level of myelosuppression can be considered a cost-effective approach which would allow the prevention and early detection of myelosuppression in this vulnerable population without compromising the efficacy of the treatment.[47]

2.8. Perspective

While the goal of personalized medicine in general, and pharmacogenetics in particular, is to deliver patient-tailored treatments that would ensure maximum efficacy with minimum toxicity, the studies presented in this review make the argument that this is not an easy task. There is a balance to consider between treatment benefits and ADRs that is controlled by multiple factors. This being said, what we can be sure of, for now, is that the more we get to know about the impact of pharmacogenetics on the variability of treatment response, the better we are able to control the outcome to the advantage of the patient. Moreover, most pharmacoeconomic analyses have indicated that screening for *TPMT* pharmacogenetics promises to be cost-effective. With the advent of next-generation sequencing and the many breakthroughs in bioinformatics, the cost of analyzing the entire human genome is bound to drop, which would allow for greater accessibility to genetic data and a larger understanding of how their interactions with each other and with other factors influence the treatment. In the meantime, it is very promising to see that most major institutions have already incorporated pre-emptive *TPMT* screening in their treatment protocols to enhance treatment outcome and the continuously emerging long-term data proving the utility of doing so should encourage other institutions to follow.

2.9. Acknowledgments

Rachid Abaji is a scholar of the Cole Foundation and the Network of Applied Medical Genetics (RMGA) and acknowledges the support of both organizations.

2.10. Disclosure

The authors report no conflicts of interest in this work.

2.11. References

1. Higgs JE, Payne K, Roberts C, Newman WG. Are patients with intermediate TPMT activity at increased risk of myelosuppression when taking thiopurine medications? *Pharmacogenomics.* 2010;11(2):177-188.

2. Dong XW, Zheng Q, Zhu MM, Tong JL, Ran ZH. Thiopurine S-methyltransferase polymorphisms and thiopurine toxicity in treatment of inflammatory bowel disease. *World J Gastroenterol.* 2010;16(25):3187-3195.

3. Kotur N, Dokmanovic L, Janic D, et al. TPMT gene expression is increased during maintenance therapy in childhood acute lymphoblastic leukemia patients in a TPMT gene promoter variable number of tandem repeat-dependent manner. *Pharmacogenomics.* 2015;16(15):1701-1712.

4. Chouchana L, Narjoz C, Beaune P, Loriot MA, Roblin X. Review article: the benefits of pharmacogenetics for improving thiopurine therapy in inflammatory bowel disease. *Aliment Pharmacol Ther.* 2012;35(1):15-36.

5. Gisbert JP, Gomollon F. Thiopurine-induced myelotoxicity in patients with inflammatory bowel disease: a review. *Am J Gastroenterol.* 2008;103(7):1783-1800.

6. Fraser AG, Orchard TR, Jewell DP. The efficacy of azathioprine for the treatment of inflammatory bowel disease: a 30 year review. *Gut.* 2002;50(4):485-489.

7. Lennard L. The clinical pharmacology of 6-mercaptopurine. *Eur J Clin Pharmacol.* 1992;43(4):329-339.

8. Lennard L. TPMT in the treatment of Crohn's disease with azathioprine. *Gut.* 2002;51(2):143-146.

9. Chouchana L, Narjoz C, Roche D, et al. Interindividual variability in TPMT enzyme activity: 10 years of experience with thiopurine pharmacogenetics and therapeutic drug monitoring. *Pharmacogenomics.* 2014;15(6):745-757.

10. Relling MV, Hancock ML, Rivera GK, et al. Mercaptopurine therapy intolerance and heterozygosity at the thiopurine S-methyltransferase gene locus. *J Natl Cancer Inst.* 1999;91(23):2001-2008.

11. Lennard L. Implementation of TPMT testing. *Br J Clin Pharmacol.* 2014;77(4):704-714.

12. Wang L, Pelleymounter L, Weinshilboum R, et al. Very important pharmacogene summary: thiopurine S-methyltransferase. *Pharmacogenet Genomics.* 2010;20(6):401-405.

13. Vora A, Mitchell CD, Lennard L, et al. Toxicity and efficacy of 6-thioguanine versus 6-mercaptopurine in childhood lymphoblastic leukaemia: a randomised trial. *Lancet.* 2006;368(9544):1339-1348.

14. Liu YP, Xu HQ, Li M, et al. Association between Thiopurine S-Methyltransferase Polymorphisms and Azathioprine-Induced Adverse Drug Reactions in Patients with Autoimmune Diseases: A Meta-Analysis. *PLoS One.* 2015;10(12):e0144234.

15. Kim HY, Lee SH, Lee MN, et al. Complete sequence-based screening of TPMT variants in the Korean population. *Pharmacogenet Genomics.* 2015;25(3):143-146.

16. Newman WG, Payne K, Tricker K, et al. A pragmatic randomized controlled trial of thiopurine methyltransferase genotyping prior to azathioprine treatment: the TARGET study. *Pharmacogenomics.* 2011;12(6):815-826.

17. Evans WE, Horner M, Chu YQ, Kalwinsky D, Roberts WM. Altered mercaptopurine metabolism, toxic effects, and dosage requirement in a thiopurine methyltransferase-deficient child with acute lymphocytic leukemia. *J Pediatr.* 1991;119(6):985-989.

18. Yates CR, Krynetski EY, Loennechen T, et al. Molecular diagnosis of thiopurine S-methyltransferase deficiency: genetic basis for azathioprine and mercaptopurine intolerance. *Ann Intern Med.* 1997;126(8):608-614.

19. Evans WE, Hon YY, Bomgaars L, et al. Preponderance of thiopurine S-methyltransferase deficiency and heterozygosity among patients intolerant to mercaptopurine or azathioprine. *J Clin Oncol.* 2001;19(8):2293-2301.

20. Karas-Kuzelicki N, Jazbec J, Milek M, Mlinaric-Rascan I. Heterozygosity at the TPMT gene locus, augmented by mutated MTHFR gene, predisposes to 6-MP related toxicities in childhood ALL patients. *Leukemia.* 2009;23(5):971-974.

21. Peregud-Pogorzelski J, Tetera-Rudnicka E, Kurzawski M, et al. Thiopurine S-methyltransferase (TPMT) polymorphisms in children with acute lymphoblastic leukemia, and the need for reduction or cessation of 6-mercaptopurine doses during maintenance therapy: the Polish multicenter analysis. *Pediatr Blood Cancer.* 2011;57(4):578-582.

22. Coenen MJ, de Jong DJ, van Marrewijk CJ, et al. Identification of Patients With Variants in TPMT and Dose Reduction Reduces Hematologic Events During Thiopurine Treatment of Inflammatory Bowel Disease. *Gastroenterology.* 2015;149(4):907-917 e907.

23. Relling MV, Hancock ML, Boyett JM, Pui CH, Evans WE. Prognostic importance of 6-mercaptopurine dose intensity in acute lymphoblastic leukemia. *Blood.* 1999;93(9):2817-2823.

24. Colombel JF, Ferrari N, Debuysere H, et al. Genotypic analysis of thiopurine S-methyltransferase in patients with Crohn's disease and severe myelosuppression during azathioprine therapy. *Gastroenterology.* 2000;118(6):1025-1030.

25. Stocco G, Cheok MH, Crews KR, et al. Genetic polymorphism of inosine triphosphate pyrophosphatase is a determinant of mercaptopurine metabolism and toxicity during treatment for acute lymphoblastic leukemia. *Clin Pharmacol Ther.* 2009;85(2):164-172.

26. Moriyama T, Nishii R, Perez-Andreu V, et al. NUDT15 polymorphisms alter thiopurine metabolism and hematopoietic toxicity. *Nat Genet.* 2016;48(4):367-373.

27. Dubinsky MC, Yang H, Hassard PV, et al. 6-MP metabolite profiles provide a biochemical explanation for 6-MP resistance in patients with inflammatory bowel disease. *Gastroenterology.* 2002;122(4):904-915.

28. Ansari A, Hassan C, Duley J, et al. Thiopurine methyltransferase activity and the use of azathioprine in inflammatory bowel disease. *Aliment Pharmacol Ther.* 2002;16(10):1743-1750.

29. Winter J, Walker A, Shapiro D, Gaffney D, Spooner RJ, Mills PR. Cost-effectiveness of thiopurine methyltransferase genotype screening in patients about to commence azathioprine therapy for treatment of inflammatory bowel disease. *Aliment Pharmacol Ther.* 2004;20(6):593-599.

30. Appell ML, Berg J, Duley J, et al. Nomenclature for alleles of the thiopurine methyltransferase gene. *Pharmacogenet Genomics.* 2013;23(4):242-248.

31. Collie-Duguid ES, Pritchard SC, Powrie RH, et al. The frequency and distribution of thiopurine methyltransferase alleles in Caucasian and Asian populations. *Pharmacogenetics.* 1999;9(1):37-42.

32. Katara P, Kuntal H. TPMT Polymorphism: When Shield Becomes Weakness. *Interdiscip Sci.* 2016;8(2):150-155.

33. Lennard L, Cartwright CS, Wade R, Vora A. Thiopurine methyltransferase and treatment outcome in the UK acute lymphoblastic leukaemia trial ALL2003. *Br J Haematol.* 2015;170(4):550-558.

34. Winter JW, Gaffney D, Shapiro D, et al. Assessment of thiopurine methyltransferase enzyme activity is superior to genotype in predicting myelosuppression following

azathioprine therapy in patients with inflammatory bowel disease. *Aliment Pharmacol Ther.* 2007;25(9):1069-1077.

35. Hindorf U, Appell ML. Genotyping should be considered the primary choice for pre-treatment evaluation of thiopurine methyltransferase function. *J Crohns Colitis.* 2012;6(6):655-659.

36. Schmiegelow K, Forestier E, Kristinsson J, et al. Thiopurine methyltransferase activity is related to the risk of relapse of childhood acute lymphoblastic leukemia: results from the NOPHO ALL-92 study. *Leukemia.* 2009;23(3):557-564.

37. Schaeffeler E, Fischer C, Brockmeier D, et al. Comprehensive analysis of thiopurine S-methyltransferase phenotype-genotype correlation in a large population of German-Caucasians and identification of novel TPMT variants. *Pharmacogenetics.* 2004;14(7):407-417.

38. Schmiegelow K, Al-Modhwahi I, Andersen MK, et al. Methotrexate/6-mercaptopurine maintenance therapy influences the risk of a second malignant neoplasm after childhood acute lymphoblastic leukemia: results from the NOPHO ALL-92 study. *Blood.* 2009;113(24):6077-6084.

39. Booth RA, Ansari MT, Loit E, et al. Assessment of thiopurine S-methyltransferase activity in patients prescribed thiopurines: a systematic review. *Ann Intern Med.* 2011;154(12):814-823, W-295-818.

40. Levinsen M, Rosthoj S, Nygaard U, et al. Myelotoxicity after high-dose methotrexate in childhood acute leukemia is influenced by 6-mercaptopurine dosing but not by intermediate thiopurine methyltransferase activity. *Cancer Chemother Pharmacol.* 2015;75(1):59-66.

41. Tamm R, Magi R, Tremmel R, et al. Polymorphic variation in TPMT is the principal determinant of TPMT phenotype: a meta-analysis of three genome-wide association studies. *Clin Pharmacol Ther.* 2016.

42. Ford L, Prout C, Gaffney D, Berg J. Whose TPMT activity is it anyway? *Ann Clin Biochem.* 2004;41(Pt 6):498-500.

43. Lennard L, Chew TS, Lilleyman JS. Human thiopurine methyltransferase activity varies with red blood cell age. *Br J Clin Pharmacol.* 2001;52(5):539-546.

44. Yan L, Zhang S, Eiff B, et al. Thiopurine methyltransferase polymorphic tandem repeat: genotype-phenotype correlation analysis. *Clin Pharmacol Ther.* 2000;68(2):210-219.

45. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *BMJ.* 2006;333(7568):597-600.

46. Almoguera B, Vazquez L, Connolly JJ, et al. Imputation of TPMT defective alleles for the identification of patients with high-risk phenotypes. *Front Genet.* 2014;5:96.

47. Relling MV, Gardner EE, Sandborn WJ, et al. Clinical pharmacogenetics implementation consortium guidelines for thiopurine methyltransferase genotype and thiopurine dosing: 2013 update. *Clin Pharmacol Ther.* 2013;93(4):324-325.

48. Lennard L, Van Loon JA, Lilleyman JS, Weinshilboum RM. Thiopurine pharmacogenetics in leukemia: correlation of erythrocyte thiopurine methyltransferase activity and 6-thioguanine nucleotide concentrations. *Clin Pharmacol Ther.* 1987;41(1):18-25.

49. Lennard L, Gibson BE, Nicole T, Lilleyman JS. Congenital thiopurine methyltransferase deficiency and 6-mercaptopurine toxicity during treatment for acute lymphoblastic leukaemia. *Arch Dis Child.* 1993;69(5):577-579.

50. Lennard L, Lilleyman JS, Van Loon J, Weinshilboum RM. Genetic variation in response to 6-mercaptopurine for childhood acute lymphoblastic leukaemia. *Lancet.* 1990;336(8709):225-229.

51. Dokmanovic L, Urosevic J, Janic D, et al. Analysis of thiopurine S-methyltransferase polymorphism in the population of Serbia and Montenegro and mercaptopurine therapy tolerance in childhood acute lymphoblastic leukemia. *Ther Drug Monit.* 2006;28(6):800-806.

52. Relling MV, Rubnitz JE, Rivera GK, et al. High incidence of secondary brain tumours after radiotherapy and antimetabolites. *Lancet.* 1999;354(9172):34-39.

53. Relling MV, Yanishevski Y, Nemec J, et al. Etoposide and antimetabolite pharmacology in patients who develop secondary acute myeloid leukemia. *Leukemia.* 1998;12(3):346-352.

54. Pui CH, Pei D, Sandlund JT, et al. Long-term results of St Jude Total Therapy Studies 11, 12, 13A, 13B, and 14 for childhood acute lymphoblastic leukemia. *Leukemia.* 2010;24(2):371-382.

55. Relling MV, Pui CH, Cheng C, Evans WE. Thiopurine methyltransferase in acute lymphoblastic leukemia. *Blood.* 2006;107(2):843-844.

56. Rocha JC, Cheng C, Liu W, et al. Pharmacogenetics of outcome in children with acute lymphoblastic leukemia. *Blood.* 2005;105(12):4752-4758.

57. Moriyama T, Relling MV, Yang JJ. Inherited genetic variation in childhood acute lymphoblastic leukemia. *Blood.* 2015;125:3988-3995.

58. Schmiegelow K, Forestier E, Hellebostad M, et al. Long-term results of NOPHO ALL-92 and ALL-2000 studies of childhood acute lymphoblastic leukemia. *Leukemia.* 2010;24(2):345-354.

59. Schmiegelow K, Bjork O, Glomstein A, et al. Intensification of mercaptopurine/methotrexate maintenance chemotherapy may increase the risk of relapse for some children with acute lymphoblastic leukemia. *J Clin Oncol.* 2003;21(7):1332-1339.

60. Thomsen JB, Schrøder H, Kristinsson J, et al. Possible carcinogenic effect of 6-mercaptopurine on bone marrow stem cells. *Cancer.* 1999;86(6):1080-1086.

61. Levinsen M, Rotevatn EØ, Rosthøj S, et al. Pharmacogenetically based dosing of thiopurines in childhood acute lymphoblastic leukemia: Influence on cure rates and risk of second cancer. *Pediatric Blood & Cancer.* 2014;61(5):797-802.

62. Lennard L, Cartwright CS, Wade R, Vora A. Thiopurine dose intensity and treatment outcome in childhood lymphoblastic leukaemia: the influence of thiopurine methyltransferase pharmacogenetics. *Br J Haematol.* 2015;169(2):228-240.

63. Stanulla M, Schaeffeler E, Flohr T, et al. Thiopurine methyltransferase (TPMT) genotype and early treatment response to mercaptopurine in childhood acute lymphoblastic leukemia. *JAMA.* 2005;293(12):1485-1489.

64. Stanulla M, Schaeffeler E, Moricke A, et al. Thiopurine methyltransferase genetics is not a major risk factor for secondary malignant neoplasms after treatment of childhood acute lymphoblastic leukemia on Berlin-Frankfurt-Munster protocols. *Blood.* 2009;114(7):1314-1318.

65. Liu YP, Wu HY, Yang X, et al. Association between thiopurine S-methyltransferase polymorphisms and thiopurine-induced adverse drug reactions in patients with inflammatory bowel disease: a meta-analysis. *PLoS One.* 2015;10(3):e0121745.

66. de Jong DJ, Derijks LJ, Naber AH, Hooymans PM, Mulder CJ. Safety of thiopurines in the treatment of inflammatory bowel disease. *Scand J Gastroenterol Suppl.* 2003(239):69-72.

67. Hindorf U, Lindqvist M, Peterson C, et al. Pharmacogenetics during standardised initiation of thiopurine treatment in inflammatory bowel disease. *Gut.* 2006;55(10):1423-1431.

68. Ansari A, Arenas M, Greenfield SM, et al. Prospective evaluation of the pharmacogenetics of azathioprine in the treatment of inflammatory bowel disease. *Aliment Pharmacol Ther.* 2008;28(8):973-983.

69. Derijks LJ, Gilissen LP, Engels LG, et al. Pharmacokinetics of 6-mercaptopurine in patients with inflammatory bowel disease: implications for therapy. *Ther Drug Monit.* 2004;26(3):311-318.

70. Meggitt SJ, Gray JC, Reynolds NJ. Azathioprine dosed by thiopurine methyltransferase activity for moderate-to-severe atopic eczema: a double-blind, randomised controlled trial. *Lancet.* 2006;367(9513):839-846.

71. van den Akker-van Marle ME, Gurwitz D, Detmar SB, et al. Cost-effectiveness of pharmacogenomics in clinical practice: a case study of thiopurine methyltransferase genotyping in acute lymphoblastic leukemia in Europe. *Pharmacogenomics.* 2006;7(5):783-792.

72. Sayani FA, Prosser C, Bailey RJ, Jacobs P, Fedorak RN. Thiopurine methyltransferase enzyme activity determination before treatment of inflammatory bowel disease with azathioprine: effect on cost and adverse events. *Can J Gastroenterol.* 2005;19(3):147-151.

73. Payne K, Newman WG, Gurwitz D, Ibarreta D, Phillips KA. TPMT testing in azathioprine: a 'cost-effective use of healthcare resources'? *Personalized Medicine.* 2008;6(1):103-113.

74. Marra CA, Esdaile JM, Anis AH. Practical pharmacogenetics: the cost effectiveness of screening for thiopurine s-methyltransferase polymorphisms in patients with rheumatological conditions treated with azathioprine. *J Rheumatol.* 2002;29(12):2507-2512.

75. Thompson AJ, Newman WG, Elliott RA, Roberts SA, Tricker K, Payne K. The cost-effectiveness of a pharmacogenetic test: a trial-based evaluation of TPMT genotyping for azathioprine. *Value Health.* 2014;17(1):22-33.

76. Fargher EA, Tricker K, Newman W, et al. Current use of pharmacogenetic testing: a national survey of thiopurine methyltransferase testing prior to azathioprine prescription. *J Clin Pharm Ther.* 2007;32(2):187-195.

77. Zelinkova Z, Derijks LJ, Stokkers PC, et al. Inosine triphosphate pyrophosphatase and thiopurine s-methyltransferase genotypes relationship to azathioprine-induced myelosuppression. *Clin Gastroenterol Hepatol.* 2006;4(1):44-49.

Table 1. Summary of selected studies which investigated the influence of TPMT pharmacogenetics on thiopurine treatment response in childhood acute lymphoblastic leukemia.

Trial Name	Design relative to TPMT genotype	6-MP Treatment Strategy	6-MP Dosing (Maintenance Phase)	Genotypes Included	Finding	Conclusion	Author, publication year (reference NO.)
Total Therapy XII	Observational / Retrospective	Individualized therapy (dose reduction in patients experiencing myelosuppression to the highest tolerable dose)	75mg/m^2/day with selective dose reduction in patients based on clinical tolerance.	*TPMT *2, *3A, *3B and *3C*	Highest cumulative incidence of toxicity-induced reductions among patients homozygous for mutant *TPMT* alleles (100%), intermediate among heterozygous (35%), and lowest among wild-type patients (7%).	*TPMT* genotype is an important determinant of 6-MP toxicity and tolerance.	Relling et al. 1999 (10)
					- Reduced TPMT activity tends to be associated with improved EFS. - 6-MP dose intensity is a significant predictor of EFS.	Treatment should aim to administer maximal tolerable protocol dose of 6-MP, particularly to patients with wild-type TPMT activity.	Relling et al. 1999 (22)
					Reduced TPMT activity tends to associate with higher risk of secondary AML	TPMT-deficient genotype may be associated with an increased risk of secondary AML.	Relling et al. 1998 (50)
					Significantly higher cumulative risk of brain tumours among patients with reduced activity *TPMT* polymorphisms, especially when combined with high doses of antimetabolites.	TPMT-deficient genotype may be associated with an increased risk of radiation-associated brain tumours.	Relling et al. 1999 (49)

Total Therapy XIIIB	Individualized / Prospective	Individualized therapy (dose reduction in patients experiencing myelosuppression to the highest tolerable dose) with pharmacogenetics compartment	75 mg/m²/day with selective dose reduction in patients with low or intermediate TPMT activity based on a strategy that involves up-front knowledge of TPMT status combined with clinical tolerance and measurement of thiopurine metabolites	*TPMT *2, *3A, *3B, and *3C*	No association between hematologic relapse (or other long-term outcomes) and TPMT status.	Considering pharmacogenetics of *TPMT* for dose adjustment of 6-MP can help reduce treatment toxicity while not compromising its efficacy.	Relling et al. 2006 (52) Rocha et al. 2005 (53)
NOPHO-ALL-92	Observational / Retrospective	Individualized randomized maintenance therapy based on clinical response and levels of metabolites.	75 mg/m²/day with subsequent dose adjustment to a target WBC	*TPMT *3A, *3B, and *3C*	- Higher risk of relapse in patients with high TPMT activity. - Increased leukemogenic risk and higher rate of SMN in patients with low activity.	No difference in OS between low vs. high TPMT activity groups as the improved EFS is offset by the higher risk of SMN.	Thomsen, J.B., et al. 1999 (57). Schmiegelow et al. 2003 (56). Schmiegelow et al. 2009 (33). Schmiegelow et al. 2009 (35). Schmiegelow et al. 2010 (55).
NOPHO-ALL-2000	Individualized / Prospective	Individualized randomized maintenance therapy based on clinical response, levels of metabolites and pharmacogenetics.	75 mg/m²/day for wild-type patients, 50 mg/m²/day for heterozygous patients and 5–10 mg/m²/day for TPMT deficient patients, with subsequent dose adjustment to a target WBC during the first year of maintenance therapy	TPMT *3A, *3B, and *3C	- Reduced risk of SMN in heterozygous patients compared to ALL-92. -Similar relapse risk between low TPMT activity patients and wild-type patients. - No difference in overall EFS (any event) or OS compared to previous protocol.	Initial 6-MP dose adjustment based on *TPMT* genotype did reduce the risk of SMN in heterozygous patients but at the expense of an increased risk of relapse.	Schmiegelow et al. 2010 (55). Levinsen et al. 2014 (58)

UK-ALL97 and ALL97/99	Individualized	Individualized randomized maintenance therapy based on TPMT activity and clinical response.	75 mg/m²/day for both *TPMT* heterozygous and wild-type patients subsequently titrated to toxicity. Patients with TPMT deficiency were titrated from a starting dose of 10% of protocol dose (7·5 mg/m²/day).	*TPMT *2, *3A, *3B, and *3C*	-Patients with *TPMT*1/*3A* genotype had better EFS than both wild-type *TPMT*1/*1* and *TPMT*1/*3C* patients. -Heterozygosity was not associated with a higher rate of second cancers. - Higher hematotoxicity in heterozygous patients compared to wild-type patients.	While, heterozygous patients had significantly more cytopenias and required dose adjustments more often than wild-type patients, *TPMT*1/*3A* patients had better EFS.	Lennard et al. 2015 (60)
UK-ALL-2003	Individualized / Prospective	Risk stratification based on MRD and individualized randomized maintenance therapy based on clinical response, TPMT activity and pharmacogenetics.	75 mg/m²/day for both *TPMT* heterozygous and wild-type patients subsequently titrated to toxicity. Patients with TPMT deficiency were titrated from a starting dose of 10% of protocol dose (7·5 mg/m²/day).	*TPMT*2, *3A, *3B, *3C and *9*	- Improved overall EFS (all *TPMT* genotypes confounded) compared to previous protocol. - No difference in EFS, RFS or OS between TPMT genotypes. - The only significant factor affecting EFS was MRD status.	Refinements of the risk stratification process and treatment strategies have reduced the influence of *TPMT* genotype on treatment outcome.	Lennard et al. 2015 (30)

ALL-BFM-2000	Observational / Prospective	Measurement of minimal residual disease load before and after 6-MP treatment in heterozygous vs. wild-type patients	- During consolidation phase, on treatment day 78, heterozygous and wild-type homozygous patients received a 4-week cycle of (60 mg/m²/day of 6-MP). Patients homozygous for mutant *TPMT* alleles were excluded from the analyses. - During maintenance phase, 50 mg/m²/day of 6-MP. The doses of Methotrexate and 6-MP were adjusted to a target WBC.	TPMT*2, *3A, *3B, *3C and *3D + direct sequencing for individuals with an intermediate activity not having any mutation in exon 5, 7 and 10	- Significant reduction in the risk of having detectable MRD in *TPMT* heterozygous patients after induction consolidation treatment compared to wild-type patients. - No difference in the rate of hematopoietic toxicity between TPMT heterozygous carriers and homozygous wild-type carriers. - No association between low TPMT activity and risk of SMN.	- *TPMT* genotype has a significant impact on MRD during induction consolidation treatment as heterozygous patients showed a better MRD response indicating an increased clearance of disease likely due to higher intensity of 6-MP effect. - *TPMT* polymorphisms do not influence the risk of developing SMN in ALL patients treated with BFM protocols.	Stanulla et al. 2005 (61) Stanulla et al. 2009 (62)

Abbreviations: TPMT, thiopurine S-methyltransferase; 6-MP, 6-mercaptopurine; EFS, event-free survival; RFS, relapse-free survival; OS, overall survival; WBC, white blood cells; SMN, secondary malignant neoplasms; MRD, minimal residual disease; AML, Acute Myeloid Leukemia.

Table 2. Summary of selected studies which investigated the influence of TPMT pharmacogenetics on thiopurine treatment response in inflammatory bowel disease.

Author and publication year (reference NO.)	Design relative to TPMT genotype	Treatment Strategy	Genotypes Included	Findings and Conclusions
Derijks *et al.* 2004 (67)	Observational	6-MP as a single oral 50-mg evening dose	*TPMT *2, *3A, *3B and *3C*	*TPMT* genotype correlated with 6-TGN concentrations. Patients with mutant alleles have higher risk of developing leukopenia.
Hindrof *et al.* 2006 (65)	Observational	Dose escalation schedule to the target dose of (2.5 mg/kg) for azathioprine and (1.25 mg/kg) for 6-MP by week 3	*TPMT *2, *3A, *3B, *3C, *3D, *4, *5, *6, *7, *8, *10, *14 and *15*	Overall thiopurine-related ADRs were significantly more common among patients with low to intermediate TPMT activity; particularly myelotoxicity in TPMT-deficient patients.
Ansari *et al.* 2008 (66)	Observational	AZA was started as near 2 mg/kg daily as possible and without dose alteration	*TPMT *3A, *3B and *3C*	Heterozygous *TPMT* genotype strongly predicts treatment withdrawal due to early-onset of ADRs. Gastric intolerance was the most frequent reason for withdrawal among heterozygous patients. Myelotoxicity and Gastric intolerance occurred significantly more frequently among heterozygous than with *TPMT* wild-type patients.
Newman *et al.* 2011 "TARGET Trial" (15)	Individualized /Prospective	***Arm-1:*** Standard dosing without genotyping ***vs.*** ***Arm-2:*** Pre-treatment *TPMT* genotyping and AZA dosing. Wild-type (1.5–3 mg/kg/day); heterozygous (25–50 mg/day) and titrate to the maintenance dose; homozygous for *TPMT* variant alleles were given alternative drugs	*TPMT *2, *3A, *3B, and *3C*	No differences between the two study arms or between heterozygous and wild-type homozygous patients with respect to the rate of stopping azathioprine due to ADRs. No difference in the rate of remission between the intervention and control groups
Coenen *et al.* 2015 "TOPIC Trial" (21)	Individualized /Prospective	***Arm-1:*** Control group. No genotyping + standard dosing: 2–2.5 mg/kg/day AZA or 1–1.5 mg/kg/day 6-MP. ***Arm-2:*** Intervention group. Genotyping + individualized dosing. *TPMT* wild-type: same as control; heterozygous: 50% of control and homozygous variant: 0-10% of control	*TPMT *2, *3A, and *3C*	No significant overall impact of TPMT-genotype guided dosing of thiopurines on treatment efficacy or on the risk of hematologic ADRs (i.e. leukopenia and thrombocytopenia) between the genotyped and not-genotyped arms. Carriers of at least one genetic variant in the pharmacogenetics arm had a significant reduction in the risk of hematologic ADRs compared to same group in the conventional arm.

Abbreviations: TPMT, thiopurine S-methyltransferase; 6-MP, 6-mercaptopurine; AZA, azathioprine; 6-TGN, 6-thioguanine nucleotides; ADRs: adverse drug reactions.

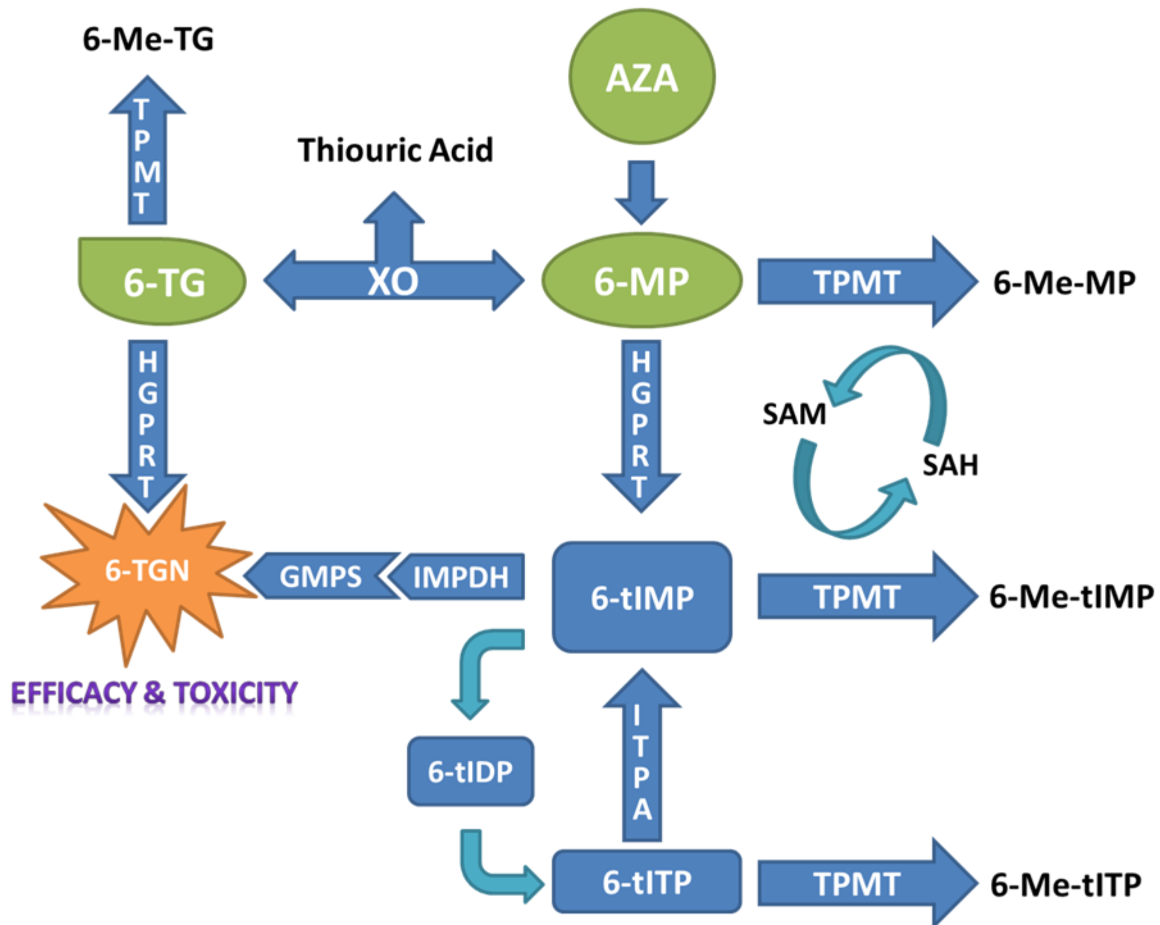


Figure 1. Metabolic pathways involved in the mechanism of action of thiopurines.

This figure illustrated the pathways involved in the metabolism of Azathioprin, 6-Mercaptopurine and Thioguanine highlighting the genes/enzymes that can potentially affect the metabolism of these drugs.

Abbreviations: AZA, Azathioprine; 6-MP, 6-Mercaptopurine; 6-TG, Thioguanine; GMPS, guanosine monophosphatase synthetase; HGPRT, hypoxanthine guanine phosphoribosyl transferase; IMPDH, inosine monophosphate dehydrogenase; ITPA, inosine triphosphate pyrophosphatase; SAH, S-adenosyl-L-homocysteine; SAM, S-adenosyl-L-methionine; TPMT, thiopurine methyltransferase; XO, xanthine oxidase; 6-Me-MP, 6-methyl-mercaptopurine; 6-Me-TG, 6-methyl-thioguanine; 6-Me-tIMP, 6-methyl-thioinosine-monophosphate; 6-Me-tITP, 6-methyl-thioinosine-triphosphate; 6-TGN, 6-thioguanine nucleotides; 6-tIDP, 6-thio-inosine diphosphate; 6-tIMP, 6-thio-inosine monophosphate; 6-tITP, 6-thio-inosine triphosphate.

Section B

Chapter 3

Pharmacogenetics of Asparaginase in Acute Lymphoblastic Leukemia

This chapter presents a comprehensive review of the published literature that investigated the pharmacogenomics of Asparaginase in the context of acute lymphoblastic leukemia treatment and summarizes the results of elegant studies carried by several internationally recognized research groups while shedding light on their reproducibility and clinical utility. It can serve as an elaborate introduction that prepares the readers for the original work that will be presented in the two following chapters (Chapter 4 and Chapter 5). It discusses the most interesting associations between genetic variants and common Asparaginase complications, highlighting the lack of external replication analysis for most findings and the conflicting, inconclusive results among the few studies that targeted the same genotype-phenotype associations.

This work was featured in the special issue of the *Cancer Drug Resistance* journal, titled "Pharmacogenetics of Cancer". This entire work was performed by me (under the supervision and guidance of Dr. Maja Krajinovic) including the literature review, selecting candidate articles to be included, and drafting the manuscript.

Review

Pharmacogenetics of asparaginase in acute lymphoblastic leukemia

Rachid Abaji[1,2], Maja Krajinovic[1,2,3]

[1]Charles Bruneau Cancer Centre, Research Center, CHU Sainte-Justine, University of Montreal, Montreal, QC, Canada. H3T 1C5
[2]Department of Pharmacology & Physiology, University of Montreal, Montreal, QC, Canada. H3C 3J7
[3]Department of Pediatrics, University of Montreal, Montreal, QC, Canada. H3T 1C5

Correspondence to: Prof. Maja Krajinovic, Departments of Pediatrics and Pharmacology, University of Montreal & Research Center, CHU Sainte-Justine, 3175 chemin de la Côte-Ste-Catherine, Montreal, QC H3T 1C5, Canada. E-mail: maja.krajinovic@umontreal.ca

How to cite this article: Abaji R, Krajinovic M. Pharmacogenetics of asparaginase in acute lymphoblastic leukemia. *Cancer Drug Resist* 2019;2:xx. DOI: 10.20517/cdr.2018.24

Received: 8 Nov 2018 | **First Decision**:12 Nov 2018 | **Revised**: 21 Jan 2019 | **Accepted**: 21 Feb 2019 | **Published**:

Science Editor: Enrico Mini **Copy Editor**: Cai-Hong Wang

3.1. Abstract

Asparaginase is a key component in leukemias and lymphomas treatment protocols and is suggested as a treatment for other malignancies in which an amino acid depletion strategy is indicated. Asparaginase intolerance is subject to inter-individual variability and can manifest as hypersensitivity reactions, pancreatitis, thrombosis as well as metabolic abnormalities, and may affect treatment outcome. Pharmacogenetics aims at enhancing treatment efficacy and safety by better understanding the genetic basis of variability and its effect on the pharmacological responses. Many groups tried to tackle the pharmacogenetics of asparaginase but the potential implementation of such findings remains debatable. In this review, we highlight the most important findings reported in studies of the pharmacogenetics of asparaginase related complications and treatment outcome in acute lymphoblastic leukemia (ALL).

Key Words: Asparaginase; Pharmacogenomics; Hypersensitivity Reactions; Pancreatitis; Relapse; Acute Lymphoblastic Leukemia; Adverse Drug Reactions.

3.2. General Introduction

3.2.1. Asparaginase and Acute Lymphoblastic Leukemia

L-asparaginase (ASNase) is a key component in leukemias and lymphomas treatment strategies and is universally incorporated into major childhood acute lymphoblastic leukemia (ALL) treatment protocols.[1-4] ALL in adults has a much lower incidence than in children, and poor survival rates in this population pose a significant challenge.[5] The incorporation of ASNase into adults and young adult protocols is still limited due to its toxicity profile in this population.[6] On the other hand, the introduction of ASNase into pediatric regimens for ALL treatment and the intensification of its use, along with dexamethasone and vincristine (VCR), is to be credited for most of the improvement in ALL treatment outcome.[1] A typical ALL treatment protocol consists of phases that focus on remission-induction, consolidation and maintenance. ASNase is usually administered during the induction phase as well as throughout the consolidation therapy where it is administered in for 20-30 weeks together with glucocorticoids and vincristine.[5, 7]

ALL accounts for 30% of pediatrics cancers and is the most common childhood malignancy in developed and underdeveloped countries.[1, 8, 9] The past few decades have witnessed a revolution in the treatment of ALL as survival rates increased considerably from less than 40% in the mid-sixties to currently exceed 90% for most international protocols.[1, 2, 10-12] This result was achieved by the creation and continuous

optimization of multi-agent protocols through evidence based medicine, refined stratification of patients into risk groups, personalized chemotherapy that exploit the differences in the characteristics between host and leukemia cells and improvement in supportive care.[5, 12-14] While these figures seem quite encouraging, there is a large margin for improvement as treatment failure, cancer relapse and treatment-related toxicities continue to jeopardize the lives of a significant percentage of children with ALL.[8] It is estimated that almost 50% of patients will experience at least one acute severe toxicity, and that a considerable percentage of mortality among leukemia patients is attributable to adverse-events of the treatment rather than the actual sickness.[2, 12, 15] In fact, these toxicities can often be life-threatening and are the primary cause of interruption or discontinuation of chemotherapy [10] and are a frequent cause of sequelae on the long-term.[2] Indeed, the recent improvement in survival rate has resulted in a gradual shift towards putting more focus on reducing the toxicity burden of chemotherapy.[2, 15]

Consequently, several research groups are investigating biomarkers that can predict the risk of treatment resistance or treatment-related adverse effects even before starting the therapy in the hope to modify the treatment in a patient-tailored manner that would increase the probability of response and reduce the risk of side-effects. This is the core goal of pharmacogenetics (PGx) which aims at enhancing treatment efficacy and safety by providing a better understanding of the genetic basis of variability and its effect on the pharmacological responses.[14, 16] Indeed, there are several success stories in which PGx discoveries have restructured the medical practice and the classical example is the genotyping of *TPMT* gene to guide the dosing of mercaptopurine which is considered

mandatory in almost all recent practice guideline.[17] Accordingly, many groups studied the pharmacogenetics of ASNase aiming to uncover the genes mediating ASNase antileukemia effect and the genetic basis for interpatient variability of response. However, the implementation of such findings in ALL management remains debatable. In this review, we highlight the most important findings reported up-to-date which tackled the PGx of ASNase-related complications and treatment outcome. We used different search-engine tools –but mainly the ones embedded in the NCBI platform- to identify eligible scientific papers that included the word asparaginase along with either the term pharmacogenomics or pharmacogenetics. Upon evaluating the content of these papers, a filtering step was applied in order to retain only the articles that specifically addressed the PGx of ASNase, which are summarized in table 1.

3.2.2. ***Mechanism of Action, Resistance & Formulations***

The exact mechanism of the anti-leukemic effect of ASNase is still not fully understood. However, it is generally accepted that this enzyme works by hydrolysing asparagine –and glutamine- in the serum, thus depleting the extracellular compartment from these amino acids essential for survival of all cells.[7, 10, 14, 18, 19] Asparagine is produced by the enzyme asparagine synthetase, encoded by the *ASNS* gene, which catalyzes the transfer of an amino group to aspartic acid to form asparagine and may thus counteract the effect of asparaginase and produce resistance as suggested by in vitro experiments conducted in leukemia cell lines and patient lymphoblasts.[5, 10, 18]

It has been hypothesized that malignant lymphoblasts have low expression of the *ASNS* gene, or alternatively, are incapable of upregulating the expression of *ASNS* when exposed to ASNase or nutritive stress; subsequently making them unable to produce enough asparagine or glutamine to meet the high demand required for their rapid growth. This renders the leukemic cells more dependent on extracellular sources of asparagine and thus more sensitive to the effect of ASNase which hence selectively kills them by depleting the media of asparagine, leading to amino acid starvation and disrupting the biosynthesis of proteins and eventually cellular apoptosis and death.[2, 7, 10, 14, 18, 19]

As for glutamine, ASNase-resistant lymphoma cells were demonstrated to have a substantial increase in glutamine synthetase activity compared to ASNase-sensitive cells - consequently increasing their production of glutamine; and thus, their proliferation capacity was less affected by low levels of extracellular glutamine.[20] Moreover, it was also shown that the transport of glutamine into the ASNase-resistant cells was significantly elevated due to an adaptive regulation response.[20] Furthermore, in a study that evaluated the effect of ASNase on glutamine-dependant lymphoid cell lines, the authors reported a relationship between cells’ sensitivity and the expression pattern of molecules involved in glutamine and asparagine metabolism.[21]

The *in vitro* and *in vivo* sensitivity to ASNase have been associated with childhood ALL prognosis.[14, 19] Inter-individual differences in *ASNS* expression levels and ALL sensitivity to ASNase were noted, which might be explained by a change in expression of *ASNS* gene itself, or genes coding for the regulators of its expression (e.g. The basic region leucine zipper *activating transcription factor 5*, *ATF5*; and *arginosuccinate synthase* 1, *ASS1*). Nonetheless, the body of evidence reporting on the associations between *ASNS* activity and ASNase resistance is conflicting.[5, 14, 18, 19] Other causes of resistance include the formation of ASNase inactivating antibodies, the secretion of asparagine from mesenchymal cells in the bone marrow, or altered expression in genes involved in apoptosis.[14, 19] A study that tested almost 2.4 million SNPs in a genome-wide association study (GWAS) approach using the HapMap lymphoblastoid cell-line, identified aspartate metabolic pathway as a contributor to ASNase sensitivity with primary involvement of *ADSL* and *DARS* genes. The authors were also able to reproduce significant associations in primary ALL leukemic blasts.[19]

Historically, three asparaginase preparations were commercially available and each of them has different pharmacokinetic properties. The original preparation was derived from *Escherichia coli* (and is referred to as *E.Coli* asparaginase), but it has been abandoned by most developed countries due to its toxicity profile (particularly allergic reactions), and the adoption of its less immunogenic pegylated form (PEG-asparaginase). While PEG-asparaginase is relatively more expensive than its parent-compound, it is considered to be a safer and more effective treatment with a prolonged duration of activity. The third product is a formulation derived from *Erwinia chrysanthemi* (*Erwinia* asparaginase) and is generally

associated with lower immunogenic properties and less toxicity. However, its pharmacokinetic profile was reported to be associated with poorer treatment outcome when compared to other formulations at a similar posology (mainly attributed to its shorter duration of activity), suggesting the need for higher doses and increased frequency of administration in order to achieve optimal asparagine depletion. Thus, its use is usually restricted to patients who develop allergic reactions *"or silent inactivation"* to the *E.Coli* /PEG-asparaginase owing to the lack of cross-reactivity,[1, 5, 7, 8, 10, 22] although it is important to mention that controversies on ASNase antibodies formation and its activity has been reported.[23] Several clinical trials have reported associations between success of ALL treatment and ASNase dose intensity or formulation.[10, 22] Of note, enzyme variants with reduced l-glutaminase coactivity are being tested for their clinical utility as antileukemic agents with potentially lower side effects (since several studies suggested that the depletion of l-glutamine may correlate with many of the side effects of the enzyme). For example, a recent study demonstrated that novel low l-glutaminase variants derived from modifications to *Erwinia* asparaginase can offer high efficacy against both T-Cell and B-Cell ALL while provoking less toxicities.[24]

3.2.3. ***Pharmacogenetics of Asparaginase***

3.2.3.1. ***Hypersensitivity Reactions, Pancreatitis & Thrombosis***

Since ASNase is a foreign protein produced in bacteria, it is not surprising that all formulations of ASNase, to varying extents, have the immunogenic potential to provoke the formation of antibodies which can be associated with clinical symptoms manifested in ASNase allergy and hypersensitivity reactions (HSRs), or can be asymptomatic but still capable of neutralizing the activity of ASNase leading to suboptimal response and thus referred to as 'silent hypersensitivity' or 'silent inactivation' which occurs in up to 30% of patients.[1, 8, 9, 12, 25] While the allergic symptoms can be mitigated through premedication with anti-histamines and corticosteroids, this still does not prevent ASNase inactivation.[12] It is important to mention that higher systemic exposure to ASNase was associated with a lower clearance of dexamethasone, and thus a higher systemic exposure and an increased risk of osteonecrosis. Nonetheless, studies also found that the formation of ASNase antibodies can increase the systemic clearance of dexamethasone, consequently reducing its serum levels and increasing the risk of Central Nervous System (CNS) relapse.[1, 26-28]

HSRs are the most common side-effect and can manifest as pain around the injection site, urticaria, flushing, fever, chills, dyspnea, bronchospasm edema/angioedema, and hypotension. They could arguably occur in as much as 75% of patients and could manifest as life-threatening anaphylactic reactions in 10% of them and usually require changing the drug formulation.[2, 5, 7-10, 12, 18, 25, 29] The incidence is dependent on different factors

which include the number of doses received, route of administration, type of formulation used, re-challenging after a period of interruption, and the administration of concomitant medications during the course of treatment.[1, 8, 9, 12, 25, 30]

One of the pioneer studies in the context of PGx of HSRs was a GWAS which aimed at identifying germline genetic polymorphisms that could contribute to the risk of allergy in an ethnically diverse population of 485 ALL children treated with ASNase on St. Jude Children's Research Hospital treatment protocol Total Therapy XV. They interrogated over 500,000 single nucleotide polymorphisms (SNPs) and had many significant hits. Essentially, the results demonstrated an overrepresentation of SNPs in genes located on chromosome 5q33 in general (which is already known to be associated with several inflammatory or autoimmune diseases), and in the *GALNT10* and *GRIA1* genes in particular. Indeed, the associations of five of the polymorphisms (i.e. rs4958381, rs10070447, rs6890057, rs4958676, and rs6889909) in *GRIA1* gene with HSRs were successfully validated in the same study in an independent replication cohort [25] and were later replicated in an independent Slovenian population of 146 pediatric ALL patients mainly treated according to one of Berlin-Frankfurt-Münster (BFM) treatment protocols.[31] Moreover, the authors reported an association between the frequency of ASNase allergy and racial ancestry; with patients of Caucasian origins developing allergic reactions at a higher frequency than those of black or Hispanic ones.[25] Another group tried to replicate the results by targeting 20 SNPs in *GRIA1* and *GALNT10* genes in a candidate-gene fashion in a group of Hungarian ALL children treated as part of the BFM Study Group. Briefly, they were unable to replicate any

of the results in the total cohort. However, interestingly, they found an opposite association between rs4958381 in *GRIA1* and reduced risk of HSRs in the T-cell ALL subgroup but not in the pre-B-cell ALL patients. Moreover, they reported significant associations of two SNPs in *GRIA1* not identified in the original work (but only in the medium risk group), which can still serve as a further evidence of the implication of the *GRIA1* gene in the modulation of the risk of ASNase induced HSRs and might suggest that the influence can vary depending on subgroups.[9]

In another study that involved a total of 1870 patients of European ancestry, the authors imputed human leukocyte antigen (HLA) alleles and searched for significant associations with ASNase hypersensitivity in childhood ALL patients from Jude Children's Research Hospital and the Children's Oncology Group. They reported a strong association of *HLA-DRB1**07:01 allele in both groups and demonstrated that *HLA-DRB1* alleles that confer high-affinity binding to ASNase epitopes contribute to the observed higher frequency of HSRs.[4]

Another GWAS was performed on a cohort of 3308 pediatric ALL patients treated according to St. Jude Children's Research Hospital (SJCRH) protocols or Children's Oncology Group protocols and demonstrated that variants within genes regulating the immune response, particularly genes involved in T-cell function, strongly influenced the risk of ASNase hypersensitivity. The authors found a strong association between a polymorphism in the nuclear factor of activated T cells 2 (NFATC2), rs6021191, and hypersensitivity to ASNase. They also reported that the association was strongest among patients receiving native E. coli ASNase as compared to PEG-ASNase and that carrier-status of this intronic variant was associated with a higher expression of the gene's messenger RNA compared to

noncarriers (both in ALL leukemic blast samples and lymphoblastoid cell-lines). Moreover, looking at the association of nonsynonymous coding variants with HSRs, they found that the most significant association was that of rs17885382 in *HLA-DRB1* which is in almost perfect linkage disequilibrium with HLA-DRB1*07:01 mentioned earlier and can be used as a confirmation of the importance of the latter in influencing the risk of ASNase hypersensitivity. Importantly, this finding extends the role of the polymorphism to non-European patients; since the new cohort was ethnically diversified as opposed to the previous one which only involved patients with European ancestry. Furthermore, the authors also demonstrated that the risk of HSRs associated with carrying the risk alleles of rs6021191 in *NFATC2* and rs17885382 in *HLA-DRB1* was additive.[30]

In a study performed on samples from 359 Hungarian childhood ALL patients treated with one of the BFM protocols and aimed at using next-generation sequencing to identify associations between ASNase hypersensitivity and polymorphisms of the Human Leukocyte Antigen (HLA) Class II region alleles, the authors further confirmed that variations in HLA-D region can influence the development of ASNase HSRs. For example, patients with *HLA-DRB1*07:01* allele or *HLA-DQB1*02:02* allele had a significantly higher risk of developing this toxicity compared to non-carriers. Moreover, a significant association with the haplotype *HLA-DRB1*07:01-HLA-DQB1*02:02* was observed as carriers of this haplotype were at higher risk than carriers of only one of the risk alleles. Furthermore, carrying the *HLA-DRB1*07:01–HLA-DQA1*02:01–HLA-DQB1*02:02* haplotype was associated with the highest risk of ASNase hypersensitivity. Of note, this study also

reported that *HLA-DQB1*02:02* allele was significantly less frequent in the proportion of patients with T-cell ALL than in pre-B-cell ALL patients.[32]

Since patients with PEG-asparaginase HSRs were demonstrated to have no ASNase enzymatic activity, a more recent study investigated genetic predisposition to PEG-asparaginase hypersensitivity in a GWAS analysis by defining the hypersensitivity phenotype as both having clinical hypersensitivity and no enzymatic activity. The genetic analysis was performed on fifty-nine cases and 772 control pediatric patients treated on the Nordic Society of Paediatric Haematology and Oncology (NOPHO) ALL2008 protocol. The study found rs73062673 polymorphism of the *CNOT3* gene to be associated with PEG-asparaginase allergy. Of note, this gene was previously shown to regulate the transcription of HLA and to act as a tumour suppressor which is frequently mutated in T-cell ALL. The study also reported the detection of two other associations involving rs9272131 and rs115360810 variants in the *HLA-DQA1* and *TAP2* genes, respectively. While these associations were not significant on a genome-wide level, they remain of a particular interest since the variants are located in a region known to be highly involved in allergic responses. These results further suggest the implication of genetic variations in the HLA region, as well as regulators of these genes, in the mechanisms leading to asparaginase hypersensitivity.[33]

Other common adverse-events to ASNase are acute pancreatitis and cerebrovascular accidents, such as thrombosis, which can occur in 18% and 5% of ALL patients, respectively; and are usually dose limiting.[2, 10, 12, 15, 18] Pancreatitis symptoms can range from being mild and self-resolving, to a more severe systemic inflammatory response

syndrome and failure of pancreatic function.[15] While the risk of mortality due to ASNase induced acute pancreatitis is relatively low, the risk of recurrence upon re-challenge is almost 50% and patients affected by it have a higher risk of developing chronic or relapsing pancreatitis as well as acute or persistent diabetes mellitus.[12, 15, 34] Clinical factors of ASNase associated pancreatitis include Native American ancestry, older age, and higher cumulative ASNase exposure.[35] While the role of genetics in predisposition to acute recurrent and/or chronic pancreatitis of different etiologies has been the focus of many studies (PRSS1, PRSS2, SPINK1, CFTR, CLDN2, CAP1),[34, 36-39] ASNase-related acute pancreatitis have only started emerging recently.

In a work that tackled the PGx of ASNase through candidate-gene approach by investigating the association between SNPs in *ASNS, ATF5* and *ASS1* genes and ASNase induced allergy and pancreatitis in a discovery cohort of 285 ALL patients and a replication cohort of 248 patients who were treated according to Dana-Farber Cancer Institute ALL Consortium protocols. The authors reported a significant association between a 14-bp tandem-repeat polymorphism *rs3832526* in *ASNS* gene and both of these toxicities as patients homozygous for the triple repeat allele (*3R*) had the complications more frequently than other genotype groups. Moreover, when analysing the effect of possible haplotypes, they found that the *ASNS* haplotype **1* harbouring double repeat (*2R*) allele conferred a protective effect from these toxicities and its association with allergy was further validated in an independent replication cohort. Furthermore, they showed that one of the subtypes of this haplotype was associated with reduced in vitro sensitivity to ASNase in lymphoblastoid cell lines.[10]

It is worth mentioning that in a study including 472 Japanese children with ALL who were treated on a protocol that included *E.*coli derived asparaginase, the authors followed a candidate-gene approach aimed at replicating the associations found with *GRIA1* rs4958351, *NFATC2* rs6021191, and *ANSN* rs3832526. The authors reported no significant associations between any of the variants and HSRs which suggests that the role of these variants might be influenced by ethnic specific differences in genetic structure surrounding them.[40]

Another work followed an exome-wide association study approach which was performed on 302 children with ALL treated according to DFCI protocols and the results were validated in an independent group of 282 patients following protocols of the same institution. The authors interrogated around 4.5 thousand SNPS distributed across 3802 genes and reported 12 associations with ASNase complications in the discovery cohort including 3 with allergy, 3 with pancreatitis and 6 with thrombosis along with a strong additive effect of combining more than one polymorphism. Interestingly, rs3809849 in the *MYBBP1A* gene was associated with allergy, pancreatitis, thrombosis, event-free survival (EFS) and overall survival while rs11556218 in *IL16* gene and rs34708521 in *SPEF2* gene were both associated with thrombosis and pancreatitis. Importantly, the association of each of these three polymorphisms with pancreatitis was replicated in the validation cohort.[41] Of note, our search results could not identify other original research work that investigated the PGx of ASNase-induced thrombosis, which could be an interesting field for future studies.

In a GWAS study of ASNase-induced pancreatitis involved ALL patients treated following St Jude Children's Research Hospital and in the Children's Oncology Group protocols. The discovery group was composed of 5,185 children and young adults with ALL and was replicated in an independent case-control group of 213 patients. While the authors reported no significant association of common variants at the GWAS level, they detected a significant association for a rare nonsense variant rs199695765 in *CPA2* gene. Interestingly, in a subsequent gene-level investigation, 16 SNPs in this gene were significantly associated with pancreatitis with around 54% of carriers of at least one of these polymorphisms ended up developing it.[35]

In another GWAS study of 700 children who were treated following the NOPHO ALL2008 protocol, the authors interrogated around 1.5 million SNPs and found 27 significant associations with ASNase related pancreatitis. rs281366 variant in *ULK2* gene showed the strongest association with pancreatitis, and interestingly, 14 of the 27 associations were of polymorphisms in this same gene. In a sub-analysis focusing on patients who were less than 10 years old, rs17179470 in *RGS6* was strongly associated with pancreatitis. Moreover, in this particular subgroup, more than half of the cases carried one of these two risk alleles and the risk of pancreatitis associated with carrying both alleles was additive. Of noteworthy, *ULK2* gene involved in autophagy, and *RGS6* regulates G-protein signaling regulating cell dynamics.[15]

In a larger and more recent multi-centric study lead by researchers from the same group, the authors investigated the risk of ASNase-associated pancreatitis in a discovery cohort of 244 cases and 1320 controls through GWAS analysis.[15] rs62228256, a variant

located in a noncoding region of the genome upstream from the *NFATC2* gene, and for which it acts as an expression quantitative trait loci (eQTL) in pancreatic tissue, had the strongest association signal with an increased risk of pancreatitis for carriers of the minor allele. However, the validation analysis in a cohort of 33 cases and 285 controls who followed one of the DFCI treatment protocols did not replicate this association. An association with pancreatitis was also detected for minor alleles of rs13228878 and rs10273639 which reside on the same haplotype and are in high linkage disequilibrium in the *PRSS1-PRSS2* locus encoding for cationic and anionic trypsinogen, respectively. The association was further confirmed in a replication analysis performed on samples from patients of the Children's Oncology Group (76 cases and 2653 controls). Of note, these variants were associated with an increase in the expression of *PRSS1* gene and they have been previously linked to alcohol-associated and sporadic pancreatitis in adults. Another interesting outcome of this study is the further validation of the association between pancreatitis risk and SNPs within genes known to regulate trypsin activation. Accordingly, minor alleles of rs17107315 in pancreatic secretory trypsin inhibitor (*SPINK1),* rs10436957 in chymotrypsin C (*CTRC),* and rs4409525 in Claudin-2 (*CLDN2)* all had significant associations with modulating the risk of ASNase-induced acute pancreatitis with directions and effects similar to the previously reported findings. The authors also applied a targeted genotyping approach to test the reproducibility of the association of the *ULK2* variant rs281366 and *RGS6* variant rs17179470 with the risk of pancreatitis previously reported by the same group but the results were not significant.[15]

3.2.3.2. ***Other Less Common Toxicities***

ASNase intolerance can also result in hepatotoxicity, abnormalities of hemostasis, hyperglycemia, hyperlipidemia, and may also affect treatment outcome since it was shown that patients who experienced a dose-limiting ASNase toxicity had a significantly worse disease-free survival.[2, 8, 10, 12, 18] However, due to the rarity of these toxicities, they were less frequently investigated. ASNase-induced hepatotoxicity is one of the most common ASNase complications in adults treated for ALL but is rarely investigated in genetic studies since most of such studies focus on the use of ASNase in pediatric patients. Given its mechanism of action, ASNase induces amino acid stress response by depleting asparagine and glutamine. This results in an excessive production of reactive oxygen species (ROS) and a subsequent increase in mitochondrial permeabilization and eventual cell apoptosis; a process that has been linked to ASNase-induced hepatotoxicity.[6] In a candidate-gene analysis that involved 190 adult ALL patients enrolled on CALGB-10102, the authors reported a significant association between homozygous carriers of the minor allele of rs4880 in *SOD2* gene, a mitochondrial enzyme that protects cells against ROS, and an increased risk of hepatotoxicity following ASNase-based treatment.[6]

*3.2.3.3. **Relapse***

Relapse is a major cause of treatment failure in pediatric ALL as it was reported to arise in 11% to 36% of patients with high-risk B-precursor ALL.[42-49] The risk of relapse and treatment toxicity can be modulated by multiple factors and differences in genetic composition among patients have recently driven considerable attention.[5, 16] Several PGx studies reported that genomic variation was associated with higher risk of relapse in ALL patients.[42, 50, 51] For example, in a GWAS that involved 2535 children with newly diagnosed ALL that aimed at targeting germline polymorphisms associated with relapse, the authors identified 5 SNPs linked to higher levels of ASNase antibodies and 2 of those were associated with a higher relapse rate.[51] In a more recent study that investigated the contribution of germline genetic factors to relapse in 2,225 children treated on Children's Oncology Group trial AALL0232, the author reported that the group of relapse SNPs in the more ASNase intensive treatment arm was overrepresented with SNPs linked to ASNase resistance or allergy.[42]

Early reports have indicated that lower exposure to ASNase during ALL treatment can result in an increased risk of relapse [52, 53] which lead a research team to hypothesising that genetic polymorphisms of genes in asparagine pathway (i.e. *ASNS*, *ATF5*, and *ASS1*) can be associated with risk of event-free survival and relapse leading to a study in 318 Caucasian children with ALL and an independent replication cohort of 267 patients.[18] Indeed, the authors identified a variant in the promotor of *ATF5* gene, rs11554772, and a higher risk of ALL relapse in patients who received *E.coli* ASNase. This

gene codes for a transcriptional factor involved in *ASNS* gene regulation. Importantly, the result was validated in the replication group and was corroborated with data on the association of the same polymorphism with higher promoter activity. Another finding was the association of a 14-bp tandem-repeat polymorphism, *rs3832526*, located in the first intron of *ASNS* gene and EFS which showed that homozygous carriers of the double repeat (2R) had a significantly lower EFS, but the association lacked significance in the validation cohort.[18] They also reported the association of polymorphisms in the ASS1 gene and EFS albeit these associations did not sustain correction for multiple testing and thus were not further investigated in the replication cohort.[18] Interestingly, the repeat polymorphism in *ASNS* gene was later linked to early response to ALL treatment following the administration of a single ASNase dose in a study of 264 Polish children with ALL. However, the association was in the opposite direction as carriers of the (3R) allele with a poor response at day 15 had an increased risk of events, hence the data suggest an interaction between this polymorphism and early response to treatment that could result in variability of EFS rates.[54]

3.2.4. MicroRNA

One area that is currently under-investigated in relation to the effect of ASNase is that of microRNAs (miRNA), with only few studies reporting associations between differences in miRNAs expression levels and response in childhood ALL.[55-57] while reports suggest that the expression of over 60% of protein coding genes is subject to regulation via miRNAs.[58] Of note, many groups linked the expression levels of specific miRNAs to clinical outcome of ALL patients. In fact, studies suggest that miRNA expression profiles can differ significantly between ALL genetic-subtypes and that drug-resistant cases are associated with unique miRNA signature. For example, one study showed that miR-454 was expressed at nearly two-fold lower levels in ASNase-resistant pediatric ALL patients when compared to ASNase-sensitive ones.[55] Another study linked miR-210 to ASNase-sensitivity as demonstrated by the expression levels dependent change in the minimum inhibitory concentration (IC_{50}), the concentration needed to block the proliferation of half of the initial cell population.[56]

3.3. General Conclusion

While it is becoming increasingly recognized that both tumor and germline genomics can influence response to treatment, the latter is less commonly used to guide treatment in oncology settings. It should be emphasized that the possibility of detecting a random signal in association studies is relatively high, which could explain the conflicting data and inconclusive results among studies that targeted the same genetic-phenotypic associations. Moreover, differences in trial settings, treatment protocols, nature of supportive care, the degree of scrutiny with which an outcomes is measured and variations in disease characteristics, among others, can influence the role of the variant in question. [10, 12] One example is the leukemic cells that carry the subtype of ALL featuring a TEL/AML1 fusion gene which were demonstrated to be more sensitive to the effect of ASNase compared to other subtypes.[59] Thus, the implication of a gene or its polymorphisms in the outcome should only be taken into consideration for clinical implementation if the association was confirmed by independent studies and further supported by functional analysis.

The translatability of pharmacogenetics findings into the clinical realm of personalized medicine remains a challenge given the complex interplay between the host and malignancy genomes. One example is the CoALL 06-97 study which incorporated a combined drug resistance profile into their risk group stratification process of 224 patients. While this profile, which was based on *in vitro* cellular resistance to prednisolone, VCR and ASNase, was previously shown to be linked to treatment response and was confirmed in several studies, the authors reported no significant difference between results of that study

and those of historical control group stratified according to conventional risk factors.[60] A lot of work needs to be done in the context of implementation of pharmacogenetics. In a study that analyzed pharmacogenomics literature of 125 drugs used in oncology, more than half of the drugs (55%) did not have pharmacogenomics data while only 12 of those which did, had actionable associations.[61] Understanding the pharmacogenetics of ASNase can help refining treatment strategies for other cancers in which asparagine and/or glutamine depletion can be indicated such as in subtypes of acute myeloid leukemia, sarcomas, pancreatic and ovarian malignancies.[62-65]

Given the recent breakthroughs in biotechnology allowing for increasingly shorter rendering time and lower costs of genotyping and sequencing services, pharmacogenetics will continue to flourish as more complex analyses will be feasible. This will enrich the pool of validated genetic markers that can predict the risk and outcome of a particular treatment and will make it possible to move away from the less-than-optimal trial-and-error approach to dosing towards the implementing PGx to guide a treatment that is tailored to the genetics of each individual.

3.4. Declarations

Acknowledgments

The Authors would like to thank the Charles Bruneau Foundation for the constant support. M. Krajinovic currently holds the grant for Cancer Research Society. R. Abaji is a scholar of the Cole Foundation and the Network of Applied Medical Genetics (RMGA) and acknowledges the support of both organizations.

Authors' contributions

R.A. conducted the literature review and drafted the article; M.K. revised the manuscript.

Availability of data and materials

Not applicable.

Financial support and sponsorship

None.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2018.

3.5. Rreferences

1. Pui CH, Evans WE. A 50-year journey to cure childhood acute lymphoblastic leukemia. *Semin Hematol* 50(3), 185-196 (2013).

2. Gervasini G, Vagace JM. Impact of genetic polymorphisms on chemotherapy toxicity in childhood acute lymphoblastic leukemia. *Front Genet* 3 249 (2012).

3. Asselin BL, Ryan D, Frantz CN *et al*. In vitro and in vivo killing of acute lymphoblastic leukemia cells by L-asparaginase. *Cancer Res* 49(15), 4363-4368 (1989).

4. Fernandez CA, Smith C, Yang W *et al*. HLA-DRB1*07:01 is associated with a higher risk of asparaginase allergies. *Blood* 124(8), 1266-1276 (2014).

5. Mei L, Ontiveros EP, Griffiths EA, Thompson JE, Wang ES, Wetzler M. Pharmacogenetics predictive of response and toxicity in acute lymphoblastic leukemia therapy. *Blood Rev* 29(4), 243-249 (2015).

6. Alachkar H, Fulton N, Sanford B *et al*. Expression and polymorphism (rs4880) of mitochondrial superoxide dismutase (SOD2) and asparaginase induced hepatotoxicity in adult patients with acute lymphoblastic leukemia. *Pharmacogenomics J* 17(3), 274-279 (2017).

7. Paugh SW, Stocco G, Evans WE. Pharmacogenomics in pediatric leukemia. *Curr Opin Pediatr* 22(6), 703-710 (2010).

8. Lopez-Santillan M, Iparraguirre L, Martin-Guerrero I, Gutierrez-Camino A, Garcia-Orad A. Review of pharmacogenetics studies of L-asparaginase hypersensitivity in acute lymphoblastic leukemia points to variants in the GRIA1 gene. *Drug Metab Pers Ther* 32(1), 1-9 (2017).

9. Kutszegi N, Semsei AF, Gezsi A *et al*. Subgroups of Paediatric Acute Lymphoblastic Leukaemia Might Differ Significantly in Genetic Predisposition to Asparaginase Hypersensitivity. *PLoS One* 10(10), e0140136 (2015).

10. Ben Tanfous M, Sharif-Askari B, Ceppi F *et al*. Polymorphisms of asparaginase pathway and asparaginase-related complications in children with acute lymphoblastic leukemia. *Clin Cancer Res* 21(2), 329-334 (2015).

11. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin* 63(1), 11-30 (2013).

12. Schmiegelow K, Muller K, Mogensen SS *et al*. Non-infectious chemotherapy-associated acute toxicities during childhood acute lymphoblastic leukemia therapy. *F1000Res* 6 444 (2017).

13. Pui CH, Mullighan CG, Evans WE, Relling MV. Pediatric acute lymphoblastic leukemia: where are we going and how do we get there? *Blood* 120(6), 1165-1174 (2012).

14. Cheok MH, Pottier N, Kager L, Evans WE. Pharmacogenetics in acute lymphoblastic leukemia. *Semin Hematol* 46(1), 39-51 (2009).

15. Wolthers BO, Frandsen TL, Abrahamsson J *et al*. Asparaginase-associated pancreatitis: a study on phenotype and genotype in the NOPHO ALL2008 protocol. *Leukemia* 31(2), 325-332 (2017).

16. Zhang G, Nebert DW. Personalized medicine: Genetic risk prediction of drug response. *Pharmacol Ther* 175 75-90 (2017).

17. Abaji R, Krajinovic M. Thiopurine S-methyltransferase polymorphisms in acute lymphoblastic leukemia, inflammatory bowel disease and autoimmune disorders: influence on treatment response. *Pharmgenomics Pers Med* 10 143-156 (2017).

18. Rousseau J, Gagne V, Labuda M *et al*. ATF5 polymorphisms influence ATF function and response to treatment in children with childhood acute lymphoblastic leukemia. *Blood* 118(22), 5883-5890 (2011).

19. Chen SH, Yang W, Fan Y *et al*. A genome-wide approach identifies that the aspartate metabolism pathway contributes to asparaginase sensitivity. *Leukemia* 25(1), 66-74 (2011).

20. Aslanian AM, Kilberg MS. Multiple adaptive mechanisms affect asparagine synthetase substrate availability in asparaginase-resistant MOLT-4 human leukaemia cells. *Biochem J* 358(Pt 1), 59-67 (2001).

21. Sugimoto K, Suzuki HI, Fujimura T *et al*. A clinically attainable dose of L-asparaginase targets glutamine addiction in lymphoid cell lines. *Cancer Sci* 106(11), 1534-1543 (2015).

22. Pui CH, Thiel E. Central nervous system disease in hematologic malignancies: historical perspective and practical applications. *Semin Oncol* 36(4 Suppl 2), S2-S16 (2009).

23. Tram Henriksen L, Gottschalk Hojfeldt S, Schmiegelow K *et al*. Prolonged first-line PEG-asparaginase treatment in pediatric acute lymphoblastic leukemia in the NOPHO ALL2008 protocol-Pharmacokinetics and antibody formation. *Pediatr Blood Cancer* 64(12), (2017).

24. Nguyen HA, Su Y, Zhang JY *et al*. A Novel l-Asparaginase with low l-Glutaminase Coactivity Is Highly Efficacious against Both T- and B-cell Acute Lymphoblastic Leukemias In Vivo. *Cancer Res* 78(6), 1549-1560 (2018).

25. Chen SH, Pei D, Yang W *et al*. Genetic variations in GRIA1 on chromosome 5q33 related to asparaginase hypersensitivity. *Clin Pharmacol Ther* 88(2), 191-196 (2010).

26. Kawedia JD, Kaste SC, Pei D *et al*. Pharmacokinetic, pharmacodynamic, and pharmacogenetic determinants of osteonecrosis in children with acute lymphoblastic leukemia. *Blood* 117(8), 2340-2347; quiz 2556 (2011).

27. Kawedia JD, Liu C, Pei D *et al*. Dexamethasone exposure and asparaginase antibodies affect relapse risk in acute lymphoblastic leukemia. *Blood* 119(7), 1658-1664 (2012).

28. Liu C, Janke LJ, Kawedia JD *et al*. Asparaginase Potentiates Glucocorticoid-Induced Osteonecrosis in a Mouse Model. *PLoS One* 11(3), e0151433 (2016).

29. Castro-Pastrana LI, Ghannadan R, Rieder MJ, Dahlke E, Hayden M, Carleton B. Cutaneous adverse drug reactions in children: an analysis of reports from the Canadian Pharmacogenomics Network for Drug Safety (CPNDS). *J Popul Ther Clin Pharmacol* 18 e106-120 (2011).

30. Fernandez CA, Smith C, Yang W *et al*. Genome-wide analysis links NFATC2 with asparaginase hypersensitivity. *Blood* 126(1), 69-75 (2015).

31. Rajic V, Debeljak M, Goricar K, Jazbec J. Polymorphisms in GRIA1 gene are a risk factor for asparaginase hypersensitivity during the treatment of childhood acute lymphoblastic leukemia. *Leuk Lymphoma* 56(11), 3103-3108 (2015).

32. Kutszegi N, Yang X, Gezsi A *et al*. HLA-DRB1*07:01-HLA-DQA1*02:01-HLA-DQB1*02:02 haplotype is associated with a high risk of asparaginase hypersensitivity in acute lymphoblastic leukemia. *Haematologica* 102(9), 1578-1586 (2017).

33. Hojfeldt SG, Wolthers BO, Tulstrup M *et al*. Genetic predisposition to PEG-asparaginase hypersensitivity in children treated according to NOPHO ALL2008. *Br J Haematol* 184(3), 405-417 (2019).

34. Kumar S, Ooi CY, Werlin S *et al*. Risk Factors Associated With Pediatric Acute Recurrent and Chronic Pancreatitis: Lessons From INSPPIRE. *JAMA Pediatr* 170(6), 562-569 (2016).

35. Liu C, Yang W, Devidas M *et al*. Clinical and Genetic Risk Factors for Acute Pancreatitis in Patients With Acute Lymphoblastic Leukemia. *J Clin Oncol* 34(18), 2133-2140 (2016).

36. Witt H, Beer S, Rosendahl J *et al*. Variants in CPA1 are strongly associated with early onset chronic pancreatitis. *Nat Genet* 45(10), 1216-1220 (2013).

37. Derikx MH, Kovacs P, Scholz M *et al*. Polymorphisms at PRSS1-PRSS2 and CLDN2-MORC4 loci associate with alcoholic and non-alcoholic chronic pancreatitis in a European replication study. *Gut* 64(9), 1426-1433 (2015).

38. Lee YJ, Kim KM, Choi JH, Lee BH, Kim GH, Yoo HW. High incidence of PRSS1 and SPINK1 mutations in Korean children with acute recurrent and chronic pancreatitis. *J Pediatr Gastroenterol Nutr* 52(4), 478-481 (2011).

39. Sobczynska-Tomaszewska A, Bak D, Oralewska B *et al*. Analysis of CFTR, SPINK1, PRSS1 and AAT mutations in children with acute or chronic pancreatitis. *J Pediatr Gastroenterol Nutr* 43(3), 299-306 (2006).

40. Tanaka Y, Urayama KY, Kawaguchi T *et al*. The Association Between L-Asparaginase Hypersensitivity and Genetic Variants in Japanese Childhood ALL Patients. *Blood* 128(22), 5141-5141 (2016).

41. Abaji R, Gagne V, Xu CJ *et al*. Whole-exome sequencing identified genetic risk factors for asparaginase-related complications in childhood ALL patients. *Oncotarget* 8(27), 43752-43767 (2017).

42. Karol SE, Larsen E, Cheng C *et al*. Genetics of ancestry-specific risk for relapse in acute lymphoblastic leukemia. *Leukemia* 31(6), 1325-1332 (2017).

43. Place AE, Stevenson KE, Vrooman LM *et al*. Intravenous pegylated asparaginase versus intramuscular native Escherichia coli L-asparaginase in newly diagnosed childhood acute lymphoblastic leukaemia (DFCI 05-001): a randomised, open-label phase 3 trial. *Lancet Oncol* 16(16), 1677-1690 (2015).

44. Pui CH, Campana D, Pei D *et al*. Treating childhood acute lymphoblastic leukemia without cranial irradiation. *N Engl J Med* 360(26), 2730-2741 (2009).

45. Conter V, Valsecchi MG, Parasole R *et al*. Childhood high-risk acute lymphoblastic leukemia in first remission: results after chemotherapy or transplant from the AIEOP ALL 2000 study. *Blood* 123(10), 1470-1478 (2014).

46. Stary J, Zimmermann M, Campbell M *et al*. Intensive chemotherapy for childhood acute lymphoblastic leukemia: results of the randomized intercontinental trial ALL IC-BFM 2002. *J Clin Oncol* 32(3), 174-184 (2014).

47. Larsen EC, Devidas M, Chen S *et al*. Dexamethasone and High-Dose Methotrexate Improve Outcome for Children and Young Adults With High-Risk B-Acute Lymphoblastic Leukemia: A Report From Children's Oncology Group Study AALL0232. *J Clin Oncol* 34(20), 2380-2388 (2016).

48. Bowman WP, Larsen EL, Devidas M *et al*. Augmented therapy improves outcome for pediatric high risk acute lymphocytic leukemia: results of Children's Oncology Group trial P9906*. Pediatr Blood Cancer* 57(4), 569-577 (2011).

49. Marshall GM, Dalla Pozza L, Sutton R *et al*. High-risk childhood acute lymphoblastic leukemia in first remission treated with novel intensive chemotherapy and allogeneic transplantation*. Leukemia* 27(7), 1497-1503 (2013).

50. Yang JJ, Cheng C, Devidas M *et al*. Ancestry and pharmacogenomics of relapse in acute lymphoblastic leukemia*. Nat Genet* 43(3), 237-241 (2011).

51. Yang JJ, Cheng C, Devidas M *et al*. Genome-wide association study identifies germline polymorphisms associated with relapse of childhood acute lymphoblastic leukemia*. Blood* 120(20), 4197-4204 (2012).

52. Silverman LB, Gelber RD, Dalton VK *et al*. Improved outcome for children with acute lymphoblastic leukemia: results of Dana-Farber Consortium Protocol 91-01*. Blood* 97(5), 1211-1218 (2001).

53. Pession A, Valsecchi MG, Masera G *et al*. Long-term results of a randomized trial on extended use of high dose L-asparaginase for standard risk childhood acute lymphoblastic leukemia*. J Clin Oncol* 23(28), 7161-7167 (2005).

54. Pastorczak A, Fendler W, Zalewska-Szewczyk B *et al*. Asparagine synthetase (ASNS) gene polymorphism is associated with the outcome of childhood acute lymphoblastic leukemia by affecting early response to treatment*. Leuk Res* 38(2), 180-183 (2014).

55. Schotte D, De Menezes RX, Akbari Moqadam F *et al*. MicroRNA characterize genetic diversity and drug resistance in pediatric acute lymphoblastic leukemia*. Haematologica* 96(5), 703-711 (2011).

56. Mei Y, Gao C, Wang K *et al*. Effect of microRNA-210 on prognosis and response to chemotherapeutic drugs in pediatric acute lymphoblastic leukemia*. Cancer Sci* 105(4), 463-472 (2014).

57. Mesrian Tanha H, Mojtabavi Naeini M, Rahgozar S, Moafi A, Honardoost MA. Integrative computational in-depth analysis of dysregulated miRNA-mRNA interactions in drug-resistant pediatric acute lymphoblastic leukemia cells: an attempt to obtain new potential gene-miRNA pathways involved in response to treatment*. Tumour Biol* 37(6), 7861-7872 (2016).

58. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs*. Genome Res* 19(1), 92-105 (2009).

59. Krejci O, Starkova J, Otova B *et al*. Upregulation of asparagine synthetase fails to avert cell cycle arrest induced by L-asparaginase in TEL/AML1-positive leukaemic cells. *Leukemia* 18(3), 434-441 (2004).

60. Escherich G, Troger A, Gobel U *et al*. The long-term impact of in vitro drug sensitivity on risk stratification and treatment outcome in acute lymphoblastic leukemia of childhood (CoALL 06-97). *Haematologica* 96(6), 854-862 (2011).

61. Wellmann R, Borden BA, Danahey K *et al*. Analyzing the clinical actionability of germline pharmacogenomic findings in oncology. *Cancer* 124(14), 3052-3065 (2018).

62. Cui H, Darmanin S, Natsuisaka M *et al*. Enhanced expression of asparagine synthetase under glucose-deprived conditions protects pancreatic cancer cells from apoptosis induced by glucose deprivation and cisplatin. *Cancer Res* 67(7), 3345-3355 (2007).

63. Lorenzi PL, Reinhold WC, Rudelius M *et al*. Asparagine synthetase as a causal, predictive biomarker for L-asparaginase activity in ovarian cancer cells. *Mol Cancer Ther* 5(11), 2613-2623 (2006).

64. Samudio I, Konopleva M. Asparaginase unveils glutamine-addicted AML. *Blood* 122(20), 3398-3400 (2013).

65. Tardito S, Uggeri J, Bozzetti C *et al*. The inhibition of glutamine synthetase sensitizes human sarcoma cells to L-asparaginase. *Cancer Chemother Pharmacol* 60(5), 751-758 (2007).

3.6. Tables & Figures

Table 1. This table summarizes the prominent studies in the literature which investigated the pharmacogenetics of asparaginase and highlights the most important finding.

Study	Method	Gene	Polymorphism	Toxicity	Discovery Cohort (N)	Internal Validation Cohort (N)	Validated	Notes & Conclusions
Chen et al. 2010	GWAS	*GRIA1*	rs4958351 rs10070447 rs6890057 rs4958676 rs6889909	HSRs	322	163	Yes	• rs4958351 had the strongest association • Carriers of the minor alleles were at increased risk of developing HSRs to ASNase.
Rajic et al. 2015	Gene-Candidate		rs4958351 rs10070447 rs6890057 rs4958676 rs6889909	HSRs	146	No	N/A	• The associations of the variants with increased risk of HSRs found in the original study by (Chen et al. 2010) were successfully replicated by (Rajic et al. 2015, N=146)
Kutszegi et al. 2015	Gene-Candidate		rs4958351 rs10070447 rs6890057 rs4958676 rs6889909	HSRs	505	No	N/A	• The associations of the variants with increased risk of HSRs found in the original study by (Chen et al. 2010) were not replicated by (Kutszegi et al. 2015, N=505).
Fernandez et al. 2014	Gene-Candidate	*HLA-DRB1*	*HLA-DRB1* *07:01	HSRs	541	1329	Yes	• The variant allele was associated with an increased risk of ASNase HSR. • Alleles that confer high-affinity binding to ASNase epitopes contribute to the observed higher frequency of HSRs.
Fernandez et al. 2015	GWAS	*HLA-DRB1*	rs17885382	HSRs	3308	No	N/A	• The variant allele was associated with an increased risk of ASNase induced HSR and is in almost in perfect linkage disequilibrium with HLA-DRB1*07:01 found in a previous study.

		NFATC2	rs6021191					• The variant is associated with an increased risk of ASNase HSR.
Kutszegi et al. 2016	Gene-Candidate	*HLA-D Region*	HLA-DRB1 *07:01 HLA-DQB1 *02:02	HSRs	359	No	N/A	• Patients with *HLA-DRB1*07:01* allele or *HLA-DQB1*02:02* allele had significantly higher risk of developing HSRs. • Carrying the *HLA-DRB1*07:01–HLA-DQA1*02:01–HLA-DQB1*02:02* haplotype was associated with the highest risk.
Højfeldt et al. 2019	GWAS	*CNOT3*	*rs73062673*	HSRs	831	No	N/A	• The minor allele of rs73062673 was associated with an increased risk of HSRs. • The study also reported two other positive associations involving rs9272131 in *HLA-DQA1* gene and rs115360810 in *TAP2* gene, albeit not significant on a genome-wide level.
Ben Tanfous et al. 2015	Gene-Candidate	*ASNS*	*rs3832526*	Allergy	285	248	Yes	• Patients homozygous for the triple repeat allele (*3R*) had the complications more frequently than other genotype groups.
				Pancreatitis			No	• ASNS haplotype **1* harbouring double repeat (*2R*) allele conferred a protective effect from these toxicities and the association with reduced risk of allergies was further validated in the replication cohort.
Abaji et al. 2017	EWAS	*MYBBP1A*	rs3809849	Pancreatitis	302	282	Yes	• This variant was also associated with allergy, thrombosis, event-free survival and overall survival.
		IL16	rs11556218				Yes	• This variant was also associated with thrombosis.
		SPEF2	rs34708521				Yes	• This variant was also associated with thrombosis.
Liu et al. 2016	GWAS	*CPA2*	Gene-Level	Pancreatitis	5185	213	Yes	• 16 SNPs in this gene were significantly associated with pancreatitis in a gene-level analysis. rs199695765 showed the strongest association.

Wolthers et al. 2017	GWAS	*ULK2*	rs281366	Pancreatitis	700	No	N/A	• 14 of the 27 associations found in the study were polymorphisms in *ULK2* gene • The variant in *RGS6* gene was associated with pancreatitis in patients less than 10 years old • The risk of pancreatitis associated with carrying the risk alleles of rs281366 and rs17179470 was additive in patients less than 10 years old.
		RGS6	rs17179470					
Wolthers et al. 2018	GWAS	*N/A*	rs62228256	Pancreatitis	1564	318	No	• rs62228256 had the strongest association signal. It is located located in a noncoding region of the genome upstream from the NFATC2 gene and acts as an eQTL for it in pancreatic tissue. • Minor alleles of all SNPs were associated with an increased risk of pancreatitis.
		PRSS1-PRSS2	rs13228878 rs10273639			2729	Yes	
Alachkar et al. 2017	Gene-Candidate	*SOD2*	rs4880	Hepato-toxicity	190	No	N/A	• Increased risk of hepatotoxicity following ASNase-based treatment for carriers of the minor allele.
Rousseau et al. 2011	Gene-Candidate	*ATF5*	rs11554772	EFS	318	267	Yes	• Carriers of the minor allele who received E.coli ASNase were at higher risk of ALL relapse and the result was corroborated through higher promoter activity.
		ASNS	*rs3832526*					• Homozygous carriers of the double repeat (2R) had significantly lower EFS.
Pastorczak et al. 2014	Gene-Candidate	*ASNS*	*rs3832526*	EFS / Response	264	No	N/A	• Carriers of the (3R) allele with a poor response at day 15 had an increased risk of events.

ALL: acute lymphoblastic leukemia; ASNase: asparaginase; ASNS: asparagine synthetase; EFS: event-free survival; EWAS: exome-wide association study; GWAS: genome-wide association study; HSRs: hypersensitivity reactions; N/A: not applicable; SNPs: single-nucleotide polymorphisms.

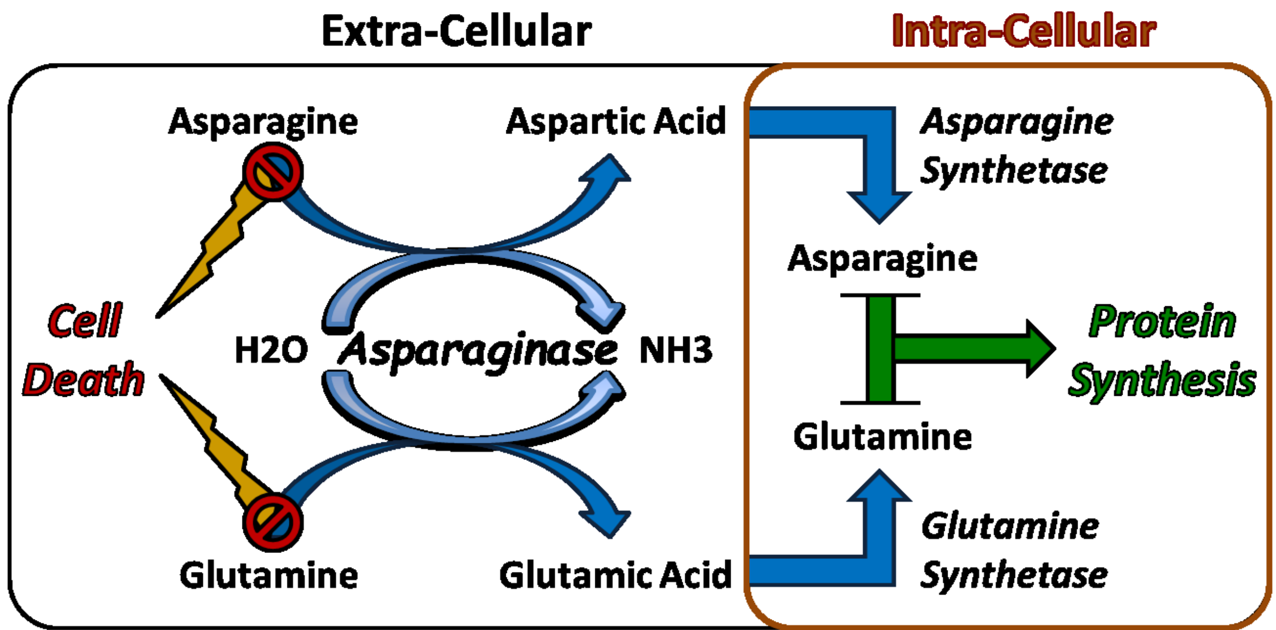


Figure 1. Mechanism of action of Asparginase.

Illustration of the mechanism of action of asparaginase as an anti-leukemic agent. The activity of asparaginase leading to the depletion of extra-cellular asparagine and/or glutamine and eventual cell death is counteracted by the intra-cellular production of these amino acids through asparagine synthetase and glutamine synthetase, respectively.

Section B

Chapter 4

Whole-exome sequencing identified genetic risk factors for asparaginase-related complications in childhood ALL patients

This chapter presents the results of an original work that was the fruit of collaboration between several research groups orchestrated by Dr. Maja Krajinovic. It was performed as a type of post-hoc analysis targeting data available from institutional acute lymphoblastic leukemia trial protocols and was aimed at finding genetic associations that could explain the observed differential sensitivity of patients to complications of asparaginase.

My involvement in this work spanned the entire project from the production of genotype libraries of the variants of interest that were identified following the initial analysis, to performing the association studies with clinical response parameters in the discovery group as well as testing their reproducibility in the validation cohort. I also analysed the extent of individual contributions of validated variants to the overall combined-effect in modulating the response and constructed the risk prediction model for pancreatitis. I drafted the manuscript under the supervision of Dr. Krajinovic, which was then revised by all authors. My contribution to this work can be estimated as 70% of the total input.

This work was published in the journal *Oncotarget*.

Whole-exome sequencing identified genetic risk factors for asparaginase-related complications in childhood ALL patients

Running title: WES-identified biomarkers for ASNase complications

Rachid Abaji,[1,7] Vincent Gagné,[1] Chang Jiang Xu[1], Jean-François Spinella,[1] Francesco Ceppi[1] Caroline Laverdière,[1,2] Jean-Marie Leclerc[1,2] Stephen E. Sallan,[3,4] Donna Neuberg,[5] Jeffery L. Kutok,[6] Lewis B. Silverman,[3,4] Daniel Sinnett,[1,2] and Maja Krajinovic✉[1,2,7]

[1]Research Center, Centre Hospitalier Universitaire Sainte-Justine, Montreal, QC;

[2]Department of Pediatrics, University of Montreal, Montreal, QC;

[3]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA;

[4]Division of Hematology/Oncology, Children's Hospital, Boston, MA;

[5]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA;

[6]Department of Pathology, Brigham and Women's Hospital, Boston, MA;

[7]Department of Pharmacology, University of Montreal, Montreal, QC;

✉Corresponding author.

Key words: acute lymphoblastic leukemia; asparaginase; exome-wide association; pharmacogenetics; whole-exome sequencing.

Financial support: Dr. Krajinovic received grants for this project from Leukemia Lymphoma Society of Canada and Canadian Institutes of Health Research (#106458). Dana-Farber Cancer Institute ALL treatment protocols are supported by the National Cancer Institute/NIH grant 5 P01CA06848.

Disclosure of Conflicts of Interest: The authors declare no competing financial interests.

Corresponding author:

Name: Maja Krajinovic

Address: Centre de Recherche, CHU Sainte-Justine, 3175 chemin de la Côte-Ste-Catherine, Montréal, QC, H3T 1C5 Canada

E-mail address: maja.krajinovic@umontreal.ca

Phone Number: 514 345-4931 #6259

Word count for abstract: **264**

Word count for text: **4,398**

Table count: **3**

Figure count: **6**

Reference count: **51**

4.1. Abstract

Allergy, pancreatitis and thrombosis are common side-effects of childhood acute lymphoblastic leukemia (ALL) treatment that are associated with the use of asparaginase (ASNase), a key component in most ALL treatment protocols. Starting with predicted functional germline variants obtained through whole-exome sequencing (WES) data of the Quebec childhood ALL cohort (N=302) we performed exome-wide association studies with ASNase-related toxicities. A subset of top-ranking variants was further confirmed by genotyping followed by validation in an independent replication group (N=282); except for thrombosis which was not available for that dataset. SNPs in 12 genes were associated with ASNase complications in discovery cohort including 3 that were associated with allergy, 3 with pancreatitis and 6 with thrombosis. The risk was further increased through combined SNPs effect ($p \leq 0.002$), suggesting synergistic interactions between the SNPs identified in each of the studied toxicities. Interestingly, rs3809849 in the *MYBBP1A* gene was associated with allergy (p= 0.0006), pancreatitis (p=0.002), thrombosis (p=0.02), event-free survival (p=0.02) and overall survival (p=0.003). Furthermore, rs11556218 in *IL16* and rs34708521 in *SPEF2* were both associated with thrombosis (p=0.01 and p=0.03, respectively) and pancreatitis (p=0.02). The association of SNPs in *MYBBP1A, SPEF2* and *IL16* genes with pancreatitis was replicated in the validation cohort ($p \leq 0.05$) as well as in combined cohort (p=0.0003, p=0.008 and p=0.02, respectively). The synergistic effect of combining risk loci had the highest power to predict the development of pancreatitis in both cohorts and was further potentiated in the combined cohort ($p=1x10^{-8}$).The present work demonstrates that using WES data is a successful "hypothesis-free" strategy for identifying significant genetic markers modulating the effect of the treatment in childhood ALL.

4.2. Introduction

Acute lymphoblastic leukemia (ALL) is the most common cancer in children and it accounts for 25% of all childhood malignancies.[1-3] Survival rates have improved significantly over time with the progressive intensification of ALL treatment and the implementation of multi-agent risk–adapted protocols. [2-4] However, a subset of patients experience treatment failure or short-term treatment-related toxicities which might result in the interruption or discontinuation of chemotherapy or can have severe, fatal, or lifelong consequences that challenge their ability to lead a normal life as future adults.[2]

Asparaginase (ASNase) was introduced as major component of ALL treatment protocols in 1970 and has been a mainstay of therapy ever since.[1-3, 5] It is an enzyme that catalyzes the hydrolysis of the amino acid asparagine (ASN) into aspartic acid and ammonia and is thus required by all cells. Cancerous lymphoblasts usually depend on extracellular sources of asparagine to support their fast growth as they have ASNS levels that are relatively lower than their needs. Thus, depletion of asparagine by ASNase reduces the capacity of protein biosynthesis in leukemia cells which selectively promotes their death.[1, 2]

Less favorable outcome in childhood ALL treatment has been associated with treatment discontinuation and the failure to receive the full course of ASNase due to treatment-related toxicities.[2, 4, 6] L-asparaginase comes from 2 bacterial sources, *Escherichia coli (E.coli)* and *Erwinia chrysanthemi.* While *E. coli*–derived enzyme generally has higher efficacy, it has been reported to have higher toxicity.[1-3] ASNase-related

treatment toxicities mostly include allergic reactions, pancreatitis and thrombotic events frequently associated with discontinuation of asparaginase treatment.[1-4]

Given the bacterial origin of asparaginase, it is not surprising that it is capable of inducing immune reactions in vivo as up to 30% of patients experience a hypersensitivity reaction to *E. coli*-derived asparaginase.[1-4, 7] While reported rates vary across literature, clinical and subclinical hypersensitivity reactions are associated with decreased asparaginase activity levels caused by neutralizing antibodies and may be influenced by the asparaginase preparation used, dose intensity, and other medications.[3, 4, 7]

Around 2–18% of patients receiving asparaginase develop pancreatitis which is usually associated with clinical symptoms along with serum amylase and/or lipase elevation reaching more than three times upper-normal limits.[3, 4] While currently known risk factors include intensive treatment and older age, the pathogenesis of asparaginase-induced pancreatitis is not yet fully understood and is thought to occur as a result of an underlying predisposition.[2, 8] Interestingly, unlike with hypersensitivity reactions the incidence of pancreatitis does not seem to be influenced, at least in some studies, by the formulation of asparaginase used.[3, 4, 8]

Thrombosis, defined as venous and/or arterial thromboembolism, has a higher incidence in paediatric oncology patients and is reported with both *E. coli*– and *Erwinia*-derived asparaginase (mainly due to interference with the hepatic synthesis of coagulation proteins) and has an overall incidence of around 5% according to recent studies.[4, 5] Many factors have been associated with the risk of thrombosis, some related to the disease, others

to the treatment (like the dose and duration of asparaginase exposure) as well as to patient specific factors such as older age, female gender, non-O blood group, obesity, inherited prothrombotic states or central venous catheter.[3, 5, 9, 10]

Being able to predict which patients will experience asparaginase-related toxicity and switching them to an alternate asparaginase formulation[4] or a different treatment protocol that does not depend heavily on asparaginase has been shown to yield superior outcomes.[8] Accordingly, using genetic markers for prospective stratification of patients at high risk of developing allergic reactions, pancreatitis or thrombosis has the potential to improve ALL treatment by identifying a patient subgroup which might benefit more from an alternative regimen.[4, 8]

Over the past decade, important advances in sequencing technology have been achieved which not only helped deciphering leukemia specific mutations,[11, 12] but also provided comprehensive information on germline polymorphisms for association studies of complex disease traits and suboptimal treatment responses.[11, 12] Here we present the results of an exome-wide association study (EWAS) that was performed on whole exome sequencing (WES) data obtained from childhood patients who received asparaginase as part of ALL treatment protocol. The results provide an insight on novel pharmacogenetic markers associated with asparaginase related allergic reactions, pancreatitis and thrombosis.

4.3. Results

4.3.1. Asparaginase-related complications

Twenty-nine patients (9.6%) received a formulation containing *Erwinia* derived asparaginase while the rest received an *E.coli* derived formulation (Table 1). The observed frequency of the asparaginase-related toxicities was comparable to those reported in the literature [2, 4, 5, 8]: 15.9% (48) patients developed allergies (with 40 of them having serious systemic reactions while the rest having mixed or local reactions); 5% (15) experienced pancreatitis (12 severe and 3 mild to moderate); and 3.3% (10) had thrombosis. Consequently, and following the treatment protocols guidelines, all patients with complications needed treatment modification, either interruption or switch to other types of asparaginase.

Toxicities in replication cohort had similar frequencies to those of the discovery cohort as there were 20.9% (59) patients with allergies (39 systemic) and 7.4% (21) with pancreatitis (14 severe). Information on thrombosis was not available. The frequency of *Erwinia*-derived asparaginase and *E.coli* formulation was also comparable to the discovery cohort.

4.3.2. Association Study

The number of predicted functional common variants recovered from WES data was 5527; from these, 4519 SNPS distributed across 3802 genes, respected Hardy-Weinberg equilibrium and were tested for an association with asparaginase-related toxicities. Out of the 115 top-ranking SNPs identified from WES data with FDR<20%, 43 were associated with allergy, 40 with pancreatitis and 32 with thrombosis (Supplemental Table S1). Given the relatively large number of hits, selective exclusion was performed to remove the SNPs found in genes that are unlikely to be involved in the pathways of studied toxicities (e.g. genes of the olfactory receptors family and other neurosensory functions as well as the ones whose expression in restricted to tissues that are irrelevant to the toxicity in question). Accordingly, and out of the remaining pool, thirty two SNPs (8 SNPs associated with allergy, 10 with thrombosis and 14 with pancreatitis) with MAF higher than 5% in discovery cohort and located in genes whose biological function could be relevant for drug responses, were selected (Figure 1 and Supplemental Table S2).

Based on genotyping results, 3 variants were associated with allergy (Table 2) Carriers of the minor allele of rs9656982 in the *SLC7A13* gene and of rs3809849 in the *MYBBP1A* gene were associated in additive manner (OR = 2.1; 95% CI, 1.1–3.9; p= 0.02 and OR = 2.4; 95% CI, 1.4–3.9; p= 0.0006, respectively), whereas the effect of rs75714066 minor allele in the *YTHDC2* gene followed the dominant model (OR = 3.1; 95% CI, 1.4–7.0; p= 0.008).

Three SNPs were significantly associated with a risk of pancreatitis (Table 2). Carriers of the minor allele of rs72755233 in the *ADAMTS17* gene and of rs3809849 in the *MYBBP1A* gene were at higher risk of pancreatitis when compared to non-carriers (OR = 5.6; 95% CI, 1.9–16.3; p= 0.002 and OR =6.9; 95% CI, 1.9–25.2; p= 0.002, respectively), whereas the SNP (rs9908032) in the *SPECC1* gene followed the additive model (OR = 3.9; 95% CI, 1.6–9.2; p= 0.0008).

Six SNPs were associated with thrombosis (Table 2). Carriers of minor alleles were predisposed to a higher risk when compared to non-carriers including rs6584356 in *PKD2L1* (OR =5.0; 95% CI, 1.2–20.7; P= 0.05); rs3742717 in *RIN3* (OR =13.8; 95% CI, 2.3–82.5; P= 0.02); rs34708521 in *SPEF2* (OR =6.1; 95% CI, 1.4–26.9; P= 0.03); rs7926933 in *MPEG1* (OR =5.7; 95% CI, 1.5–22.1; P= 0.01); rs11556218 in *IL16* (OR =7.4; 95% CI, 1.8–31.2; P= 0.01) and rs62619938 in *SLC39A12* (OR =4.4; 95% CI, 1.6–11.7; P= 0.0005).

In the light of their positive association, each SNP was tested for possible associations with the two other side-effects. Interestingly, on the top of their association with allergy and pancreatitis, homozygote carriers of the variant rs3809849 allele in the *MYBBP1A* gene were associated with a higher risk of thrombosis (OR= 6.8; 95% CI, 1.3–36.5; p= 0.02; Figure 2a); whereas, rs11556218 in *IL16* and rs34708521 in *SPEF2* were, in addition to thrombosis, also correlated with pancreatitis (OR =3.1; 95% CI, 1.1–8.6; p= 0.02 and OR =3.4; 95% CI, 1.1–10.6; p= 0.02; Figures 2b and 2c, respectively).

The risk of any toxicity increased in additive manner with the minor C allele of the rs3809849 SNP in the *MYBBP1A* gene (OR= 2.7; 95% CI, 1.7-4.3; p= $3x10^{-5}$; Figure 3a). The same SNP was significantly associated with less favorable disease outcomes as homozygous C allele carriers had a reduced EFS (OR =3.2; 95% CI, 1.4–7.4; p= 0.02; Figure 3b) and OS (OR =5.3; 95% CI, 1.8–15.8; p= 0.003; Figure 3b).

In the multivariate analysis, only the association of rs34708521 in *SPEF2* gene with thrombosis lost significance (OR=4.3; 95% CI, 0.8–22.3; p=0.08), whereas other associations remained significant in their respective models (Supplemental Table S3).

4.3.3. Replication Analysis

Out of the 6 significant associations with allergy and pancreatitis that were confirmed by genotyping in the discovery cohort, the association between rs3809849 in the *MYBBP1A* gene and pancreatitis was replicated in the DFCI cohort (OR =2.8; 95% CI, 1.1–7.1; p= 0.05, Figure 5a). Interestingly, the positive associations that were observed between rs11556218 in *IL16* and rs34708521 in *SPEF2* and the higher risk of pancreatitis were also seen in DFCI cohort (OR =6.7; 95% CI, 1.1–41.5; p= 0.05 in patients with mild and moderate pancreatitis and OR =3.4; 95% CI, 1.1–10.5; p= 0.02, Figures 5b and 5c, respectively). More significant associations were noted for rs3809849 and rs34708521 when analyses were performed in the cohort combining discovery and replication set (p= 0.0003 and p= 0.008, respectively, Supplemental Table S4). The significant associations with allergies were not replicated, whereas those with thrombosis were not tested since the data were not available in the validation group.

4.3.4. Combined Effect Model

We next investigated the combined effect of the top-ranked SNPs in each of the toxicities. In this model, a significant correlation was observed between the number of variant alleles carried and the increase in the risk of each of the toxicities. For allergy, the risk associated with an additive effect was 2.5 (95% CI, 1.6–3.9; p = $4X10^{-5}$, Figure 4a), whereas the presence of 2 or more variant alleles was associated with a 6.5-fold increase in the risk of experiencing allergic reactions as compared to not carrying any variant allele (OR = 6.5; 95% CI, 2.7-15.6; p = $1X10^{-5}$, Figure 4a). Similar effect was noted for thrombosis (OR for additive effect =4.0; 95% CI, 1.5–10.6; p= 0.002, Figure 4b). As for pancreatitis, the addition of all 3 variants in the model increased the risk 6-fold (OR =5.9; 95% CI, 2.4–14.4; p= $7x10^{-6}$, Figure 4c) with carriers of at least two variant alleles being almost 28 times more at risk as compared to those without any variant allele (OR = 27,9; 95% CI, 3,5-224,3; p = $3X10^{-5}$, Figure 4c).

In an attempt to increase the discrimination ability of the model, rs11556218 in *IL16* and rs34708521 in *SPEF2* that were initially investigated for their association with thrombosis but later found to be also associated with pancreatitis, were added to the analysis. In this new comprehensive model with five variants, the groups of 0, 1, 2 and 3 or more variant alleles were compared. The association between the number of minor alleles and the increase in the risk of pancreatitis was directly proportional (OR =5; 95% CI, 2.4–10.2; P= $5x10^{-7}$, Supplemental figure S1).

The model combining the 3 SNPs associated with pancreatitis (i.e. rs72755233 in *ADAMTS17*, rs3809849 in *MYBBP1A* and rs9908032 in *SPECC1*) was also replicated in the validation cohort (OR =2.2; 95% CI, 1.1–4.6; p= 0.02, Figure 5d), as also was the comprehensive model with the five variants (OR =2.6; 95% CI, 1.3–5.4; p= 0.005, Supplemental Figure S1). The association was further potentiated in the combined cohort ($p=2x10^{-6}$ and $p=1x10^{-8}$ for the models containing 3 and 5 SNPs, respectively; Supplemental figure S2).

4.3.5. Risk Prediction

To assess the performance of the comprehensive combined-effect model in predicting the risk of ASNase-induced pancreatitis, we used the weighted genetic risk score (wGRS) method.[13] A risk score was assigned to each patient by taking the sum of the weighted score of each risk allele across the 5 loci. We then applied these values derived from the discovery cohort to assign the risk scores to patients in the validation cohort. The performance of the model in the discovery, replication and combined cohorts, is summarized in Table 3. The discriminatory ability of the model is reflected by the area under the ROC curve derived from the wGRS. The best sensitivity/specificity values were derived from the OR values greater than 11 corresponding to at least two associated SNPs. The model was successfully validated in the replication and combined cohorts.

In order to evaluate the efficiency and reproducibility of the model in assigning patients to risk categories, the patients were divided into 4 groups based on the weighted genetic risk scores. Patients who had a score of 0 (indicating the absence of any risk allele) were considered the standard risk category, whereas those who had higher scores were divided into 3 equal groups corresponding to low, intermediate and high risk based on their individually assigned cumulative OR. Distribution of the patients with pancreatitis was compared across the groups and between the two cohorts. The distribution of patients with pancreatitis in the replication cohort (which was based on the predicted ORs) was similar to that of patients from the discovery cohort (who were classified according to their observed ORs), Figure 6. Patients predicted to have the highest risk of pancreatitis (thus assigned to group H) had substantially higher frequency of patients who actually developed pancreatitis and the observed OR of this group was significantly greater than that of the standard risk group (Figure 6).

4.4. Discussion

Using WES data we identified common genetic variants significantly associated with asparaginase-related side-effects. The rs3809849 in the *MYBBP1A* gene was associated both with allergy and pancreatitis; the significant association with pancreatitis was replicated in the validation cohort. The same SNP was also associated with thrombosis as well as reduction in EFS and OS in discovery cohort. The observed association with EFS and OS could be the result of treatment interruption due to the development of side-effects or could be mediated by ASNase deactivation in the case of allergic reactions. In either situation, the patients would consequently receive a lower ASNase dose intensity, which has been previously shown to be associated with less favourable outcome.[2, 4, 6] Another possible hypothesis involves an increased clearance of dexamethasone driven by anti-asparaginase antibodies which ultimately reduces the overall exposure to this drug and is associated with higher risk of relapse.[14] The effect of other confounding factors such as, for example, leukemia specific mutations, cannot be however ruled out.

MYBBP1A gene encodes MYB Binding Protein 1a which is important for early embryonic development as well as many other cellular processes including mitosis, cell cycle control, response to nuclear stress, synthesis of ribosomal DNA and tumoral suppression via modulation of the p53 activity.[15, 16] MYBBP1a also acts as a co-repressor of the nuclear factor kappaB (NF-kB),[17, 18] a transcription factor activated in response to inflammatory and stress signals, apoptosis and cellular proliferation. Interestingly, a key role of NF-kB in the development of acute pancreatitis has been recently documented.[19]

To our knowledge, this is the first study demonstrating an association between *MYBBP1A* gene and the risk of pancreatitis. In general, rs3809849 in *MYBBP1A* gene was rarely investigated. There is only one study which found significant association of this SNP with higher risk of tuberculosis.[18]

Another interesting observation is that 2 loci that were initially investigated for their possible association with thrombosis also showed significant and reproducible associations with pancreatitis. Accordingly, G allele carriers of the rs11556218 SNP in the *IL16* gene and carriers of the A allele in the rs34708521 SNP of the *SPEF2* gene, were at higher risk of pancreatitis in both discovery and replication cohorts. The association with *IL16* is of particular interest because *IL16* gene codes for interleukin-16, a multifactorial cytokine involved in inflammatory and autoimmune diseases as well as cancer risk.[20] In the past few years, rs11556218 has been found to be associated with a wide range of conditions such as endometriosis,[21] Alzheimer's Disease,[22] emphysema,[23] coronary artery disease,[24] ischemic stroke,[25] systemic lupus erythematous,[26] chronic hepatitis B infection,[27] osteoarthritis,[20] overall cancer risk as well as particular cancer types.[28] *SPEF2* stands for "Sperm Flagellar 2" gene which encodes for a protein that is required for correct axoneme development.[29] Even though the association of this gene with thrombosis and pancreatitis might seem counterintuitive, we are tempted to speculate that this might be mediated by the role this gene has in protein dimerization activity and the fact that the protein it encodes is significantly overexpressed in platelets.[30] This finding should be investigated in future studies.

Our analysis also suggests that synergistic interactions might exist between the SNPs identified in each of the studied toxicities, which could explain the markedly significant associations and high odd-ratios in the combined SNPs models. Same combined effect was noted for pancreatitis in the replication set. When all associated SNPs were regarded together, either in combined or comprehensive model, they could explain almost all cases of pancreatitis in both patients' groups. This was further supported by the model based on wGRS that displayed the best discrimination ability between individuals with and without pancreatitis and confidence limits were substantially above random predictions. Importantly, similar sensitivity and specificity values were observed in the discovery and replication cohorts at odds ratio greater than the chosen threshold which reflects the stability of the model. Furthermore, the prediction model using wGRS values derived from the discovery cohort to assign patients of the validation cohort into risk groups was able to detect far more patients at risk of pancreatitis than any of the SNPs considered alone. In fact, the group of patients predicted to have the highest risk based on their calculated wGRS had a substantial overrepresentation of individuals with pancreatitis compared to all other groups and a significantly higher OR compared to the standard risk group.

This indicates that it would be important to further investigate the utility of using sets of SNPs, rather than individual variants. This EWAS added novel genetic markers to the existing pool of pharmacogenetics modifiers of ASNase treatment that were previously described by several groups including ours, using GWAS and candidate-gene studies (ex. *ATF5* and EFS,[31] ASNS and allergy/pancreatitis,[2] *GRIA1* and hypersensitivity,[32] *HLA-DRB1*0701* and allergy,[33] *CPA2* and pancreatitis[8]). Collectively, this rapidly growing

pool of markers might become more efficient in explaining the observed inter-individual variability in morbidities associated with anti-leukemia treatment which can eventually help developing genotype guided interventions for patients predisposed to such toxicities.[34]

As per the impact of the sources of ASNase used, the results did not differ significantly when samples of patients who received *Erwinia*-derived ASNase were excluded from the analysis. The only noteworthy observation was related to the association of *IL16* with pancreatitis. On the top of the association with mild-moderate pancreatitis shown earlier in replication cohort (when both ASNase formulations were confounded), *IL16* SNP also showed a significant association with overall pancreatitis in the group treated only with E*. coli* derived formulation in the replication cohort. This difference can be due to the fact that patients treated with *E. coli* ASNase usually have higher rates of ASNase related toxicities.[1, 2] Likewise, the addition of other factors (age, sex, protocol, risk groups) in multivariate model did not affect the results since all of the presented associations remained significant in the multivariate analysis, with the sole exception of rs34708521 in *SPEF2* gene with thrombosis.

There are several limitations to our study. The analyses were done retrospectively as clinical data were inferred from the patients' medical charts. The distribution of treatment protocols and risk groups varied significantly between the cohorts, which could have introduced variability as patients might have received different ASNase doses. The sample size of the discovery cohort was relatively small and the selected FDR threshold of <20% was relaxed, which might have increased the number of false-positives, possibly reflected in

the high number of EWAS hits. However, the fact that several associations were successfully reproducible in the independent validation cohort supports the validity of the findings. Furthermore, the analysis in the context of a larger sample size provided by the combined cohort further supports the correlation between the SNPs in *MYBBP1A*, *IL16* and *SPEF2* with pancreatitis as the associations gained more significance in the pooled sample. Finally, this study aimed primarily to identify genetic markers that put the patients at risk of developing treatment-related toxicities commonly associated with the use of asparaginase; however, the treatment included other chemotherapeutic agents which makes it difficult to estimate the magnitude of the interaction between asparaginase alone and the genetic composition, requiring experiments in cell lines and animal models to further support the observations.

In conclusion, using WES data in the context of association study was a successful "hypothesis-free" strategy which allowed identifying significant genetic associations with asparaginase-related toxicities in children treated for ALL. Results for pancreatitis were replicated in the independent validation cohort. Even though interesting associations with thrombosis were observed, no replication studies were done due to logistic limitations. Thus, it would be valuable replicating further those results.

4.5. Patients and methods

4.5.1. Study population and endpoints in the analysis

Discovery cohort consisted of 302 children of European descent from the well-established Quebec Childhood ALL (QcALL) cohort who were diagnosed with childhood ALL at the Sainte-Justine University Hospital Centre (SJUHC), Montreal, QC, Canada, between January 1989 and July 2005. All patients received ASNase as part of the Dana-Farber Cancer Institute ALL Consortium protocols DFCI 87-01, 91-01, 95-01, or 00-01 (Table1).[2, 6, 31, 35] In 95-01 and 00-01, one dose of asparaginase was administered during remission induction, and in all protocols it was administered for 20–30 consecutive weeks during consolidation phase. Details about asparaginase doses and formulation are provided elsewhere.[31, 35] Retrospective review of the medical files was conducted to obtain information on ASNase-related toxicity. Hypersensitivity reactions were defined as adverse local or general manifestations from exposure to asparaginase (flushing, erythema, rash, urticaria, drug fever, dyspnoea, symptomatic bronchospasm, oedema or angio-oedema).[2] Pancreatitis was identified according to the diagnostic criteria of the institution and the guidelines of respective protocols which involved pancreatic enzyme elevation of higher than 3-fold the normal levels along with other clinical signs and symptoms that confirm the diagnosis.[2, 36] Thrombosis was determined by clinical symptoms and confirmed by radiologic imaging based on institutional guidelines.[2, 37]

The replication cohort consisted of 282 children who share similar characteristics with the discovery cohort and who were treated according to the 95-01 and 00-01 protocols. All participants had been previously recruited at one of the nine remaining Dana Farber consortium institutions (i.e. DFCI cohort excluding the SJUHC patients). Information on ASNase related allergy and pancreatitis were available for these patients. Clinical characteristics of both the discovery and replication cohorts are shown in Table1.

Written informed consent was obtained in accordance with the Declaration of Helsinki from all participants and/or their parents or legal guardians. Institution ethics committees approved the study.

4.5.2. Whole Exome Sequencing (WES)

DNA was extracted from peripheral blood or bone marrow samples obtained after remission from 244 childhood ALL patients (QcALL cohort)[38] using standard protocols as described previously.[39] Whole exomes were captured in solution with Agilent's SureSelect Human All Exon 50Mb kits, and sequenced on the Life Technologies SOLiD System (patients mean coverage ~35X). Reads were aligned to the hg19 reference genome using SOLiD LifeScope software. PCR duplicates were removed using Picard.[40] Base quality score recalibration was performed using the Genome Analysis ToolKit (GATK)[41] and QC Failure reads were removed. Cleaned BAM files were used to create pileup files using SAMtool.[42]

Germline variants have been called using SNooPer[43] a variant caller based on a machine learning algorithm that uses a subset of variant positions from the sequencing output for which the class is known, either actual variation or sequencing error, to train a data-specific model.

The annotation of the identified germline variants was performed using ANNOVAR.[44] Only missense, nonsense and variations in splicing sites were conserved. The predicted effect of missense variants on the protein function was assessed *in silico* using Sift (≤0.05) [45] and Polyphen2 (≥0.5).[46] Minor allele frequencies (MAF) higher than 5% were derived from the 1000 Genomes (European population) [47] and the NHLBI GO Exome Sequencing Project (European population, ESP).[48]

Fisher's Exact test (allelic association) and Cochran-Armitage trend test, implemented in PLINK[49], were used for an association study. Adjustment for multiple testing was performed by bootstrap false discovery rate (FDR)[50] method; the SNPs retained for further analysis had FDR lower than 20%.

4.5.3. Validation of top-ranking EWAS signals by Genotyping

Genotyping of top ranking EWAS signals was either performed at the McGill University and Génome Québec Innovation Centre through Sequenom genotyping platform or by allele-specific oligonucleotides (ASOs) hybridization as described earlier.[51] Comparison between genotypes and ASNase related complication was performed for each of the SNPs by χ^2 test or Fisher test. For significant associations, the genetic model that was

most representative of the effect of the variant (i.e. additive, dominant, or recessive) was tested as well. The genotype-associated risk was expressed as odds ratio (OR) with 95% confidence interval (CI). Survival differences in terms of event-free-survival (EFS) and overall survival (OS) were estimated using Kaplan-Meier analysis for patients with different genotypes and were assessed using log-rank test. Patients were followed for up to five years after the last therapeutic dose and an event was defined as induction failure, relapse, second malignancy or death from any cause. Combined effect of associated SNPs was tested by recoding genotypes as having none, one or two and more alleles at risk. Logistic regression was used for multivariate analysis which included beside genotypes: sex, age (< 10 years or ≥ 10 years), risk (standard or high), DFCI protocol and asparaginase formulation (E.coli or Erwinia) as categorical variables. Statistical analyses were performed with IBM SPSS Statistics for Windows, Version 22.0. (IBM Corp. Armonk, NY).

4.5.4. Risk Prediction

Weighted genetic risk score (wGRS) method was used to predict the risk of developing ASNase induced pancreatitis based on the cumulative combined effect of all SNPs found to be associated with this toxicity in the current study. The wGRS was estimated from the number of risk alleles by calculating the sum of weighted ln(OR) for each allele as explained elsewhere.[13] The performance of the comprehensive model in classifying patients based on their individual wGRS was assessed by calculating the area under the receiver operator characteristic (ROC) curves.

4.6. Authorship Contributions

M.K. designed the study; R.A. and V.G., performed experiments; F.C. performed medical chart reviews; C.L. JM.L., S.E.S., D.N. J.K, L.B.S and D.S. contributed to sample and clinical data collection and interpretation; D.S. supervised whole exome sequencing; R.A, J-F.S., CJ.X, V.G. and M.K. performed the data analysis; R.A. drafted the article; All authors revised the manuscript.

4.7. Acknowledgements

We thank all patients and their parents who consented to participate in genetics studies related to leukemia. Leukemia Lymphoma Society of Canada, Canadian Institutes of Health Research, Charles Bruneau Foundation and Dana-Farber Cancer Institute supported this study.

4.8. Disclosure of Conflicts of Interest

The authors declare no competing financial interests.

4.9. Funding

Dr. Krajinovic received grants for this project from Leukemia Lymphoma Society of Canada and Canadian Institutes of Health Research (#106458). Dana-Farber Cancer Institute ALL treatment protocols are supported by the National Cancer Institute/NIH grant 5 P01CA06848.

4.10. References

1. Ko RH, Jones TL, Radvinsky D, Robison N, Gaynon PS, Panosyan EH, Avramis IA, Avramis VI, Rubin J, Ettinger LJ, Seibel NL, Dhall G. Allergic reactions and antiasparaginase antibodies in children with high-risk acute lymphoblastic leukemia: A children's oncology group report. Cancer. 2015; 121: 4205-11. doi: 10.1002/cncr.29641.

2. Ben Tanfous M, Sharif-Askari B, Ceppi F, Laaribi H, Gagne V, Rousseau J, Labuda M, Silverman LB, Sallan SE, Neuberg D, Kutok JL, Sinnett D, Laverdiere C, et al. Polymorphisms of asparaginase pathway and asparaginase-related complications in children with acute lymphoblastic leukemia. Clin Cancer Res. 2015; 21: 329-34. doi: 10.1158/1078-0432.CCR-14-0508.

3. Pui CH, Mullighan CG, Evans WE, Relling MV. Pediatric acute lymphoblastic leukemia: where are we going and how do we get there? Blood. 2012; 120: 1165-74. doi: 10.1182/blood-2012-05-378943.

4. Hijiya N, van der Sluis IM. Asparaginase-associated toxicity in children with acute lymphoblastic leukemia. Leuk Lymphoma. 2016; 57: 748-57. doi: 10.3109/10428194.2015.1101098.

5. Duarte X, Esteves S, Neto AM, Pereira F. Incidence and risk factors for Central Nervous System thrombosis in paediatric acute lymphoblastic leukaemia during intensive asparaginase treatment: a single-centre cohort study. Br J Haematol. 2016; 174: 280-91. doi: 10.1111/bjh.14048.

6. Silverman LB, Gelber RD, Dalton VK, Asselin BL, Barr RD, Clavell LA, Hurwitz CA, Moghrabi A, Samson Y, Schorin MA, Arkin S, Declerck L, Cohen HJ, et al. Improved outcome for children with acute lymphoblastic leukemia: results of Dana-Farber Consortium Protocol 91-01. Blood. 2001; 97: 1211-8. doi:

7. Vrooman LM, Kirov, II, Dreyer ZE, Kelly M, Hijiya N, Brown P, Drachtman RA, Messinger YH, Ritchey AK, Hale GA, Maloney K, Lu Y, Plourde PV, et al. Activity and Toxicity of Intravenous Erwinia Asparaginase Following Allergy to E. coli-Derived Asparaginase in Children and Adolescents With Acute Lymphoblastic Leukemia. Pediatr Blood Cancer. 2016; 63: 228-33. doi: 10.1002/pbc.25757.

8. Liu C, Yang W, Devidas M, Cheng C, Pei D, Smith C, Carroll WL, Raetz EA, Bowman WP, Larsen EC, Maloney KW, Martin PL, Mattano LA, Jr., et al. Clinical and Genetic Risk Factors for Acute Pancreatitis in Patients With Acute Lymphoblastic Leukemia. J Clin Oncol. 2016; 34: 2133-40. doi: 10.1200/JCO.2015.64.5812.

9. Mizrahi T, Leclerc JM, David M, Ducruet T, Robitaille N. ABO Group as a Thrombotic Risk Factor in Children With Acute Lymphoblastic Leukemia: A Retrospective Study of

523 Patients. J Pediatr Hematol Oncol. 2015; 37: e328-32. doi: 10.1097/MPH.0000000000000333.

10. Nowak-Gottl U, Wermes C, Junker R, Koch HG, Schobess R, Fleischhack G, Schwabe D, Ehrenforth S. Prospective evaluation of the thrombotic risk in children with acute lymphoblastic leukemia carrying the MTHFR TT 677 genotype, the prothrombin G20210A variant, and further prothrombotic risk factors. Blood. 1999; 93: 1595-9. doi:

11. Stranger BE, Stahl EA, Raj T. Progress and promise of genome-wide association studies for human complex trait genetics. Genetics. 2011; 187: 367-83. doi: 10.1534/genetics.110.120907.

12. Mullighan CG. The molecular genetic makeup of acute lymphoblastic leukemia. Hematology Am Soc Hematol Educ Program. 2012; 2012: 389-96. doi: 10.1182/asheducation-2012.1.389.

13. De Jager PL, Chibnik LB, Cui J, Reischl J, Lehr S, Simon KC, Aubin C, Bauer D, Heubach JF, Sandbrink R, Tyblova M, Lelkova P, the steering committees of studies evaluating I-b, et al. Integrating genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility. Lancet neurology. 2009; 8: 1111-9. doi: 10.1016/S1474-4422(09)70275-3.

14. Kawedia JD, Liu C, Pei D, Cheng C, Fernandez CA, Howard SC, Campana D, Panetta JC, Bowman WP, Evans WE, Pui CH, Relling MV. Dexamethasone exposure and asparaginase antibodies affect relapse risk in acute lymphoblastic leukemia. Blood. 2012; 119: 1658-64. doi: 10.1182/blood-2011-09-381731.

15. Mori S, Bernardi R, Laurent A, Resnati M, Crippa A, Gabrieli A, Keough R, Gonda TJ, Blasi F. Myb-binding protein 1A (MYBBP1A) is essential for early embryonic development, controls cell cycle and mitosis, and acts as a tumor suppressor. PLoS One. 2012; 7: e39723. doi: 10.1371/journal.pone.0039723.

16. George B, Horn D, Bayo P, Zaoui K, Flechtenmacher C, Grabe N, Plinkert P, Krizhanovsky V, Hess J. Regulation and function of Myb-binding protein 1A (MYBBP1A) in cellular senescence and pathogenesis of head and neck cancer. Cancer Lett. 2015; 358: 191-9. doi: 10.1016/j.canlet.2014.12.042.

17. Owen HR, Elser M, Cheung E, Gersbach M, Kraus WL, Hottiger MO. MYBBP1a is a novel repressor of NF-kappaB. J Mol Biol. 2007; 366: 725-36. doi: 10.1016/j.jmb.2006.11.099.

18. Cai L, Deng SL, Liang L, Pan H, Zhou J, Wang MY, Yue J, Wan CL, He G, He L. Identification of genetic associations of SP110/MYBBP1A/RELA with pulmonary tuberculosis in the Chinese Han population. Hum Genet. 2013; 132: 265-73. doi: 10.1007/s00439-012-1244-5.

19. Rakonczay Z, Jr., Hegyi P, Takacs T, McCarroll J, Saluja AK. The role of NF-kappaB activation in the pathogenesis of acute pancreatitis. Gut. 2008; 57: 259-67. doi: 10.1136/gut.2007.124115.

20. Luo SX, Li S, Zhang XH, Zhang JJ, Long GH, Dong GF, Su W, Deng Y, Liu Y, Zhao JM, Qin X. Genetic polymorphisms of interleukin-16 and risk of knee osteoarthritis. PLoS One. 2015; 10: e0123442. doi: 10.1371/journal.pone.0123442.

21. Azimzadeh P, Khorram Khorshid HR, Akhondi MM, Shirazi A. Association of interleukin-16 polymorphisms with disease progression and susceptibility in endometriosis. Int J Immunogenet. 2016; 43: 297-302. doi: 10.1111/iji.12281.

22. Khoshbakht T, Soosanabadi M, Neishaboury M, Kamali K, Karimlou M, Bazazzadegan N, Khorram Khorshid HR. An Association Study on IL16 Gene Polymorphisms with the Risk of Sporadic Alzheimer's Disease. Avicenna J Med Biotechnol. 2015; 7: 128-32. doi:

23. Bowler RP, Bahr TM, Hughes G, Lutz S, Kim YI, Coldren CD, Reisdorph N, Kechris KJ. Integrative omics approach identifies interleukin-16 as a biomarker of emphysema. OMICS. 2013; 17: 619-26. doi: 10.1089/omi.2013.0038.

24. Tong Z, Li Q, Zhang J, Wei Y, Miao G, Yang X. Association between interleukin 6 and interleukin 16 gene polymorphisms and coronary heart disease risk in a Chinese population. J Int Med Res. 2013; 41: 1049-56. doi: 10.1177/0300060513483405.

25. Liu XL, Du JZ, Zhou YM, Shu QF, Li YG. Interleukin-16 polymorphism is associated with an increased risk of ischemic stroke. Mediators Inflamm. 2013; 2013: 564750. doi: 10.1155/2013/564750.

26. Xue H, Gao L, Wu Y, Fang W, Wang L, Li C, Li Y, Liang W, Zhang L. The IL-16 gene polymorphisms and the risk of the systemic lupus erythematosus. Clin Chim Acta. 2009; 403: 223-5. doi: 10.1016/j.cca.2009.03.016.

27. Romani S, Hosseini SM, Mohebbi SR, Kazemian S, Derakhshani S, Khanyaghma M, Azimzadeh P, Sharifian A, Zali MR. Interleukin-16 gene polymorphisms are considerable host genetic factors for patients' susceptibility to chronic hepatitis B infection. Hepat Res Treat. 2014; 2014: 790753. doi: 10.1155/2014/790753.

28. Mo C-J, Peng Q-L, He Y, Wang J, Xie L, Li T-J, Li S, Qin X. Positive Association Between IL-16 rs11556218 T/G Polymorphism and Cancer Risk: a Meta-analysis. Asian Pacific Journal of Cancer Prevention. 2014; 15: 4697-703. doi: 10.7314/apjcp.2014.15.11.4697.

29. Sironen A, Kotaja N, Mulhern H, Wyatt TA, Sisson JH, Pavlik JA, Miiluniemi M, Fleming MD, Lee L. Loss of SPEF2 function in mice results in spermatogenesis defects and primary ciliary dyskinesia. Biol Reprod. 2011; 85: 690-701. doi: 10.1095/biolreprod.111.091132.

30. GeneCards. SPEF2 (Sperm Flagellar 2).

31. Rousseau J, Gagne V, Labuda M, Beaubois C, Sinnett D, Laverdiere C, Moghrabi A, Sallan SE, Silverman LB, Neuberg D, Kutok JL, Krajinovic M. ATF5 polymorphisms influence ATF function and response to treatment in children with childhood acute lymphoblastic leukemia. Blood. 2011; 118: 5883-90. doi: 10.1182/blood-2011-05-355560.

32. Chen SH, Pei D, Yang W, Cheng C, Jeha S, Cox NJ, Evans WE, Pui CH, Relling MV. Genetic variations in GRIA1 on chromosome 5q33 related to asparaginase hypersensitivity. Clin Pharmacol Ther. 2010; 88: 191-6. doi: 10.1038/clpt.2010.94.

33. Fernandez CA, Smith C, Yang W, Date M, Bashford D, Larsen E, Bowman WP, Liu C, Ramsey LB, Chang T, Turner V, Loh ML, Raetz EA, et al. HLA-DRB1*07:01 is associated with a higher risk of asparaginase allergies. Blood. 2014; 124: 1266-76. doi: 10.1182/blood-2014-03-563742.

34. Mei L, Ontiveros EP, Griffiths EA, Thompson JE, Wang ES, Wetzler M. Pharmacogenetics predictive of response and toxicity in acute lymphoblastic leukemia therapy. Blood Rev. 2015; 29: 243-9. doi: 10.1016/j.blre.2015.01.001.

35. Silverman LB, Stevenson KE, O'Brien JE, Asselin BL, Barr RD, Clavell L, Cole PD, Kelly KM, Laverdiere C, Michon B, Schorin MA, Schwartz CL, O'Holleran EW, et al. Long-term results of Dana-Farber Cancer Institute ALL Consortium protocols for children with newly diagnosed acute lymphoblastic leukemia (1985-2000). Leukemia. 2010; 24: 320-34. doi: 10.1038/leu.2009.253.

36. Kearney SL, Dahlberg SE, Levy DE, Voss SD, Sallan SE, Silverman LB. Clinical course and outcome in children with acute lymphoblastic leukemia and asparaginase-associated pancreatitis. Pediatr Blood Cancer. 2009; 53: 162-7. doi: 10.1002/pbc.22076.

37. Grace RF, Dahlberg SE, Neuberg D, Sallan SE, Connors JM, Neufeld EJ, Deangelo DJ, Silverman LB. The frequency and management of asparaginase-related thrombosis in paediatric and adult patients with acute lymphoblastic leukaemia treated on Dana-Farber Cancer Institute consortium protocols. Br J Haematol. 2011; 152: 452-9. doi: 10.1111/j.1365-2141.2010.08524.x.

38. Healy J, Richer C, Bourgey M, Kritikou EA, Sinnett D. Replication analysis confirms the association of ARID5B with childhood B-cell acute lymphoblastic leukemia. Haematologica. 2010; 95: 1608-11. doi: 10.3324/haematol.2010.022459.

39. Baccichet A, Qualman SK, Sinnett D. Allelic loss in childhood acute lymphoblastic leukemia. Leuk Res. 1997; 21: 817-23. doi:

40. BroadInstitiute. GITHUB-PICARD.

41. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The Genome Analysis Toolkit: a MapReduce

framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20: 1297-303. doi: 10.1101/gr.107524.110.

42. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25: 2078-9. doi: 10.1093/bioinformatics/btp352.

43. Spinella JF, Mehanna P, Vidal R, Saillour V, Cassart P, Richer C, Ouimet M, Healy J, Sinnett D. SNooPer: a machine learning-based method for somatic variant identification from low-pass next-generation sequencing. BMC Genomics. 2016; 17: 912. doi: 10.1186/s12864-016-3281-2.

44. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010; 38: e164. doi: 10.1093/nar/gkq603.

45. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009; 4: 1073-81. doi: 10.1038/nprot.2009.86.

46. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet. 2013; Chapter 7: Unit7 20. doi: 10.1002/0471142905.hg0720s76.

47. Genomes Project C, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA. A map of human genome variation from population-scale sequencing. Nature. 2010; 467: 1061-73. doi: 10.1038/nature09534.

48. Server EV. (Seattle, WA.

49. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81: 559-75. doi: 10.1086/519795.

50. Storey JD, Taylor JE, Siegmund D. Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2004; 66: 187-205. doi: 10.1111/j.1467-9868.2004.00439.x.

51. Labuda D, Krajinovic M, Richer C, Skoll A, Sinnett H, Yotova V, Sinnett D. Rapid detection of CYP1A1, CYP2D6, and NAT variants by multiplex polymerase chain reaction and allele-specific oligonucleotide assay. Anal Biochem. 1999; 275: 84-92. doi: 10.1006/abio.1999.4293.

4.11. Tables

Table 1. Characteristics of the discovery and the replication cohort.

Cohort Characteristics		QcALL	DFCI	p-Value
Total Included		**302**	**282**	
Sex	Female	139 (46%)	129 (45,7%)	1
	Male	163 (54%)	153 (54,3%)	
WBC	< $50x10^3/\mu L$	257 (85,1%)	229 (81,2%)	0,2
	> $50x10^3/\mu L$	45 (14,9%)	53 (18,8%)	
Age	< 10 years	242 (80,1%)	230 (81,6%)	0,7
	≥ 10 years	60 (19,9%)	52 (18,4%)	
Risk	Standard	151 (50%)	173 (61,3%)	0,007
	High	151 (50%)	109 (38,7%)	
Source of Asparaginase	E. Coli	273 (90,4%)	261 (92,6%)	0,4
	Erwinia	29 (9,6%)	21 (7,4%)	
DFCI Protocol	00-01	111 (36,8%)	187 (66,3%)	$6x10^{-5}$
	95-01	119 (39,4%)	95 (33,7%)	
	91-01	55 (18,2%)	-	-
	87-01	17 (5,6%)	-	

Table 2. Top-ranking signals from the exome-wide association study confirmed by genotyping.

Toxicity	Gene_SNP Genotype	Complication		OR (95%-CI)	P	Model	Complication		OR (95%-CI)	P
		+	-				+	-		
Allergy	***SLC7A13*_rs9656982: A > G***									
	AA	37 (77,1%)	217 (87,2%)	1	1				2,1 (1,1-3,9)	**0,02**
	AG	8 (16,7%)	30 (12,1%)	1,6 (0,7-3,7)	0,3					
	GG	3 (6,3%)	2 (0,8%)	8,8 (1,4-54,5)	0,03					
	MYBBP1A*_rs3809849: G > C									
	GG	20 (41,7%)	160 (65%)	1	1				2,4 (1,4-3,9)	**$6x10^{-4}$**
	GC	23 (47,9%)	79 (32,1%)	2,3 (1,2-4,5)	0,01					
	CC	5 (10,4%)	7 (2,9%)	5,7 (1,7-19,7)	0,01					
	***YTHDC2*_rs75714066: G > C**									
	GG	37 (77,1%)	232 (91,3%)	1	1	**GG**	37 (77,1%)	232 (91,3%)	1	-
	GC	11 (22,9%)	21 (8,3%)	3,3 (1,5-7,4)	0,005	**GC+CC**	11 (22,9%)	22 (8,7%)	3,1 (1,4-7,0)	**0,008**
	CC	0 (0%)	1 (0,4%)	NA	-					
Pancreatitis	***ADAMTS17*_rs72755233: G > A**									
	GG	7 (46,7%)	232 (83,2%)	1	1	**GG**	7 (46,7%)	232 (83,1%)	1	-
	GA	8 (53,3%)	45 (16,1%)	5,9 (2-17,1)	0,002	**GA+AA**	8 (53,3%)	47 (16,9%)	5,6 (1,9-16,3)	**0,002**
	AA	0 (0%)	2 (0,7%)	NA	-					
	***MYBBP1A*_rs3809849: G > C**									
	GG	3 (20%)	177 (63,4%)	1	1	**GG**	3 (20%)	177 (63,4%)	1	-
	GC	12 (80%)	90 (32,3%)	7,9 (2,2-28,6)	0,0005	**GC+CC**	12 (80%)	102 (36,6%)	6,9 (1,9-25,2)	**0,002**
	CC	0 (0%)	12 (4,3%)	NA	-					
	SPECC1*_rs9908032: C > G									
	CC	8 (53,3%)	228 (80,6%)	1	1				3,9 (1,6-9,2)	**$8x10^{-4}$**
	CG	5 (33,3%)	53 (18,7%)	2,7 (0,8-8,5)	0,1					
	GG	2 (13,3%)	2 (0,7%)	28,5 (3,6-228,8)	0,009					

Thrombosis										
	***PKD2L1*_rs6584356: C > A**									
	CC	7 (70%)	257 (92,1%)	1	1	**CC**	7 (70%)	257 (92,1%)	1	-
	CA	2 (20%)	22 (7,9%)	3,3 (0,7-17)	0,2	**CA+AA**	3 (30%)	22 (7,9%)	5 (1,2-20,7)	**0,05**
	AA	1 (10%)	0 (0%)	NA	-					
	***RIN3*_rs3742717: C > T**									
	CC	6 (60%)	219 (77,7%)	1	1	**CC+CT**	8 (80%)	277 (98,2%)	13,8 (2,3-82,5)	**0,02**
	CT	2 (20%)	58 (20,6%)	1,3 (0,2-6,4)	1					
	TT	2 (20%)	5 (1,8%)	14,6 (2,3-91)	0,02	**TT**	2 (20%)	5 (1,8%)		
	***SPEF2*_rs34708521: G > A**									
	GG	5 (62,5%)	242 (91%)	1	1	**GG**	5 (62,5%)	242 (91%)	1	-
	GA	3 (37,5%)	23 (8,7%)	6,3 (1,4-28,1)	0,03	**GA+AA**	3 (37,5%)	24 (9%)	6,1 (1,4-26,9)	**0,03**
	AA	0 (0%)	1 (0,4%)	NA	-					
	SLC39A12*_rs62619938: C > T									
	CC	6 (60%)	262 (91%)	1	1				4,4 (1,6-11,7)	**$5x10^{-4}$**
	CT	3 (30%)	23 (8%)	5,7 (1,3-24,3)	0,04					
	TT	1 (10%)	3 (1%)	14,6 (1,3-161)	0,1					
	***MPEG1*_rs7926933: G > A**									
	GG	4 (44,4%)	234 (82,1%)	1	1	**GG**	4 (44,4%)	234 (82,1%)	1	-
	GA	5 (55,6%)	45 (15,8%)	6,5 (1,7-25,1)	0,009	**GA+AA**	5 (55,6%)	51 (17,9%)	5,7 (1,5-22,1)	**0,01**
	AA	0 (0%)	6 (2,1%)	NA	-					
	***IL16*_rs11556218: T > G**									
	TT	4 (50%)	238 (88,2%)	1	1	**TT**	4 (50%)	238 (88,1%)	1	-
	TG	4 (50%)	30 (11,1%)	7,9 (1,9-33,4)	0,009	**TG+GG**	4 (50%)	32 (11,9%)	7,4 (1,8-31,2)	**0,01**
	GG	0 (0%)	2 (0,7%)	NA	-					

Table 3. Performance of the comprehensive genetic model in predicting the risk of pancreatitis.

Cohort	AUC ± SD.	95% CI	P	Sensitivity	Specificity
QcALL	0,80 ± 0,062	68,1 ~ 92,6	$1x10^{-4}$	71%	81%
DFCI	0,78 ± 0,076	63,0 ~ 92,9	$3x10^{-3}$	70%	77%
Combined	0,80 ± 0,049	70,1 ~ 89,1	$1x10^{-6}$	71%	79%

4.12. Table and Figure Legends

Table 1. Characteristics of the discovery and the replication cohort.

QcALL, Quebec Childhood ALL cohort; DFCI, Dana-Farber Cancer Institute ALL Consortium cohort.

Table 2. Top-ranking signals from the exome-wide association study confirmed by genotyping.

The SNPs are presented as a change from major to minor alleles. OR, odds ratio; CI, confidence interval. Analysis in both co-dominant model and a model that best fits the data are presented. The final models are either dominant, recessive or additive; the latter is indicted by asterisk. NA, not analyzed due to low numbers.

Table 3. Performance of the comprehensive genetic model in predicting the risk of pancreatitis.

The data were extracted from the receiver operator characteristic (ROC) curves of the comprehensive model for pancreatitis which combines the 5 SNPs associated with this toxicity. The curves were produced by plotting the sensitivity against (1-specificty) of the model using weighted genetic risk scores to estimate the area under the curve in each cohort. The sensitivity and specificity reported in this table are based on an odds ratio greater than 11 for the risk of developing pancreatitis.

AUC, Area Under the Curve; SD, standard deviation; QcALL, Quebec Childhood ALL cohort; DFCI, Dana-Farber Cancer Institute ALL Consortium cohort.

Figure 1. The selection process following the exome-wide association study.

Top-ranking signals from the EWAS (N=115) were filtered through a multi-step selection process explained on the right-side of the figure. Each circle contains all the SNPs that are inside of it, including the ones in the smaller circles. Inner circle represent significant associations with one of the 3 asparaginase related toxicities (N=12) retained for analysis in replication cohort. rs3809849 in *MYBBP1A* was significantly associated both with allergy and pancreatitis in the EWAS study.

Figure 2. Top-ranking EWAS signals common for several asparaginase-related toxicities

SNPs that showed significant associations with one of the asparaginase-related toxicities were further tested for possible associations with the remaining side-effects. Association with thrombosis in **a)** and pancreatitis in **b)** and **c).** The studied association with the OR and 95% CI in brackets is indicated on the top of the graph. The frequency of patients with and without toxicity is represented by red and blue bars, respectively. The number of patients is shown on the top of each bar and the genotypes are indicated at the bottom of the graphs.

Figure 3. Association of rs3809849 in *MYBBP1A* gene with ASNase-related toxicities and with event free- and overall survival.

a) The frequency of patients with at least one asparaginase-related toxicity and without any toxicity is represented by the red and blue part of the bar, respectively. The number of samples per category is displayed inside of the bars. The OR with the 95% CI is given when

compared to patients with no variant allele (top of the graph) and across all genotype groups (bottom of the graph). **b)** The p-values obtained by the log rank test for the difference across genotypes are provided on each plot. The number of patients represented by each genotype and number of patients with event (in brackets) are indicated next to each curve. Hazard-ratios (HR) obtained through Cox-regression analysis are given with 95% CI.

Figure 4. Combined-effect model.

Combined-effect model of the variants associated with allergy (a), thrombosis (b) and pancreatitis (c).

Each bar represents the number of the variant alleles (i.e. none, one, two or more). The frequency of patients with and without toxicity is represented by the red and blue part of the bar, respectively. The number of samples per category is displayed inside of the bars. The OR with the 95% CI is given when compared to patients with no variants allele (top of the graph) and across genotype groups with increasing number of minor alleles (bottom of the graph).

Figure 5. Replication analysis in the independent validation cohort.

Association of pancreatitis with genetic variations in *MYBBP1A (a), IL16 (b)*, *SPEF2 (c)* and in combined effect model (d).

The frequency of patients with and without pancreatitis in a), b) and c) is represented by red and blue bars, respectively. The number and the genotypes are indicated. Combined-effect model in d) includes SNPs identified for association with pancreatitis through EWAS

of discovery cohort (i.e. rs72755233 in *ADAMTS17*, rs3809849 in *MYBBP1A* and rs9908032 in *SPECC1*). Each bar represents the number of the variant alleles present (i.e. none, one, two or more). The frequency of patients with and without toxicity is represented by the red and blue part of the bar, respectively. The number of samples per category is displayed inside of the bars. The OR with the 95% CI is given when compared to patients with no variants allele (top of the graph) and across groups (bottom of the graph).

Figure 6. Distribution of patients with pancreatitis among risk groups established using wGRS from the comprehensive combined-effect model in QcALL & DFCI cohort. Risk groups (S, standard; L, low; I, intermediate and H, high) represent the categorical distribution of weighted genetic risk scores (wGRS) of the Comprehensive Combined-effect model containing the 5 SNP associated with pancreatitis in this study (i.e. rs72755233 in *ADAMTS17*, rs3809849 in *MYBBP1A*, rs9908032 in *SPECC1*, rs11556218 in *IL16* and rs34708521 in *SPEF2)*. The wGRS values in **a)** were calculated from the discovery cohort and were used to predict the odds ratios in the validation cohort **b)**. The frequency of patients with pancreatitis in each risk group is displayed as a blue lined histogram reflecting the percentage out of the total number of cases. Log(OR) for pancreatitis susceptibility for each risk group (red circle) with a 95% confidence interval and the p-value for the trend across the groups are provided. The groups correspond to the following OR cut-off values: S (1); L (>1); I (>3.4) and Q4 (>10.3) as predicted from the QcALL cohort. The observed ORs per risk group in the DFCI cohort are also provided.

4.13. Figures

Figure 1

Figure 1. The selection process following the exome-wide association study.

Figure 2

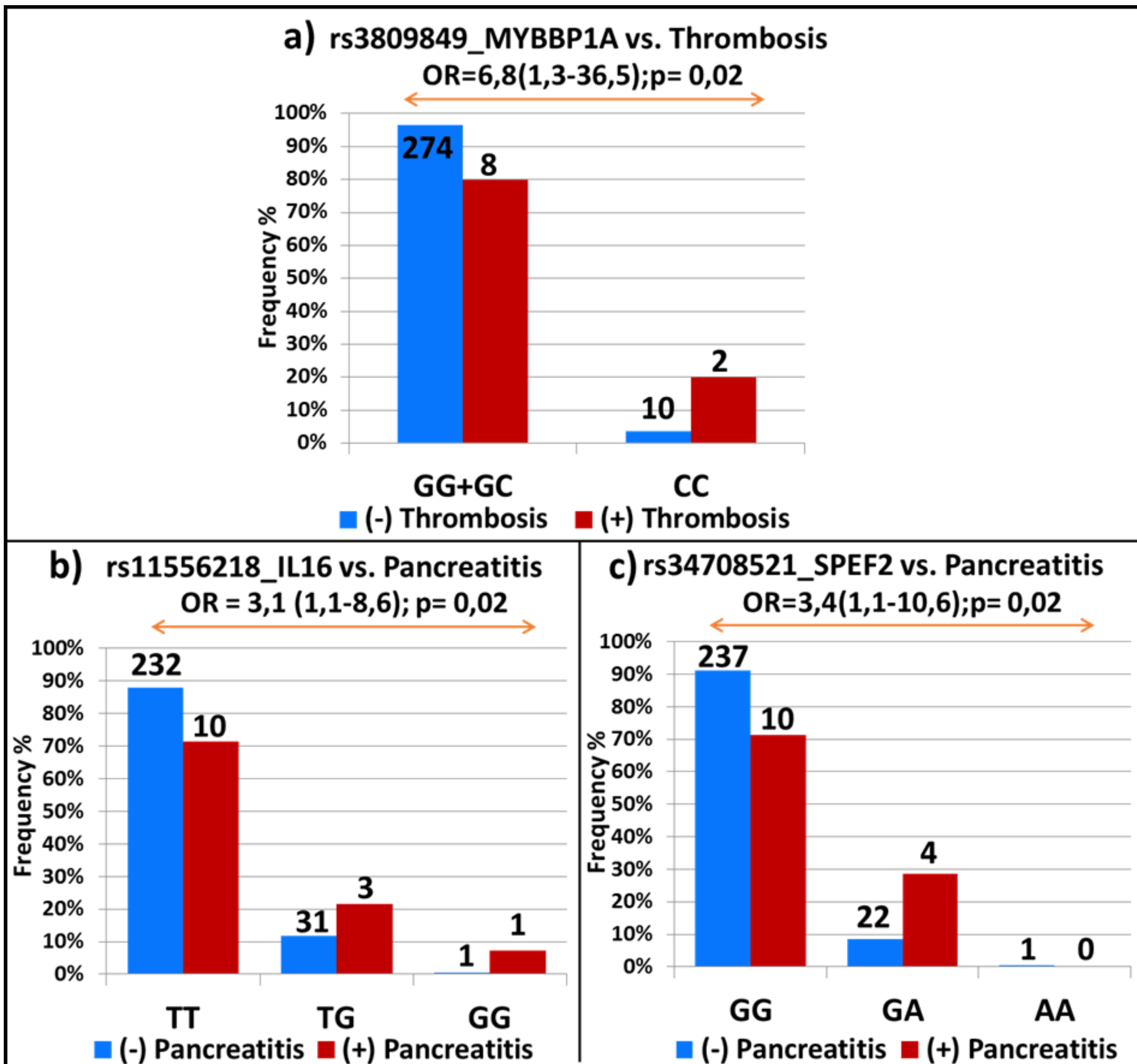


Figure 2. Top-ranking EWAS signals common for several asparaginase-related toxicities

Figure 3

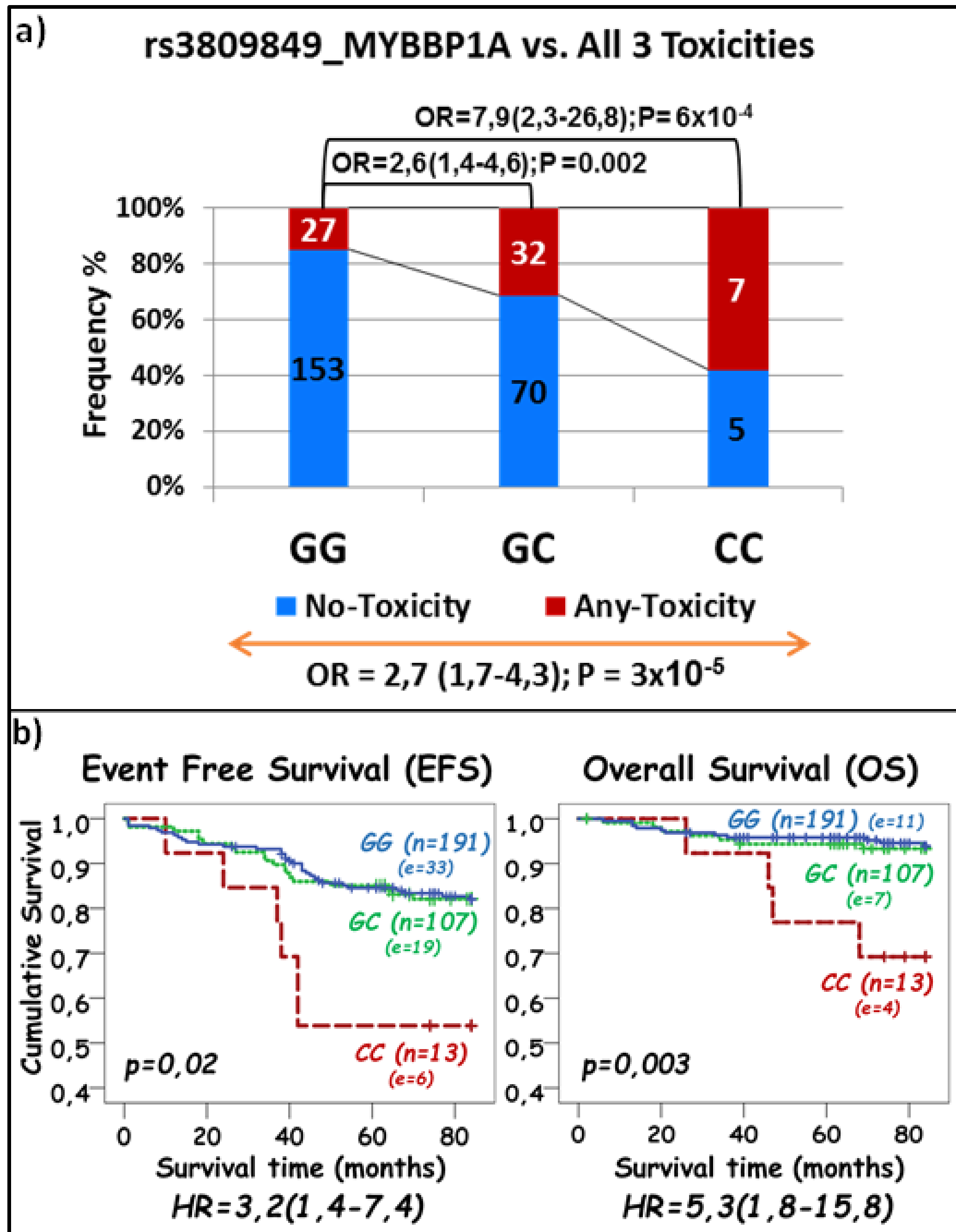


Figure 3. Association of rs3809849 in MYBBP1A gene with ASNase-related toxicities and with event free- and overall survival.

Figure 4

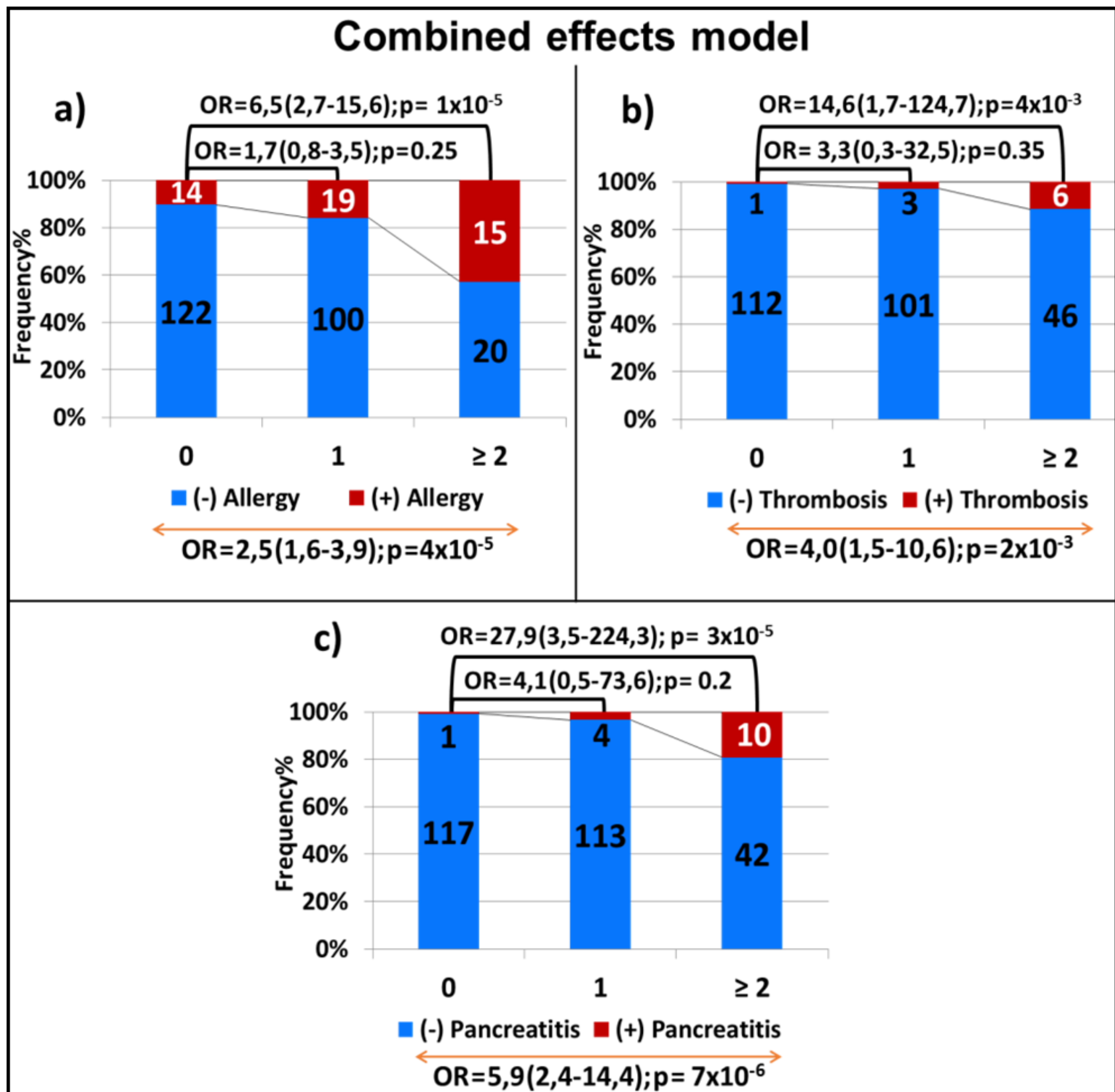


Figure 4. Combined-effect model.

Figure 5

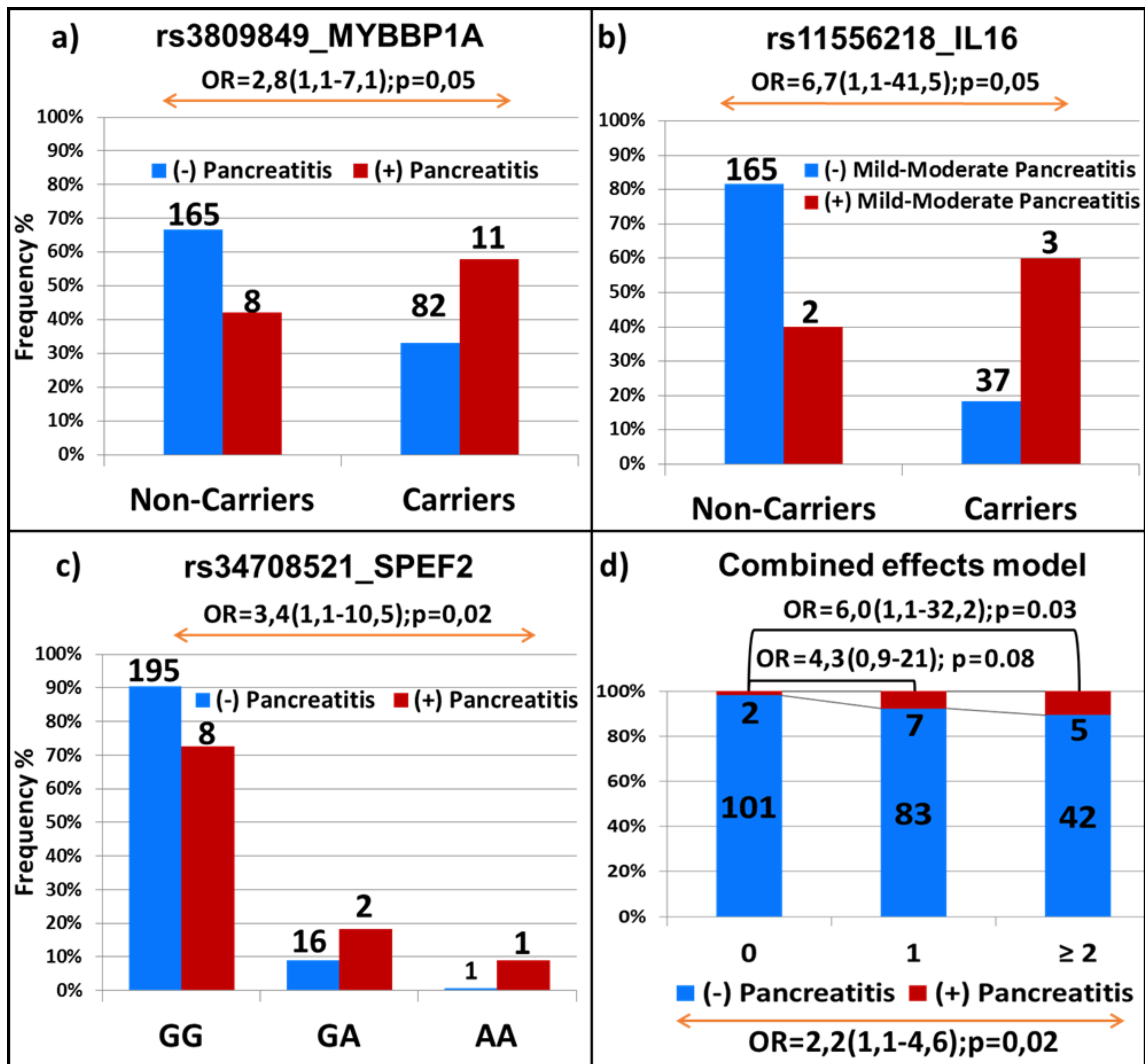


Figure 5. Replication analysis in the independent validation cohort.

Figure 6

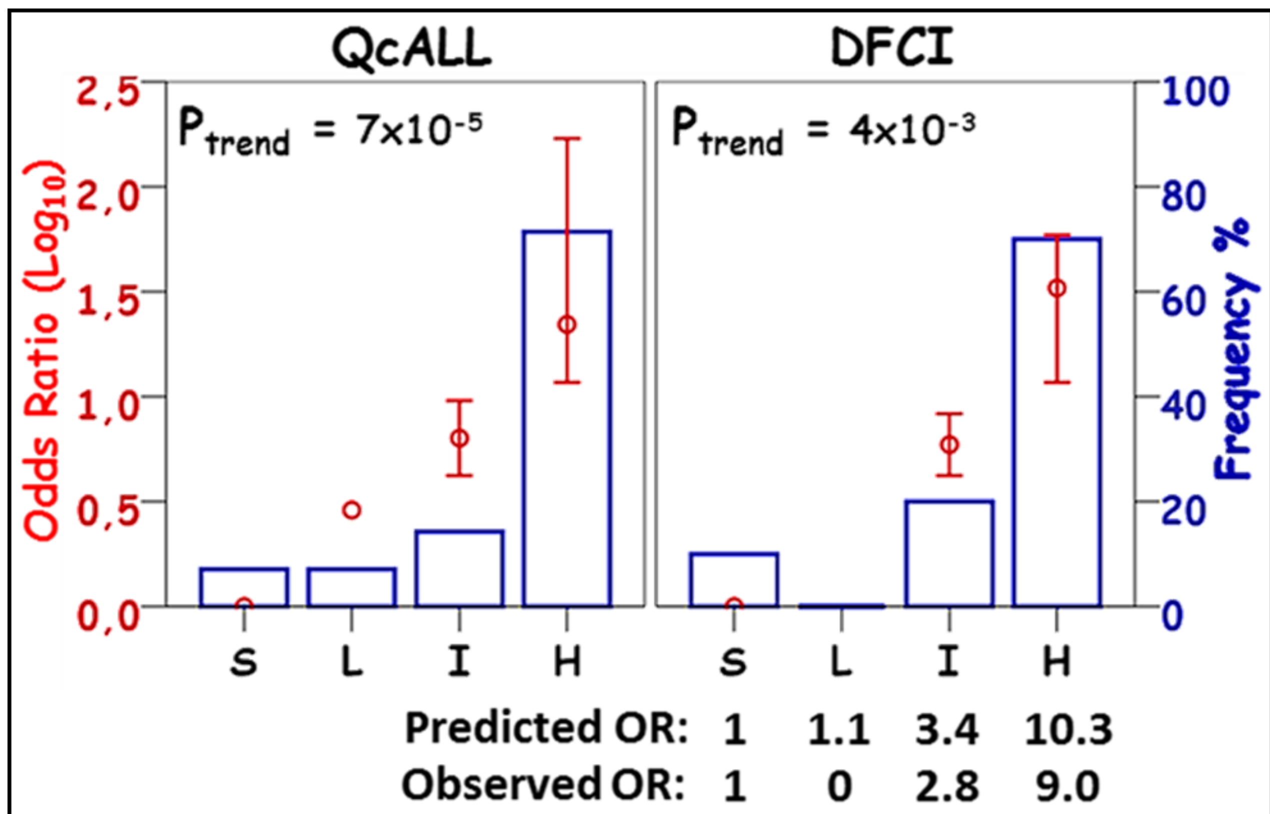


Figure 6. Distribution of patients with pancreatitis among risk groups established using wGRS from the comprehensive combined-effect model in QcALL & DFCI cohort.

4.14. Supplemental Material

Supplemental Table S1. Association of WES data with asparaginase-related complications

Gene_SNP	Toxicity	Minor Allele Frequency (%)	P-value Allelic Association (Sequencing Data)	FDR	P-value Association by Genotype (Sequencing Data)	P-value Association by Genotype (Genotyping Data)*
KIAA1107_rs565156: C > T	Allergy	1%	1,23E-05	1,4%	2,00E-03	-
MAP3K1_rs702689: G > C	Allergy	1%	1,23E-05	1,4%	2,00E-03	-
CPNE1_rs12481228: G > C	Allergy	2%	1,96E-05	1,5%	8,31E-04	-
SLC7A13_rs9656982: A > G	**Allergy**	**9%**	**1,34E-04**	**7,4%**	**9,91E-04**	***1,63E-02***
KCNJ15_rs2230033: G > A	Allergy	50%	3,57E-04	8,9%	1,81E-03	-
OTOF_rs4665855: G > A	Allergy	32%	5,18E-04	8,9%	1,53E-03	-
HLA-DPA1_rs199711661: T > C	**Allergy**	**5%**	**8,57E-04**	**8,9%**	**3,42E-03**	1,58E-01
CETP_rs5880: G > C	Allergy	3%	1,01E-03	8,9%	8,00E-04	-
PRR15_rs112093295: C > A	**Allergy**	**6%**	**1,25E-03**	**8,9%**	**3,66E-03**	1,29E-01
ZNF880_rs14048: G > A	Allergy	2%	1,43E-03	8,9%	1,30E-03	-
GTPBP5_rs6062133: G > A	Allergy	2%	1,50E-03	8,9%	1,40E-03	-
ALOXE3_rs3027229: G > C	Allergy	6%	1,61E-03	8,9%	1,33E-03	-
PKD2L2_rs1880458: G > A	Allergy	1%	1,92E-03	8,9%	3,88E-02	-
PCDHA7_rs61730623: G > A	Allergy	1%	2,31E-03	9,8%	2,20E-03	-
OR2M7_rs7555310: A > G	Allergy	1%	2,40E-03	9,8%	2,30E-03	-
LILRB3_rs61734493: C > A	Allergy	1%	2,40E-03	9,8%	2,30E-03	-
ARAP3_rs1031904: C > G	Allergy	2%	2,42E-03	9,8%	1,78E-02	-
SLC22A25_rs11231397: C > G	Allergy	8%	2,51E-03	9,9%	1,38E-02	-
FLG_rs12405278: G > A	Allergy	3%	2,55E-03	9,9%	2,20E-03	-
C17orf80_rs745143: T > C	Allergy	50%	2,88E-03	10,9%	7,69E-03	-
KRT72_rs11170183: C > A	Allergy	4%	3,29E-03	12,1%	2,47E-02	-
LILRB2_rs386056: C > T	**Allergy**	**7%**	**3,47E-03**	**12,6%**	**3,13E-03**	1,62E-01
NACAD_rs61740895: G > A	Allergy	21%	3,91E-03	14,0%	2,40E-02	-
RP1L1_rs4841399: G > C	Allergy	4%	5,10E-03	18,0%	3,84E-02	-
ACSM2A_rs1133607: C > T	Allergy	3%	5,55E-03	19,0%	4,70E-03	-
MYBBP1A_rs3809849: G > C	**Allergy**	**11%**	**5,60E-03**	**19,0%**	**2,07E-03**	***2,34E-03***
CD6_rs11230563: C > T	Allergy	2%	5,66E-03	19,0%	3,82E-02	-
FBXL6_rs61746974: C > G	Allergy	3%	6,07E-03	19,8%	2,19E-02	-
DYNC2H1_rs17301182: C > T	Allergy	7%	6,07E-03	19,8%	1,09E-02	-
OR5D18_rs297081: A > G	Allergy	2%	1,24E-06	0,2%	6,00E-04	-
KCNMB3_rs7645550: C > T	Allergy	14%	1,20E-04	7,3%	1,10E-03	-

COL9A3_rs61734651: C > T	Allergy	4%	1,55E-04	7,3%	1,21E-03	-
SNX15_rs495820: C > T	Allergy	2%	1,88E-04	7,3%	9,14E-03	-
CFB_rs4151667: T > A	Allergy	2%	2,42E-04	7,3%	5,83E-03	-
FBXO24_rs11768465: C > T	**Allergy**	**11%**	**2,92E-04**	**7,3%**	**1,25E-03**	6,06E-01
TTC3_rs61999340: C > G	Allergy	6%	3,23E-04	7,3%	2,00E-04	-
YTHDC2_rs75714066: G > C	**Allergy**	**5%**	**3,51E-04**	**7,3%**	**2,00E-04**	***9,63E-03***
OR52J3_rs58664826: G > A	Allergy	7%	3,89E-04	7,3%	1,36E-03	-
SENP6_rs17414086: C > T	**Allergy**	**14%**	**6,00E-04**	**7,8%**	**1,26E-03**	6,98E-01
KIAA1551_rs3759302: T > A	Allergy	8%	7,54E-04	9,1%	7,83E-03	-
NPHS1_rs3814995: C > T	Allergy	5%	1,33E-03	15,0%	1,67E-02	-
FERMT1_rs62200482: G > A	Allergy	7%	1,43E-03	15,1%	1,17E-02	-
P2RY11,PPAN-P2RY11_rs3745601: G > A	Allergy	2%	1,92E-03	19,1%	2,13E-02	-
PKD2L1_rs6584356: C > A	**Thrombosis**	**6%**	**1,47E-09**	**0,0%**	**6,92E-12**	***2,88E-07***
MYO15A_rs712270: A > T	Thrombosis	2%	1,51E-09	0,0%	2,08E-09	-
PKD1L2_rs7185774: C > T	Thrombosis	11%	2,10E-06	0,0%	4,45E-07	-
RIN3_rs3742717: C > T	**Thrombosis**	**5%**	**1,87E-05**	**0,1%**	**8,00E-06**	***1,02E-03***
C2orf61_rs815804: G > T	Thrombosis	14%	1,95E-05	0,1%	3,85E-07	-
MYH7B_rs3746435: G > C	Thrombosis	4%	5,51E-05	0,3%	4,38E-05	-
CCDC135_rs3809611: C > T	Thrombosis	21%	1,36E-04	0,6%	3,18E-04	-
DHX37_rs11057939: C > T	Thrombosis	12%	2,32E-04	1,0%	1,07E-03	-
NRN1L_rs73594554: G > A	Thrombosis	9%	3,49E-04	1,5%	1,10E-06	-
SPEF2_rs34708521: G > A	**Thrombosis**	**5%**	**4,04E-04**	**1,6%**	**3,00E-04**	***2,30E-02***
ESYT2_rs2305475: A > G	**Thrombosis**	**5%**	**4,04E-04**	**1,6%**	**5,03E-04**	4,39E-01
POU6F2_rs2074936: C > T	Thrombosis	5%	4,89E-04	1,9%	1,06E-04	-
EPPK1_rs11781942: G > A	Thrombosis	14%	7,45E-04	2,7%	1,79E-08	-
NUP153_rs61744976: G > C	Thrombosis	16%	2,26E-03	6,7%	6,60E-03	-
CCDC41_rs74340001: G > A	Thrombosis	7%	2,66E-03	7,6%	1,16E-04	-
TLR3_rs3775291: C > T	**Thrombosis**	**28%**	**2,76E-03**	**7,7%**	**1,95E-03**	7,93E-02
PRR16_rs17853861: C > A	Thrombosis	16%	2,76E-03	7,7%	5,30E-04	-
FAM26F_rs11544160: G > A	Thrombosis	7%	3,41E-03	9,3%	2,30E-03	-
SFI1_rs16989291: T > C	Thrombosis	7%	3,81E-03	9,7%	2,99E-03	-
SSC5D_rs925878: C > T	Thrombosis	11%	3,86E-03	9,7%	1,53E-02	-
SLC39A12_rs62619938: C > T	**Thrombosis**	**7%**	**4,45E-03**	**10,4%**	**4,08E-03**	***2,27E-03***
TMEM123_rs11547915: C > A	Thrombosis	7%	4,45E-03	10,4%	3,10E-03	-
MPEG1_rs7926933: G > A	**Thrombosis**	**7%**	**4,45E-03**	**10,4%**	**7,47E-03**	***7,33E-03***
IL16_rs11556218: T > G	**Thrombosis**	**7%**	**4,45E-03**	**10,4%**	**7,47E-03**	***4,16E-03***
CSTL1_rs3746737: C > T	**Thrombosis**	**7%**	**4,45E-03**	**10,4%**	**2,78E-03**	***1,54E-02***
FAM198A_rs3732858: G > A	Thrombosis	7%	5,90E-03	12,6%	5,79E-03	-

LRRN2_rs11588857: G > A	Thrombosis	12%	6,04E-03	12,6%	1,27E-06	-
BRD8_rs11750814: G > A	Thrombosis	12%	6,37E-03	12,9%	1,70E-03	-
SGCG_rs17314986: G > A	Thrombosis	13%	8,59E-03	16,4%	7,00E-03	-
C20orf85_rs17440813: A > G	Thrombosis	2%	2,94E-11	0,0%	3,66E-07	-
PRR5L_rs62621409: A > G	**Thrombosis**	**5%**	**5,32E-04**	**9,6%**	**1,73E-05**	***1,18E-02***
F13A1_rs5988: C > G	Thrombosis	21%	9,43E-04	12,1%	8,48E-04	-
PKD1L1_rs76100363: G > A	Pancreatitis	2%	3,72E-08	0,0%	3,27E-07	-
OR5K3_rs13068323: G > A	Pancreatitis	10%	3,19E-05	0,5%	2,22E-04	-
PARP15_rs12489170: G > A	**Pancreatitis**	**11%**	**6,32E-05**	**0,9%**	**5,81E-08**	5,31E-02
ADAMTS17_rs72755233: G > A	**Pancreatitis**	**11%**	**1,57E-04**	**1,8%**	**8,27E-05**	***1,25E-03***
FBXL6_rs61746974: C > G	Pancreatitis	3%	1,86E-04	2,0%	1,25E-04	-
ELL3_rs2277531: G > C	Pancreatitis	6%	3,53E-04	3,6%	4,32E-04	-
DNAH9_rs3744581: A > G	Pancreatitis	13%	4,65E-04	3,8%	3,00E-04	-
PDZRN4_rs285584: G > A	**Pancreatitis**	**9%**	**6,21E-04**	**4,1%**	**5,65E-03**	5,35E-01
SPATA21_rs41269193: G > T	Pancreatitis	4%	6,49E-04	4,1%	1,42E-06	-
PREX1_rs41283558: C > G	Pancreatitis	13%	7,71E-04	4,7%	1,78E-04	-
PYCRL_rs2242089: C > T	**Pancreatitis**	**10%**	**9,73E-04**	**5,8%**	**1,30E-03**	9,52E-02
CCDC8_rs2279517: C > G	Pancreatitis	4%	1,18E-03	6,7%	9,00E-04	-
GSTZ1_rs7975: G > A	Pancreatitis	4%	1,29E-03	7,1%	2,44E-03	-
DNHD1_rs4282961: C > A	Pancreatitis	23%	1,99E-03	10,7%	2,55E-03	-
SEPT_4_rs17741424: T > A	**Pancreatitis**	**11%**	**2,30E-03**	**11,9%**	**1,17E-03**	
AKAP13_rs4075256: T > C	Pancreatitis	40%	2,60E-03	13,1%	1,16E-02	-
TJP2_rs77236826: A > G	**Pancreatitis**	**8%**	**2,75E-03**	**13,5%**	**9,62E-07**	2,95E-01
MYBBP1A_rs3809849: G > C	**Pancreatitis**	**11%**	**3,47E-03**	**16,1%**	**3,24E-03**	***2,34E-03***
OR52R1_rs7941731: A > G	Pancreatitis	35%	4,13E-03	18,5%	1,69E-02	-
PHLPP2_rs61733127: A > G	Pancreatitis	16%	4,20E-03	18,5%	1,44E-03	-
OR4D2_rs74730740: C > T	Pancreatitis	8%	4,64E-03	19,0%	9,33E-03	-
DFNB31_rs12339210: G > C	Pancreatitis	8%	4,64E-03	19,0%	9,33E-03	-
MUC16_rs12150888: G > T	**Pancreatitis**	**25%**	**4,65E-03**	**19,0%**	**1,24E-03**	4,36E-01
F13A1_rs5987: C > T	Pancreatitis	4%	3,35E-05	0,7%	2,20E-05	-
C19orf59_rs72996468: A > G	Pancreatitis	4%	4,73E-05	0,8%	1,00E-03	-
MAVS_rs7262903: C > A	Pancreatitis	16%	1,11E-04	1,4%	2,16E-04	-
NPSR1_rs7809642: C > T	Pancreatitis	5%	3,62E-04	2,9%	4,04E-06	-
GJB7_rs35259282: C > T	**Pancreatitis**	**5%**	**4,08E-04**	**2,9%**	**1,86E-06**	1,87E-01
FCRL6_rs61823162: C > T	Pancreatitis	9%	5,66E-04	2,9%	5,91E-04	-
LRRC31_rs35923425: C > G	**Pancreatitis**	**6%**	**1,29E-03**	**5,0%**	**1,33E-03**	2,78E-01
MMP17_rs11835665: G > A	Pancreatitis	6%	1,75E-03	5,9%	2,23E-03	-
RAB3GAP2_rs2289189: C > G	**Pancreatitis**	**7%**	**3,19E-03**	**9,3%**	**2,20E-03**	6,00E-02

SPECC1_rs9908032: C > G	**Pancreatitis**	**11%**	**3,84E-03**	**10,4%**	**1,37E-06**	***4,98E-05***
CARD10_rs9610775: C > T	Pancreatitis	7%	3,96E-03	10,4%	1,89E-02	-
UBD_rs2076485: A > G	**Pancreatitis**	**23%**	**4,42E-03**	**11,0%**	**3,26E-05**	9,22E-02
PCDH15_rs11004439: A > C	Pancreatitis	17%	5,37E-03	13,0%	3,54E-03	-
ERCC6_rs2228527: T > C	Pancreatitis	24%	5,84E-03	13,6%	2,15E-02	-
HRG_rs2228243: A > G	**Pancreatitis**	**18%**	**5,97E-03**	**13,6%**	**5,74E-03**	8,95E-01
C3orf20_rs9821143: G > A	Pancreatitis	31%	6,39E-03	13,6%	8,97E-03	-
GTF2A1L_rs940389: G > C	Pancreatitis	33%	1,18E-02	18,4%	6,81E-04	-

The SNPs selected for validation through genotyping are highlighted and those that remained significant are depicted in dark grey color. *p value reflects the difference across genotype groups regardless of genetic model. Further analysis in accordance to appropriate models is presented in Table 2. CSTL1 and PRR5L were not considered further, as association did not follow any genetic model.

Supplemental Table S2. Function and differential protein expression of genes selected for further investigation.

Gene	Toxicity	Full Name	Function	Protein Differential Expression
SLC7A13	**Allergy**	Solute Carrier Family 7 Member 13	Amino acid transmembrane transporter activity	Kidney Cortex and Testis
HLA-DPA1	**Allergy**	Major Histocompatibility Complex, Class II, DP Alpha 1	Central role in the immune system	B-lymphocyte, Lymph node and Lung
PRR15	**Allergy**	Proline Rich 15	May have a role in proliferation and/or differentiation	Plasma, Platelet, Fetal ovary, Heart, and Fetal testis
LILRB2	**Allergy**	Leukocyte Immunoglobulin Like Receptor B2	Immunoregulatory interactions between Lymphoid and non-Lymphoids cell and Immune System	Neutrophil, Monocytes and Cervix
MYBBP1A	**Allergy**	MYB Binding Protein 1a	Role in various cellular processes including response to nucleolar stress, tumor suppression and synthesis of ribosomal DNA	T-lymphocyte, Pancreas, Peripheral blood mononuclear cells and Heart
FBXO24	**Allergy**	F-Box Protein 24	Phosphorylation-dependent ubiquitination (ubiquitin-protein transferase activity)	Pancreatic juice, Platelets and Testis
YTHDC2	**Allergy**	YTH Domain Containing 2	Nucleic acid binding and helicase activity. Linked to susceptibility to pancreatic cancer in human patients	Pancreatic juice, Ovary and Cerebrospinal fluid
SENP6	**Allergy**	SUMO1/Sentrin Specific Peptidase 6	Cysteine-type peptidase activity and SUMO-specific protease activity	CD8 T-cells, Peripheral blood mononuclear cells and Testis
PKD2L1	**Thrombosis**	Polycystin 2 Like 1	Calcium-regulated nonselective cation channel	Plasma and Colon
RIN3	**Thrombosis**	Ras Interaction/Interference Protein 3	Vesicle-mediated transport, GTPase activator activity and Rab guanyl-nucleotide exchange factor activity	Peripheral blood mononuclear cells, Monocytes, NK cells, B-lymphocyte and Lymph node
SPEF2	**Thrombosis**	Sperm Flagellar 2	Protein dimerization activity	Platelet, Fetal Brain and Fetal ovary
ESYT2	**Thrombosis**	Extended Synaptotagmin 2	Calcium ion binding and phosphatidylinositol binding and may play a role in cellular lipid transport	Peripheral blood mononuclear cells

TLR3	**Thrombosis**	Toll Like Receptor 3	Induces the activation of NF-kappaB and the production of type I interferons. Fundamental role in pathogen recognition and activation of innate immunity	Megakaryocytes, Platelets, Immature Dendritic cells, Pancreas and Nasal epithelium
SLC39A12	**Thrombosis**	Solute Carrier Family 39 Member 12	Metal ion transmembrane transporter activity and zinc ion transmembrane transporter activity. Thought to be involved in platelet function	Heart, Retina and Frontal cortex
MPEG1	**Thrombosis**	Macrophage Expressed 1	Cell cycle. Pathogen Recognition and Activation of the Innate Immune Response.	Peripheral blood mononuclear cells, Monocytes and Testis
IL16	**Thrombosis**	Interleukin 16	Cytokine activity, Chemoattractant, a modulator of T cell activation, and an inhibitor of HIV replication	Lymph node, CD8 T-cells, T-lymphocyte, and Peripheral blood mononuclear cells
CSTL1	**Thrombosis**	Cystatin Like 1	Cysteine-type endopeptidase inhibitor activity found in a variety of human fluids and secretions.	Testis
PRR5L	**Thrombosis**	Proline Rich 5 Like	Ubiquitin protein ligase binding. Related pathways are mTOR signalling and PI3K / Akt Signaling	Spleen, Colon, and Brain
PARP15	**Pancreatitis**	Poly(ADP-Ribose) Polymerase Family Member 15	NAD+ ADP-ribosyltransferase activity: transfers ADP-ribose from nicotinamide dinucleotide (NAD) to Glu/Asp residues on the substrate protein	Adipocyte and B-lymphocyte
ADAMTS17	**Pancreatitis**	A Disintegrin-Like And Metalloprotease (Reprolysin Type) With Thrombospondin Type 1 Motif, 17	Endopeptidase and Metalloendopeptidase activity	Plasma and Esophagus
PDZRN4	**Pancreatitis**	PDZ Domain Containing Ring Finger 4	Ubiquitin-protein transferase activity and ubiquitin protein ligase activity. Potential role as tumor suppressor	Platelet
PYCRL	**Pancreatitis**	Pyrroline-5-Carboxylate Reductase-Like	Pyrroline-5-carboxylate reductase activity. Involved in Arginine and proline metabolism	Multiple tissues
SEPT_4	**Pancreatitis**	Septin 4	GTPase activity. Apoptosis Modulation and Signaling. Localized to the mitochondria, and has a role in apoptosis and cancer. May play a role in cytokinesis and platelet secretion	Frontal cortex, Spinal cord, Retina, and Spleen

TJP2	**Pancreatitis**	Tight Junction Protein 2	Blood-Brain Barrier and Immune Cell Transmigration: VCAM-1/CD106 Signaling Pathways	Peripheral blood mononuclear cells, Platelet, Nasal epithelium and Cervix
MYBBP1A	**Pancreatitis**	MYB Binding Protein 1a	Role in various cellular processes including response to nucleolar stress, tumor suppression and synthesis of ribosomal DNA	Pancreas, T-lymphocyte, Peripheral blood mononuclear cells and Heart
MUC16	**Pancreatitis**	Mucin-16	Role in immune system by providing a protective, lubricating barrier against particles and infectious agents at mucosal surfaces.	Platelet, Liver, Cervix and Breast. Expression is significantly increased in Pancreatic Cancer
GJB7	**Pancreatitis**	Gap Junction Protein Beta 7	Vesicle-mediated transport and Gap junction trafficking. Contributes to leukemia cell communication and chemosensitivity	Unspecified
LRRC31	**Pancreatitis**	Leucine Rich Repeat Containing 31	Unspecified	Thyroid, Tonsil and Esophagus
RAB3GAP2	**Pancreatitis**	RAB3 GTPase Activating Protein Subunit 2	Vesicle-mediated transport. Involved in regulated exocytosis of neurotransmitters and hormones	Lymph node
SPECC1	**Pancreatitis**	Sperm Antigen With Calponin Homology And Coiled-Coil Domains 1	Is a novel fusion partner to PDGFRB in juvenile myelomonocytic leukemia with t(5;17)(q33;p11.2).	Peripheral blood mononuclear cells, Lung, Testis and some cancer cell lines
UBD	**Pancreatitis**	Ubiquitin D	Regulates IRE1α/JNK-dependent apoptosis in pancreatic beta cells. Role in regulation of NF-kappa-B signaling	Islet of Langerhans, Monocytes, Liver, Ovary
HRG	**Pancreatitis**	Histidine Rich Glycoprotein	Binds heme, dyes and divalent metal ions. Involved in inhibition of fibrinolysis and the reduction of inhibition of coagulation. Normalizes tumor vessels and promotes antitumor immunity	Serum, Plasma, Monocytes and Synovial fluid

The name, function and differential protein expression is provided for each of the genes selected for confirmation by genotyping after the initial filtration of the EWAS signals. The genes were selected based on a biological function or expression profile that could be relevant to the studied toxicity.

Supplemental Table S3. Multivariate analysis for primary and secondary associations confirmed by genotyping

SNP	Toxicity	P	OR (95%-CI)
SLC7A13_rs9656982: A > G	Allergy	0,03	2,1 (1,1-4,0)[a]
MYBBP1A_rs3809849: G > C	Allergy	0,002	2,3 (1,4-3,9)[a]
YTHDC2_rs75714066: G > C	Allergy	0,003	3,7 (1,6-8,7)[d]
ADAMTS17_rs72755233: G > A	Pancreatitis	0,002	5,5 (1,9-16,5)[d]
MYBBP1A_rs3809849: G > C	Pancreatitis	0,003	7,3 (2,0-26,9)[d]
SPECC1_rs9908032: C > G	Pancreatitis	0,002	4,2 (1,7-10,5)[a]
PKD2L1_rs6584356: C > A	Thrombosis	0,05	5,1 (1,0-26,1)[d]
RIN3_rs3742717: C > T	Thrombosis	0,01	13,9 (1,7-115,3)[r]
SPEF2_rs34708521: G > A	Thrombosis	0,08	4,3 (0,8-22,3)[d]
SLC39A12_rs62619938: C > T	Thrombosis	0,004	5,8 (1,8-19,1)[a]
MPEG1_rs7926933: G > A	Thrombosis	0,02	5,2 (1,2-21,7)[d]
IL16_rs11556218: T > G	Thrombosis	0,02	6,0 (1,3-27,7)[d]
MYBBP1A_rs3809849: G > C	Thrombosis	0,01	12,1 (1,6-100,5)[r]
SPEF2_rs34708521: G > A	Pancreatitis	0,05	3,3 (1,0-10,8)[a]
IL16_rs11556218: T > G	Pancreatitis	0,03	3,4 (1,2-10,3)[a]
SNP	**Survival**	**P**	**HR (95%-CI)**
MYBBP1A_rs3809849: G > C	Event Free	0,007	3,8 (1,4-9,8)[r]
MYBBP1A_rs3809849: G > C	*Overall	0,002	7,6 (2,0-28,4)[r]

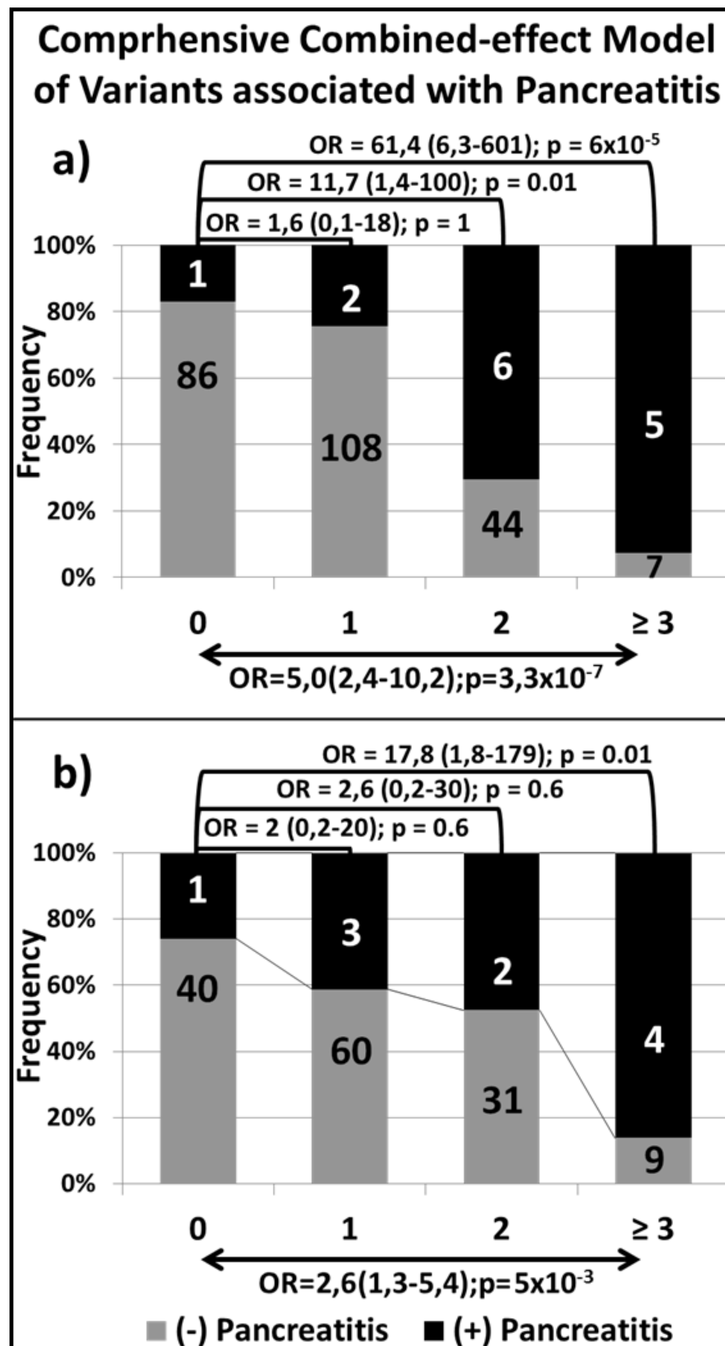
The regression models included genotypes and as covariates, age, sex, risk, DFCI protocol ASNase formulation Genotypes were coded according to genetic models presented in Table 2 (a: additive, d: dominant and r: recessive in OR column). OR, odds ratio; HR, hazard ratio.

Supplemental Table S4. Combined-cohort analysis performed for SNPs with reproducible associations with pancreatitis.

Gene_SNP Genotype	Complication		OR (95%-CI)	P	Combined Model	Complication		OR (95%-CI)	P
	+	-				+	-		
MYBBP1A_rs3809849: G > C									
GG	11 (32,4%)	342 (65,0%)	1	1	GG	11 (32,4%)	342 (65,0%)	1	*Ref.*
GC	22 (64,7%)	160 (30,4%)	4,3 (2,0-9,0)	0.00008	GC+CC	23 (67,6%)	184 (35,0%)	3,9 (1,8-8,1)	**0.0003**
CC	1 (2,9%)	24 (4,6%)	1,3 (0,2-10,5)	1					
SPEF2_rs34708521: G > A									
GG	18 (72,0%)	396 (90,8%)	1	1	GG	18 (72,0%)	396 (90,8%)	1	*Ref.*
GA	6 (24,0%)	38 (8,7%)	3,5 (1,3-9,3)	0.02	GA+AA	7 (28,0%)	40 (9,2%)	3,8 (1,5-9,8)	**0.008**
AA	1 (4%)	2 (0,5%)	11 (0,9-127)	0.1					
IL16_rs11556218: T > G									
TT	18 (66,7%)	391 (85,4%)	1	1	TT	18 (66,7%)	391 (85,4%)	1	*Ref.*
TG	8 (29,6%)	61 (13,3%)	2,8 (1,2-6,8)	0.02	TG+GG	9 (33,3%)	67 (14,6%)	2,9 (1,3-6,8)	**0.02**
GG	1 (3,7%)	6 (1,3%)	3,6 (0,4-31,6)	0.3					

The combined-cohort represents the pooled samples from the discovery and replication cohort (QcALL+DFCI) which gives rise to a cohort with a larger sample size of 584 patients. The SNPs are presented as a change from major to minor alleles. OR, odds ratio; CI, confidence interval. Analysis in both co-dominant and dominant models are presented.

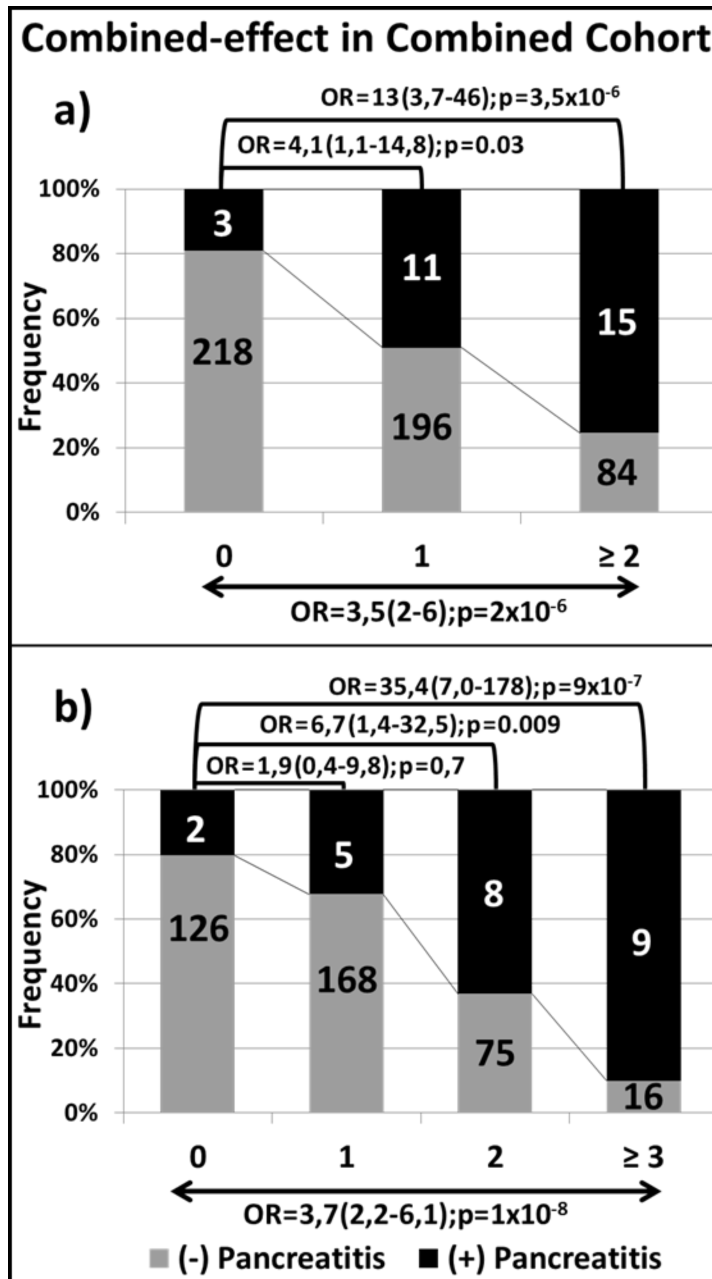
Supplemental Figure-S1



Supplemental figure S1. Comprehensive combined-effect model of all SNPs significantly associated with pancreatitis.

The combined effect of all five SNPs found to have significant associations with pancreatitis throughout the study (i.e. rs72755233 in *ADAMTS17*, rs3809849 in *MYBBP1A,* rs9908032 in *SPECC1,* rs11556218 in *IL16* and rs34708521 in *SPEF2*) was investigated in both **a)** the discovery cohort and **b)** the replication cohort. The groups of 0, 1, 2 and 3 or more variant alleles were compared. The association between the number of minor alleles and the increase in the risk of pancreatitis was directly proportional (OR and 95%CI at the bottom of the graph). Each bar represents the number of the variant alleles (i.e. none, one, two, three or more). The frequency of patients with and without pancreatitis is represented by the black and grey part of the bar. The number of samples per category is displayed inside of the bar. Carriers of 3 or more variant alleles were associated with a significant increase in the risk of pancreatitis (OR and 95% CI at the top of the graph).

Supplemental Figure-S2

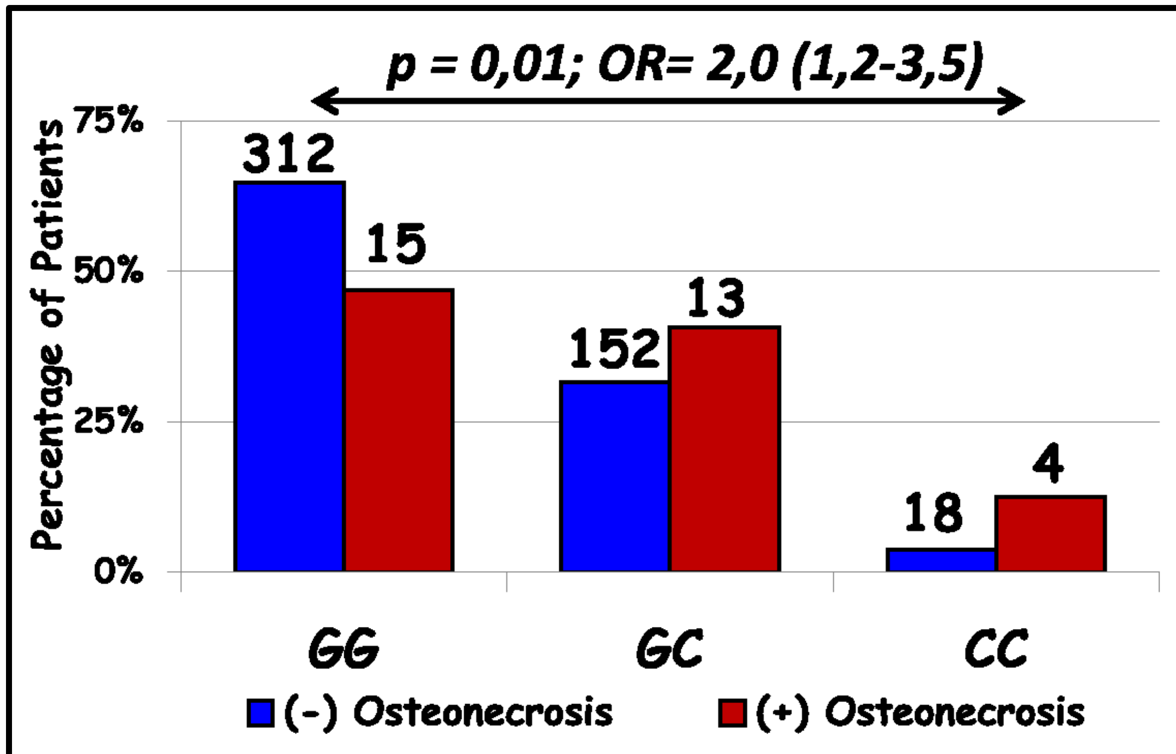


Supplemental figure S2. Combined-effect analysis in Combined-cohort for the 3 SNPs and 5 SNPs significantly associated with pancreatitis.

The combined effect of SNPs associated with pancreatitis was analysed in the combined cohort both for **a)** the model containing only the three SNPs significantly associated with pancreatitis in the EWAS (i.e. rs72755233 in *ADAMTS17*, rs3809849 in *MYBBP1A* and rs9908032 in *SPECC1*) and **b)** the comprehensive model additionally containing rs11556218 in *IL16* and rs34708521 in *SPEF2*. The patients were assigned to groups based on the number of risk alleles indicated at the bottom of each bar. The association between the number of minor alleles and the increase in the risk of pancreatitis was directly proportional (OR and 95%CI at the bottom of the graph). The frequency of patients with and without pancreatitis is represented by the black and grey part of the bar. The number of samples per category is displayed inside of the bar. The OR with 95% CI for each risk group (as compared to the group of patients not having any of the risk alleles) is displayed at the top of the graph.

4.15. Unpublished Data

4.15.1. rs3809849 in *MYBBP1A* gene and risk of osteonecrosis



Unpublished Data Figure U1. Association of the genotype of rs3809849 polymorphism in the *MYBBP1A* gene with the risk of osteonecrosis during ALL treatment.

This association analysis was performed in the combined DFCI cohort described earlier in chapter. The p value of the association in additive model, along with the odd-ratio and 95% confidence interval (in brackets) are indicated on the top of the graph. The frequency of patients with and without osteonecrosis is represented by red and blue bars, respectively. The number of patients is shown on the top of each bar and the genotypes are indicated at the bottom of the graph.

Unpublished Data Table U1. Stratification analysis of the association of rs3809849 polymorphism in the MYBBP1A gene with the risk of osteonecrosis based in clinical subgroups in the combined-DFCI cohort.

Stratification	Subgroup	ON	MYBBP1A_rs3809849: G > C*			*p* (Pearson's)	*p* (Trend)
			GG	GC	CC		
DFCI Protocol	(2000 + 1995)	(+)	279 (65,6%)	129 (30,4%)	17 (4%)	0,04	0,01
		(-)	11 (64,4%)	11 (31,1%)	3 (4,4%)		
	(1987 + 1991)	(+)	33 (57,9%)	23 (40,4%)	1 (1,8%_	0,2	0,5
		(-)	4 (57,1%)	2 (28,6%)	1 (14,3%)		
Sex	Female	(+)	134 (60,9%)	76 (34,5%)	10 (4,5%)	0,005	0,01
		(-)	7 (41,2%)	6 (35,3%)	4 (23,5%)		
	Male	(+)	178 (67,9%)	76 (29%)	8 (3,1%)	0,3	0,4
		(-)	8 (53,3%)	7 (46,7%)	0 (0%)		
Age	< 10 Years	(+)	260 (65,3%)	122 (30,7%)	16 (4%)	0,002	0,002
		(-)	8 (40%)	8 (40%)	4 (20%)		
	≥ 10 Years	(+)	52 (61,9%)	30 (35,7%)	2 (2,4%)	0,8	0,9
		(-)	7 (58,3%)	5 (41,7%)	0 (0%)		
Risk	Standard	(+)	171 (62,9%)	91 (33,5%)	10 (3,7%)	0,008	0,05
		(-)	7 (50%)	4 (28,6%)	3 (21,4%)		
	High	(+)	141 (67,1%)	61 (29%)	8 (3,8%)	0,1	0,08
		(-)	8 (44,4%)	9 (50%)	1 (5,6%)		
Source of ASNase	E. Coli	(+)	274 (63,1%)	143 (32,9%)	17 (3,9%)	0,04	0,03
		(-)	15 (48,4%)	12 (38,7%)	4 (12,9%)		
	Erwinia	(+)	38 (79,2%)	9 (18,8%)	1 (2,1%)	0,1	0,1
		(-)	0 (0%)	1 (100%)	0 (0%)		
Presence of Allergic Reactions	(-) Allergies	(+)	263 (66,9%)	116 (29,5%)	14 (3,6%)	0,04	0,02
		(-)	12 (48%)	10 (40%)	3 (12%)		
	(+) Allergies	(+)	49 (55,1%)	36 (40,4%)	4 (4,5%)	0,5	0,3
		(-)	3 (42,9%)	3 (42,9%)	1 (14,3%)		

ASNase, asparaginase; DFCI, Dana-Farber Cancer Institute ALL Consortium; ON, osteonecrosis.

* The analysis was performed assuming an additive genetic model.

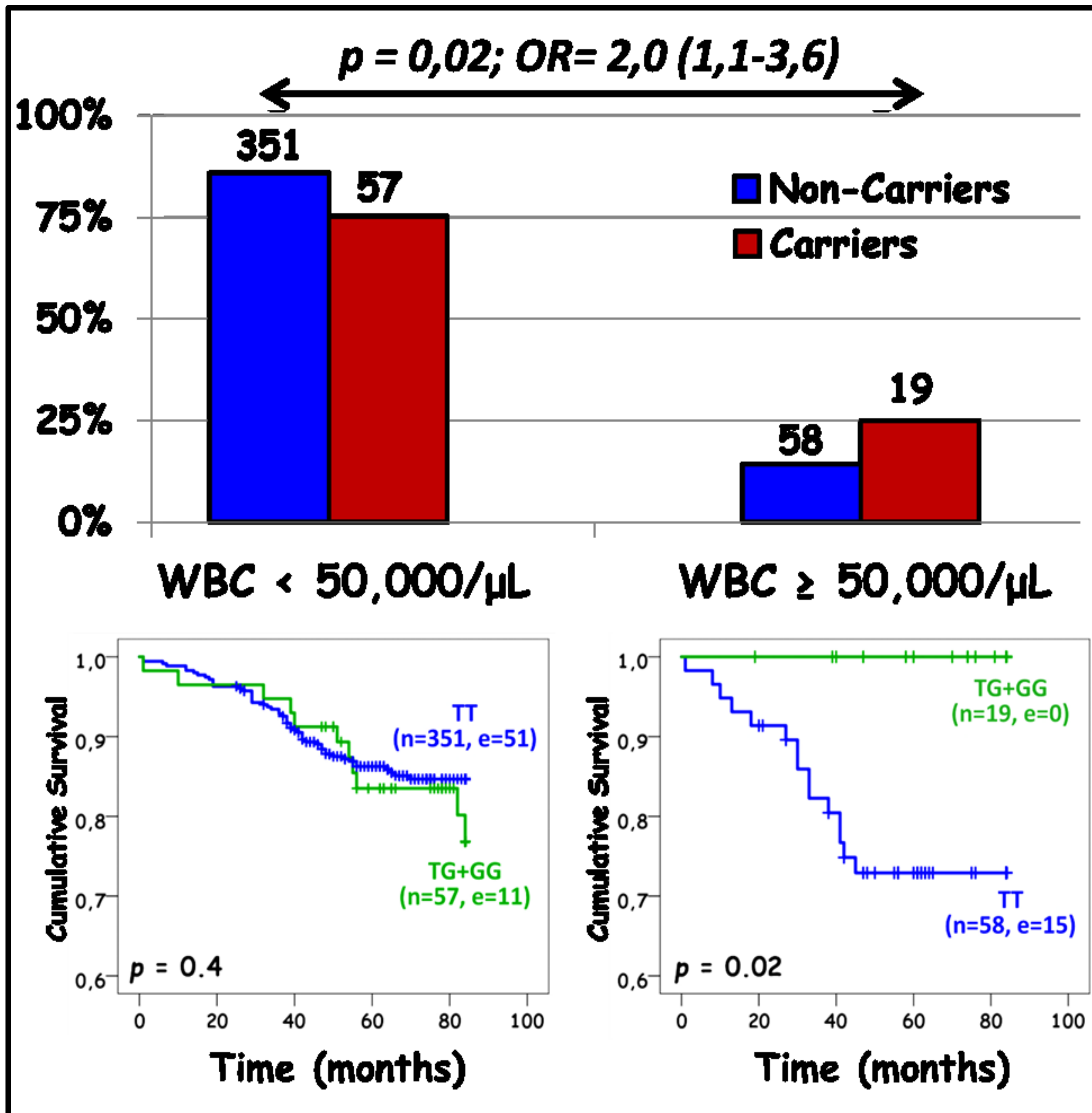
Unpublished Data Table U2. Multi-variant analysis for the risk of osteonecrosis associated adjusting for rs3809849_MYBBP1A polymorphism genotype along with clinical factors.

Risk Factor	p (Wald)	OR (95% CI)
MYBBP1A_rs3809849 Genotype	0,02	2,0 (1,1 - 3,5)
Age (<10 vs. ≥10 years)	0,02	3,6 (1,2 - 10,4)
Corticosteroid Used (dexamethasone vs. prednisone)	0,03	2,4 (1,1 - 5,4)
Treatment Protocol (new vs. old)	0,5	1,4(0,5 - 3,9)
Sex (female vs. male)	0,5	0,8 (0,4 - 1,6)
Source of Asparaginase (E.coli vs. Erwinia)	0,5	0,5 (0,1 - 4,2)
Risk (standard vs. high)	0,6	0,7 (0,3 - 2,1)
Presence of Allergic Reactions (yes vs. no)	0,8	1,1 (0,4 - 2,7)

Description of the results:

The rs3809849 *MYBBP1A* polymorphism showed a significant association with the risk of osteonecrosis in patients treated for childhood acute lymphoblastic leukemia. This association showed an additive effect in the entire cohort as the increase in the risk of osteonecrosis was proportional to the number of copies of the minor allele carried by the patient (Unpublished Data Figure U1). Moreover, the association was even stronger in certain subgroups stratified based on the clinical features of patients or the characteristics of the treatment protocol (Unpublished Data Table U1). Furthermore, the association of rs3809849 genotype with the risk of osteonecrosis maintained its significance when tested in a multi-variant model that incorporated all other factors that can possibly alter this outcome which can indicate that the observed association is independent from these factors (Unpublished Data Table U2).

4.15.2. rs11556218 in IL16 gene and event-free survival



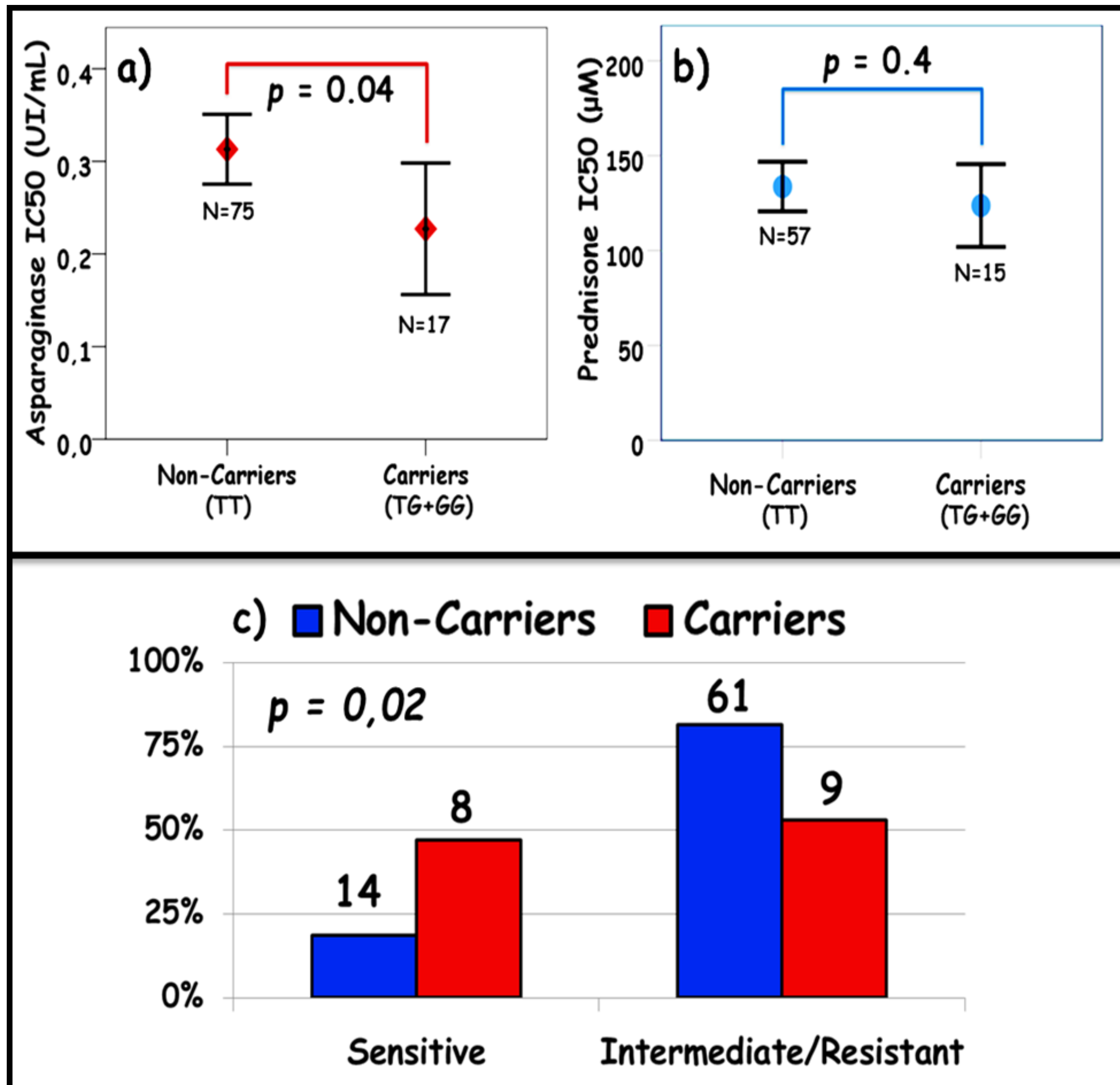
Unpublished Data Figure U2. Association of rs11556218 in IL16 gene with white blood cell count at presentation & event-free survival.

This association analysis was performed in the combined DFCI cohort described earlier in this chapter. The p value of the association comparing carriers (red) to non-carriers (blue), along with the odd-ratio and 95% confidence interval (in brackets) are indicated on the top of the graph. The frequency of patients with WBC count above and below the threshold of 50,000 cell/μL is represented by bars and the number of patients in each category is displayed on the top of each bar. The results of event-free survival analysis are presented under the respective groups. The p-values obtained by the log rank test for the difference between carriers (green curve) and non-carriers (blue curve) are provided on each plot. The number of patients represented by each group and the number of patients with event (in brackets) are indicated next to each curve.

Description of the results:

The carrier status of the variant allele of rs11556218 in *IL16* gene was significantly associated with the initial number of white blood cells at the time of the diagnosis. Carriers of this variant were at an increased risk (twice as high) of having a cell count that surpasses the threshold for high-risk stratification, and which is usually indicative of a poorer prognosis. Nonetheless, the event-free survival analysis indicated that the particular group of patients who present a high white blood cell count but who also harbour the variant allele tend to have significantly better outcome when it comes to even-free survival than patients who present an elevated cell count but not having the variant allele. There was no genotype-dependant difference in survival for the group of patients that had low white blood cell count at presentation.

4.15.3. rs11556218 in IL16 gene and drug sensitivity



Unpublished Data Figure U3. Cell viability assay in lymphoblastoid cell-lines in relation to rs11556218 *IL16* gene polymorphism.

Differential sensitivity of lymphoblastoid cell-lines (LCLs) to **a)** ASNase and **b)** Prednisone, relative to the carrier status of the variant allele G of rs11556218 in the IL16 gene as reflected by the change in the 48h minimum inhibitory concentration IC_{50}. The test was performed by screening an already existing in-house library of LCL derived from 91 individuals of Northern and Western Europe ancestry (CEU). The p value of the difference was estimated by non-parametric methods using Mann-Whitney test and is provided on the top of each graph. The number of samples per category is provided. **c)** Quartile distribution of IC_{50} values were compared by genotypes using chi-square. The p-value of the association is provided inside of the graph and the number of samples per category is displayed on the top of the bars. Each group represents a sensitivity state to ASNase derived from the IC_{50} results. The frequency of carriers of the variant allele is given in red and of non-carriers in blue.

Description of the results:

The results suggest that carrying the variant allele of rs11556218 polymorphism in the *IL16* gene renders cells more sensitive to treatment with ASNase as demonstrated by the significant reduction in IC_{50}, the minimum concentration needed to inhibit the growth of half of the initial population of cells following 48 hours of treatment. This effect was not present upon treatment with prednisone, which can support a drug specific role for this polymorphism in response to ASNase treatment. The quartile distribution of the carriers versus non-carriers shows an over-representation of carriers in the ASNase sensitive category as compared with the intermediate/resistant group.

Section B

Chapter 5

Characterization of the functional impact of MYBBP1A gene on asparaginase sensitivity and risk of pancreatitis following exome-wide association study results

This work represents a follow-up study aimed to use in vitro cell-based assays to investigate the functional impact of the *MYBBP1A* gene whose variant was shown to be associated with multiple complications of asparaginase. It reports the application of gene editing techniques to produce a *MYBBP1A* gene knock-out pancreatic cell line and the impact of this genetic modification on cellular proliferation, morphology and sensitivity to asparaginase.

This work is currently under preparation for publication. I take a lot of pride in this project as it taught me how to extract new research ideas from current data, employ them towards establishing novel hypotheses and design experiments that would validate them. Indeed, my contribution to this project involved the formulation of the hypotheses, the conceptualisation of the experiments, the coordination between the different key players, performing the experiments, analysing the results, and drafting the manuscript. This can be estimated as 80% of the work presented in this chapter.

Characterization of the functional impact of MYBBP1A gene on asparaginase sensitivity and risk of pancreatitis following exome-wide association study results

Running Title: MYBBP1A and pancreatitis: a functional follow-up

Rachid Abaji (1, 2), Vincent Roux (1), Ismahène Reguieg Yssaad (1,2), Paloma Kalegari (3, 4)

Vincent Gagné (1), Romain Gioia (1), Gerardo Ferbeyre (3, 4) Christian Beauséjour (1, 2)

and Maja Krajinovic (1, 2, 5) ✉

1- CHU Sainte-Justine Research Center, Montreal, QC, Canada;

2- Department of Pharmacology & Physiology, University of Montreal, Montreal, QC, Canada;

3- Department of Biochemistry and Molecular Medicine, University of Montreal, Montreal, QC, Canada;

4- University of Montreal Hospital Research Centre (CRCHUM), University of Montreal, Montreal, QC, Canada;

5- Department of Pediatrics, University of Montreal, Montreal, QC, Canada.

✉ Corresponding author.

Corresponding author:

Name: Maja Krajinovic

Address: Centre de Recherche, CHU Sainte-Justine, 3175 chemin de la Côte-Ste-Catherine, Montréal, QC, H3T 1C5 Canada

E-mail address: maja.krajinovic@umontreal.ca

Phone Number: 514 345-4931 #6259

5.1. Abstract

We previously identified a single-nucleotide polymorphism in the *MYBBP1A* gene that was associated in the exome-wide association study with the risk of asparaginase (ASNase)-induced acute pancreatitis in children with acute lymphoblastic leukemia (ALL)

Here we aimed to understand how the deletion of this gene would translate into differences in treatment response through cell-based functional analysis. We produced knockouts of the *MYBBP1A* gene in the PANC1 pancreatic cancer cell-line using CRISPR-CAS9 technology and tested the changes in cell proliferation capacity, sensitivity to drug treatment, colony formation potential and cellular morphology.

MYBBP1A knockout cells had a longer doubling time compared to the controls and their proliferation capacity was significantly lower ($p < 0.05$ at day 4 and $p < 0.01$ at day 5). Moreover, the deletion of this gene was associated with more sensitivity to ASNase, reflected by a significant 30% reduction in the inhibitory concentration 48h after drug challenge (IC_{50} = 0.30 IU/ml; 95% CI, 0.24-0.38 vs. IC_{50} = 0.42 IU/ml; 95% CI, 0.35-0.48, respectively; p=0.01); an effect that was not present upon treatment with vincristine (p=0.7),. Furthermore, the knockout cells exhibited a significant reduction in colony formation as compared to wild type cells prior (16% reduction; N = 159 ± 13.6 SEM, vs. N = 239 ± 14.2 SEM, respectively; p=0.001); and after ASNase treatment (13% reduction; p=0.02).

Cell cycles analysis demonstrated that gene deletion resulted in a specific blockage at the S-phase (47.5% vs 35.7%), along with an induction of apoptosis (23% vs 6.8%) in edited vs wild-type PANC1 cells, respectively. ASNase exposure blocked the cells in the G1/S checkpoint 48 hours post-incubation and further induced apoptosis and provoked cell necrosis at 96 hours of exposure. Also, the deletion of the gene from PANC1 cells was associated with a change in the morphology of the cells which seems to reflect a more malignant, mesenchymal phenotype concomitant with a 3.5-fold increase in Vimentin expression.

The results of this functional follow-up study further support a functional role of the *MYBBP1A* gene in modulating the risk of acute pancreatitis associated with the administration of ASNase during childhood ALL treatment.

Keywords:

MYBBP1A; asparaginase; pancreatitis; acute lymphoblastic leukemia; PANC1; CRISPR-CAS9; knockout.

5.2. Introduction

We previously reported the results of an exome-wide association study which identified a list of single-nucleotide polymorphisms (SNPs) associated with adverse drug reactions related to the administration of asparaginase (ASNase) during acute lymphoblastic leukemia treatment (ALL). Of those, rs3809849 in the *MYBBP1A* gene was associated with multiple major complications of ASNase treatment including pancreatitis, allergy, thrombosis, event-free survival (EFS) and overall survival (OS); and the association with pancreatitis was replicated in the independent validation cohort of that study.

Acute pancreatitis is a common dose-limiting toxicity that can occur in up to 18% of ALL patients.[1-4] The extent of pancreatitis severity varies from mild, self-resolving symptoms, to a severe systemic inflammatory response syndrome and failure of pancreatic function that can eventually precipitate an acute or persistent diabetes mellitus.[3-5]

MYBBP1A codes for Myb-binding protein 1A, a nucleolar protein implicated in stress response and carcinogenesis.[6] It was first recognized for its ability to interact with the leucine zipper of *c-myb* proto-oncogene product and to suppress its transactivation activity.[7, 8] Several studies have later described a functional role for this protein in essential biological functions such as cell division, cell proliferation, apoptosis and synthesis of ribosomal DNA.[6, 9-17] *MYBBP1A* has key roles in mitosis and tumor suppressor activity as its down-regulation influences several genes involved in regulating chromosomal segregation and cell cycle.[9] It was also identified as a substrate of Aurora-B kinase, thus further linking it to mitosis.[15] It physically interacts with several nuclear transcription

factors, such as the PGC-1α, AhR, NFκB, and p53.[10, 12, 14, 18, 19] *MYBBP1A* was also identified as a negative regulator of ribosomal RNA expression and as an integral part of the epigenetic mechanisms controlling ribosomal DNA.[20, 21]

In this study, we aimed to investigating the functional impact of knocking out this gene from PANC1 pancreatic cells and how this affects cellular behaviour and the response of pancreatic cells to ASNase treatment.

5.3. Materials and Methods

5.3.1. PANC1 cell line acquisition and maintenance

PANC1 cell line was courtesy of Dr. Gerardo Ferbeyre's Lab and cells were maintained using DMEM-based growth medium: DMEM (Wisent Inc.) + 10% fetal bovine serum (FBS) (Sigma - Aldrich) + 1% Primocin and incubated in 5% CO2 at 37°C conditions. The medium was changed every 3 to 4 days and the cells were passed when at 90% confluency.

5.3.2. MYBBP1A Knock-out PANC1 cell line production

5.3.2.1. Cas9 cloning into the PANC1 cell genome

Viral production of LentiCas9-Blast, a plasmid with lentiviral backbone that expresses human codon-optimized S. *pyogenes* Cas9 protein along with the blasticidin resistance gene, was performed as described in the virus production section of the Supplemental Methods. This virus was then used to infect PANC1 cells in order to induce constitutive expression of Cas9 protein in PANC1 cells (PANC1-Cas9). Briefly, 50 µL of the LentiCas9-Blast virus-containing supernatant was added to $5x10^5$ cells in a 6-wells plate and a volume of 1 mL of culture medium with the presence of polybrene (8mg/mL). The cells were incubated overnight at 37°C and the following day the culture medium was changed to (DMEM + 10% FBS + primocin). The transfected cells were allowed to grow for 24 hours before the antibiotic (blasticidin) was added to select for positive clones which were then confirmed by immunoblotting.

5.3.2.2. Knocking-out of MYBBP1A

A guide RNA targeting DNA sequence within the first exon of the *MYBBP1A* gene was designed and cloned into the backbone of a pLentiGuide vector as described in the MYBBP1A single-guide RNA expression vector construction section of the Supplemental Methods. pLentiGuide is a plasmid with lentiviral backbone that expresses S. *pyogenes* CRISPR chimeric RNA element along with the puromycin resistance gene. To induce the

production of MYBBP1A knock-out cells, 5x10^5 Cas9-expressing cells were plated in 1 mL of culture medium (DMEM + 10% FBS + primocin) in a 6-well plate. Next, 50 µL of the supernatant containing the gRNA-expressing vector (produced as mentioned above) was added along with 1 µl of polybrene (8mg/mL). The cells were incubated overnight at 37°C and the medium was changed to regular culture medium the following day. After 24 hours, the regular culture medium was replaced with the selection medium containing the antibiotic mix (blasticidin + puromycin) in order to select positive clones that both express Cas9 and the CRISPR RNA element. A mismatch assay was performed using GeneArt cleavage detection kit according to the manufacturer's protocol (Thermo Fisher scientific). MYBBP1A Knockout PANC1-Cas9 cells (PANC1-Cas9-ΔMYBBP1A) were then sorted by clonal selection whereby cells were separated into single cells using flow-cytometry techniques and each cell was individually transferred into a separate well of a 96-wells plate and left to establish a single-cell based colony. A Western Blot with anti-MYBBP1A antibodies was then performed to characterize positive clones.

5.3.3. Proliferation capacity assay

Measurement of the proliferation capacity was performed by plating in a 96-well plate (1×10^3 cells per well) and observing their relative growth over 6 consecutive days. On the day of plating, as well as on days 1, 2, 3, 4 and 5 post-plating, 10 µL of WST-1 cell viability reagent (Roche Diagnostics) was added to each well for a total reaction volume of 100 µL. After 2 hours of incubation, the absorbance was measured at a wavelength of 435nm using ClarioStar microplate reader (BMG LABTECH), which directly correlates to the

number of viable cells. The background absorbance was determined by adding the WST-1 reagent to wells that contained the culture medium without cells and the value obtained was used to adjust for the background noise of the other wells by subtraction. Each condition was carried in quadruplicate and repeated at least 3 independent times. The normalised data were then fitted into a non-linear regression curve using the Exponential Growth Equation allowing to evaluate the difference between the growth curves and to calculate the doubling-time. The mean absorbance, along with the standard error, was calculated at each time-point and the difference between each two cell lines (i.e. PANC1-WT vs. PANC1-Cas9-WT, PANC1-Cas9-WT vs. PANC1-Cas9-ΔMYBBP1A, and PANC1-WT vs. PANC1-Cas9-ΔMYBBP1A) was evaluated using Student's t-test. The proliferation curves were produced using GraphPad (Prism version 5.0).

5.3.4. IC_{50} determination assay

PANC1 cells in-vitro sensitivity to E.*coli* asparaginase (ASNase) and vincristine (VCR) relative to *MYBBP1A* gene deletion was assessed by calculating the drug concentration resulting in 50% inhibition of cell growth (IC_{50}). Briefly, for each of the cell lines (PANC1-Cas9-WT and PANC1-Cas9-ΔMYBBP1A), $1x10^3$ cells were plated in separate 96-wells plate and treated with either ASNase (10 concentrations ranging from 0.0001 to 100 I.U/mL) or VCR (7 concentrations ranging from 0.001uM to 1000uM). In order establish the baseline proliferation capacity of the cell line, a positive control containing the cells and the culture medium without any drug was used to serve as the drug-free, 100% viability reference. A

negative control containing the culture medium and the drug, but without any cells, was used to determine the cell-free background absorption. Following 48 hours of incubation, 10 µL of WST-1 cell viability reagent (Roche Diagnostics) was added to each well for a total reaction volume of 100 µL, and after 2 hours, the absorbance was measured at a wavelength of 435nm using ClarioStar microplate reader (BMG LABTECH). The value corresponding to the reading of the negative control was subtracted from other readings to compensate for the background absorption. The viability of cells at each drug concentration was calculated as a percentage of the viability relative to drug-free positive control. The experiment was performed in duplicates for each condition and repeated 6 separate times. IC_{50} values were estimated individually for each of the repetitions using GraphPad (Prism version 5.0) software by fitting sigmoidal dose-response curves for the two drugs. Obtained values were correlated to the presence or absence of the *MYBBP1A* gene (PANC1-Cas9-WT and PANC1-Cas9-ΔMYBBP1A, respectively) using Student's t-test.

5.3.5. Colony formation assay

The capacity of cells to produce colonies was evaluated in the presence and the absence of ASNase. For assay without the drug, 500 cells were plated in a 6-well plate and covered with culture medium (DMEM + 10% FBS + Primocin) and incubated at 37°C in a humidified incubator for two weeks. For assay with the drug, 1000 cells were incubated for 48 hours in a culture medium containing ASNase (0.45 IU/ml; corresponding to the IC_{50} value determined in the previous experiment). Following the drug challenge, the medium was aspirated and replaced with ASNase-free culture medium and the cells were incubated

for 12 more days (a total of 14 days from initial plating). At day 14, colony detection was performed by staining cells with *Methylene Blue* solution for at least 30 minutes. After staining, the plates were washed and air dried and colonies were counted manually against a bright background. The experiment was performed at least 4 times for each condition and the difference in the number of colonies at day 14 between the two cell lines (with and without ASNase) was evaluated using the Student's t-test.

5.3.6. Flow Cytometry analysis

The impact of the genetic modification and ASNase exposure on PANC1 cell-cycle and apoptosis/necrosis was evaluated using a double staining flow-cytometry approach. PANC1-Cas9-WT and PANC1-Cas9-ΔMYBBP1A cells were cultured simultaneously for 48 hours with or without ASNase and harvested in drug-free culture medium just before the assay. Cells were analyzed with FACS CANTO (BD Immunocytometry) and the results were processed using BD FACSDiva™ or Flowjo™ Software. Experiments were performed on at least three independent cultures.

Click-iT™ EdU Alexa Fluor™ 647 Flow Cytometry Assay Kit (Invitrogen) was used following the manufacturer's protocol which employs 5′ethynyl-2′-deoxyuridine (EdU) and 7-amino-actinomycin D (7-AAD) staining. Briefly, cells were harvested and pulsed with 10 μM EdU for 1.5 hour and Anti-EdU-Antibody conjugated to Alexa Fluor® 647 was used to stain EdU positive cells while 7-AAD was used to stain DNA. The percent of cells in each phase was calculating following the display of the results as bivariate distribution of EdU content versus DNA content. The percent of cells in the S-phase was calculated by gating

EdU positive cells, while the percentage in the G0/G1-phase was calculated based on EdU negative cells with low 7-AAD signal and that of the G2/M-phase based on EdU negative cells with high 7-AAD signal.

For the apoptosis analysis, cells were harvested, washed and the dead and apoptotic cells were detected by Annexin V/ Propidium Iodide Solution (PI) staining using APC Annexin V Apoptosis Detection Kit with PI (BioLegend) following the manufacturer's instructions. Results were displayed as bivariate distribution of Annexin V staining versus PI staining. The percentage of early apoptotic cells was calculated by gating Annexin V positive and PI negative cells while that of cells in the late apoptotic phase was based on Annexin V positive and PI positive cells. Necrotic cells' percentage was represented by the Annexin V negative and PI negative cell population.

5.3.7. Epithelial-Mesenchymal Transition characterization

Total RNA extracts were prepared in TransZol (Civic Bioscience). Total RNA was reverse transcribed using 5X All-In-One RT MasterMix (Abmgood) on 2 μg of total RNA in 20 μL final volume according to the kit's instructions. Before proceeding to qPCR, reverse transcription products were diluted 10-fold in RNAse free water. Real-Time PCR (Roche Applied Science) was performed using SYBR Green technologies as described previously.[22] The ΔΔCT method in a light Cycler 480 (Roche) was used to relative target-gene quantification. The mRNA expressions were measured relative to the mRNAs of two housekeeping genes (HMBS and TBP). Sequences of qPCR primers can be made available upon request.

5.4. Results

5.4.1. MYBBP1A Knock-out PANC1 cell line production & characterization

The production of Cas9 protein expressing wildtype PANC1 cells (PANC1-Cas9-WT) was confirmed by Western blotting (Figure 1a) and these cells were then used for further experiments to produce the gene knock-out. The result of the mismatch assay indicated approximately 30% efficiency within the cell population used for clonal selection (Supplementary Figure S1). Five clones were probed for MYBBP1A protein expression by immunoblotting using Anti-MYBBP1A antibodies (Figure 1b). Four clones (PANC1-Cas9-ΔMYBBP1A clones 1, 2, 3 and 4) were selected for further analysis as they showed a significant reduction in MYBBP1A protein expression relative to the wild-type control (Supplementary Figure S2).

Of note, there was no significant difference between wild-type PANC1 cells expressing the Cas9 protein and those not expressing it in any of the experiments performed (data not shown). Therefore, all following results presented in this work show only the comparison between the PANC1 cells that express Cas9 without the gene deletion (PANC1-Cas9-WT) and with the deletion (PANC1-Cas9-ΔMYBBP1A); since the former ones were used to produce the latter ones.

5.4.2. Proliferation capacity

The effect of MYBBP1A knock-out on cellular proliferation capacity was evaluated over a period of 6 days and the results indicated that the two growth curves were significantly different from each other ($p < 0.0001$). Doubling time was longer for PANC1-Cas9-ΔMYBBP1A cells as compared to PANC1-Cas9-WT cells (38 Hours; 95% CI, 30.6-51.2 vs. 34 Hours; 95% CI, 29.2–40.8, respectively). The results demonstrate a visible divergence in growth curves over time starting at day 2 post plating and this difference became significant at day 4 (OD_{435nm} = 2.15 vs. 1.47; $p = 0.02$; for controls and knock-out cells, respectively; Figure 1c) and continued to increase overtime (OD_{435nm} = 3.28 vs 1.94 at day 5; $p = 0.001$; Figure 1c).

5.4.3. IC_{50} determination and in-vitro sensitivity

The result of the IC_{50} determination assay indicate that following 48 hours of drug challenge, the deletion of the *MYBBP1A* gene significantly correlates with the *in vitro* sensitivity of PANC1 cells to ASNase (p = 0.01). Knockout cells needed a lower concentration of ASNase to inhibit the growth of half of the original population than needed by the controls (IC_{50} = 0.30 IU/ml; 95% CI, 0.24-0.38 vs. IC_{50} = 0.42 IU/ml; 95% CI, 0.35-0.48, respectively; Figure 2b). Treatment with VCR did not shown any significant difference (p = 0.7) relative to the presence or absence of the gene (IC_{50} = 3.7 μM; 95% CI, 2.31-5.96 vs. IC_{50} = 4.1 μM; 95% CI, 2.56-6.58, respectively; Figure 2b). This selective increase in sensitivity to ASNase of PANC1-Cas9-ΔMYBBP1A was maintained even after 96 hours of incubation with the drug (IC_{50} = 0.16 IU/ml; 95% CI, 0.13-0.19 vs. IC_{50} = 0.23 IU/ml; 95% CI, 0.16-0.30, for the edited vs unedited cell line respectively; Supplemental Figure S3).

5.4.4. *Colony formation*

The results of the clonogenic assay demonstrate that when the MYBBP1A expression was disrupted, PANC1 cells formed visibly less colonies than the control (Figure 3a). This difference was statistically significant (16% reduction; p = 0.001), even without drug treatment (N = 159 ± 13.6 SEM, vs. N = 239 ± 14.2 SEM), for the knock-out and the wild-type, respectively (Figure 3b). Moreover, upon microscopic examination of these colonies, the knock-out cells exhibited a considerable change in cell morphology (Figure 3a). Furthermore, even upon challenge with ASNase for two days at a dose corresponding to the average IC_{50} value of the control determined above, the reduction in colony formation capacity was still significant (13% reduction; p = 0.02; N = 226 ± 1.5 SEM, vs. N = 77 ± 2.5 SEM for control and knock-out cells, respectively; Figure 3b). Interestingly, we could observe a strong synergistic effect when comparing the colony formation capacity of untreated wild-type control cells to that of the ∆MYBBP1A cells exposed to the treatment (47.8% vs. 15.7%, respectively; a reduction of 32%; $p < 0.0001$).

5.4.5. *Flow Cytometry analysis*

The cell-cycle analysis results suggest that PANC1-Cas9-ΔMYBBP1A cells undergo a specific block of the cell cycle at the S phase as the percentage of cells in this phase was 12% higher when compared to PANC1-Cas9-WT cells (47.5% vs 35.7%, respectively) along with a decrease in the number of cells in the G0/G1 phase (41.5% vs 49.9%) and G2/M phase (14.4% vs 11%), as demonstrated in Figure 4a. Interestingly, these differences between the two cell lines and their directions were maintained even after 48 hours of ASNase exposure

(15.3% vs 7.2% in the S phase, 78.6% vs 84.8% in the G0/G1 phase and 6.1% vs 8% in the G2/M phase) for PANC1-Cas9-ΔMYBBP1A and PANC1-Cas9-WT cells, respectively (Figure 4a). However, following the treatment, a significant shift in the ratios of cells across the three cell cycle phases can be observed in both cell lines with the most notable change being the significant increase of the number of cells in the G0/G1 phase at the expense of the two other phases, thus suggesting that the drug induces a cell cycle block at the G0-G1/S checkpoint, Figure 4a.

The flow cytometry results also suggest that knocking MYBBP1A out of PANC1 cells reduces total cell viability, since the proportion of healthy cells (non-apoptotic and non-necrotic) was 75.4% vs 91.5% of the total cell population in PANC1-Cas9-ΔMYBBP1A and PANC1-Cas9-WT cells, respectively. The Annixin V/7-AAD staining experiment revealed that this difference stems from a significant increase in the percentage of cells undergoing apoptosis, as 23% of PANC1-Cas9-ΔMYBBP1A cells vs 6.8% of PANC1-Cas9-WT cells were apoptotic and therefore positively stained with Annixin V. Of these, 19.6% and 3.9%, respectively, were early apoptotic cells (negatively stained for PI), while 3.4% and 2.9%, respectively, were late apoptotic cells (negatively stained for PI) as illustrated in Figure 4b. The percentage of cells undergoing necrosis did not change between the edited and wild-type cell lines, since 1.7% vs 1.6%, respectively, stained negative for Annixin V but positive for PI (Figure 4b). Similar results were obtained following 48 hours of exposure to ASNase and the distribution of cells was as follows: 76.7% vs 90.3% healthy; 18.5% vs 5.9% early apoptotic; 3.4% vs 2.6% late apoptotic; and 1.3% vs 1.1% necrotic, in PANC1-Cas9-ΔMYBBP1A and PANC1-Cas9-WT cells, respectively (Figure 4b). After 96 hours of incubation with ASNase, a significant reduction in the number of healthy cells in both cell

lines compared to untreated cells was observed (61% of PANC1-Cas9-ΔMYBBP1A and 68% of PANC1-Cas9-WT), while a further increase in the percentage of apoptotic (38.9% and 31.8%) and necrotic (7.1% vs 6.1%) cells was recorded for each cell line, respectively (Figure 4b). Furthermore, by comparing the percentage of unhealthy cells in each of the conditions to that of the untreated wild-type cells, a strong additive effect can be observed for combining gene deletion, ASNase exposure and longer duration of treatment (p = 0,0003).

5.4.6. Epithelial-Mesenchymal Transition

When performing the colony formation assay, an intriguing observation was that the colonies formed out of PANC1-Cas9-ΔMYBBP1A cells were visibly less dense and more diffused (as they displayed a much lighter blue colour after staining) compared to colonies of the PANC1-Cas9-WT control cell line. In fact, the lack of the *MYBBP1A* gene expression seems to have provoked a distinctive change in cellular morphology similar to that seen during an Epithelial-Mesenchymal Transition (EMT). Accordingly, we investigated the levels of N-Cadherin, Vimentin and ZEB, markers associated with the cellular transition to a mesenchymal state. The results suggest a general increase in the relative mRNA expression levels for all three markers (Figure 5a), with Vimentin showing a significant increase of 3.5-fold in PANC1-Cas9-ΔMYBBP1A compared to PANC1-Cas9-WT cells (p = 0.05). Interestingly, following 48 hours of incubation with ASNase, the levels of Vimentin increased significantly in PANC1-Cas9-WT (7.7-fold; p = 0.007) but not in PANC1-Cas9-ΔMYBBP1A (1.4-fold; p = 0.6), as shown in Figure 5b.

5.5. Discussion

The *MYBBP1A* gene is a relatively newly discovered protein coding gene but it is involved in several essential biological functions as demonstrated in many studies.[6, 9-17] Its functional role spans cell proliferation, cell division, apoptosis and synthesis of ribosomal DNA, among others.[6, 9-17] Moreover, it is now well understood that the nucleolus functions as a stress sensor that can detect changes in ribosomal RNA content and consequently controls MYBBP1A translocation as a mechanism to modulate stress response.[12] Recently, the results of an exome-wide association study suggested the involvement of rs3809849 polymorphism in the *MYBBP1A* gene with the risk of multiple toxicities related to the administration of ASNase as part of ALL treatment protocols and even affecting the treatment outcome. Particularly interesting was the association with acute pancreatitis which was replicated in an independent validation cohort.[23] In this follow-up analysis, our goal was to investigate the impact of *MYBBP1A* gene deletion on the behaviour of pancreatic cells using PANC1 cell-line as a model.

The MYBBP1A knockout cells demonstrated what can be interpreted as 30% increase in sensitivity to ASNase, as reflected by a significantly lower IC_{50} after 48 hours (and 96 hours) of incubation with this drug. Importantly, this difference in sensitivity was unique to ASNase since the deletion of the *MYBBP1A* gene did not affect the *in-vitro* sensitivity to VCR, which exerts its function through distinct, ASNase-independent mechanisms.

Moreover, these cells had longer doubling time as compared to the control cells and the proliferation curves of the two cell lines were significantly different at day 4 of plating and beyond. This is in line with previous studies showing a decreased cellular growth in other cell lines upon MYBBP1A gene silencing, such as in head and neck squamous cell carcinoma [10] and HeLa cells.[9, 24] Also, the colony formation assay demonstrated a significant reduction in the number of colonies established by the knockout cells as compared to the controls, which further implies that abolishing *MYBBP1A* expression in PANC1 cells reduces their capacity to reproduce, possibly rendering them less capable of compensating for damage caused by stress. This effect was further potentiated when cells were challenged with ASNase at a concentration corresponding to the IC_{50} of the control cells, also supporting the observation that the knockout cells are more sensitive to the effect of this drug compared with the controls. This reduction in the clonogenic potential upon supressing *MYBBP1A* expression is consistent with the results found in hepatocellular carcinoma cells;[25] however, it is in contrast with the findings of other studies that reported an increase in the clonogenic potential of other cell lines upon *MYBBP1A* down-regulation including NIH3T3 cells,[9] and breast cancer cells.[11] Such discrepancy can be explained by the observation that the role of *MYBBP1A* gene in cellular viability and proliferation seems to be context-dependent, as an opposing effect of its downregulation was reported in different cell lines.[9] The silencing of this gene in the mouse embryonic stem cells, mouse embryonic fibroblasts and human HeLa cells was associated with a rapid entry into senescence and reduced proliferation capacity, while its down-regulation in immortalized NIH3T3 primary mouse embryonic fibroblast cells increased their growth rate and caused more potent Ras-driven tumors.[9] Indeed, the level of expression of

MYBBP1A was variably correlated with patient survival probability in different cancers. Low levels of expression of this gene are associated with lower survival in pancreatic cancer, while in renal cancer, melanoma and thyroid cancer, it was the higher levels of expression that showed a worsening prognosis (Supplemental Figure S4).[26]

The cell cycle analysis results show that knocking-out MYBBP1A in PANC1 cells results in a blockage at the S/G2-M checkpoint, suggesting a slower growth and a reduced capacity to complete the process of cytokinesis and enter into mitosis; which could explain the observed reduction in the PANC1-Cas9-ΔMYBBP1A cells' proliferation rate compared to PANC1-Cas9-WT cells even in the absence of ASNase treatment. Moreover, the results also show that ASNase exposure blocks the cells in the G0-G1/S checkpoint, plausibly due to the incapacity of cells to move forward with the protein synthesis process required for cytokinesis as a result of the depletion of the amino acid asparagine caused by the action of the drug. Interestingly, an additive effect can be observed for combining the genetic alteration and the treatment with ASNase, resulting in a 2-fold increase in the percentage of PANC1-Cas9-ΔMYBBP1A cells blocked in the S phase compared to PANC1-Cas9-WT, and an eventual decrease of the total number of cells reaching the G2/M phase undergoing mitosis (as indicated by the red-dotted line in Figure 4a). These results are in contrast with other studies that investigated the role of *MYBBP1A* gene in cell cycle control of other cell lines as one study in HeLa cells found that the suppression of gene expression resulted in a blockage at G2/M,[9] while another found the blockage at G1/S in hepatocellular carcinoma cells;[25] consistent with the growing body of evidence suggesting a cell-type dependent role for *MYBBP1A* gene.

The apoptosis experiment provides evidence on the involvement of MYBBP1A protein in PANC1 cells apoptosis, since knocking-out MYBBP1A in these cells resulted in a significant reduction of the healthy cells population, with a concomitant significant increase in the percentage of cells undergoing apoptosis (but not necrosis); consistent with the results found in another study on HeLa cells.[9] Interestingly, this difference was maintained between the two cell line following 48 hours of incubation with ASNase, but the percentage of apoptotic or necrotic cells did not increase. However, following 96 hours of ASNase exposure, there was a marked increase in the percentage of both apoptotic and necrotic cells in both cell lines. In fact, when comparing the increase in the percentage of unhealthy cells from each of the conditions to that of untreated wild-type cells, an additive effect can be observed for the combination of genetic alteration, treatment with ASNase and duration of exposure (Figure 4b).

By combining data from the cell cycle and apoptosis/necrosis analyses, it can be hypothesised that the effect of MYBBP1A gene deletion on PANC1 cells results from a specific cell cycle blockage at the S-phase, along with an induction of apoptosis; thus reducing the proliferation rate and the clonogenic capacity of the KO cells. Moreover, combined with ASNase exposure, this genetic modification provokes an additional cell cycle arrest at G0/G1 following two days of treatment, and a stronger induction of apoptotic reactions, as well as cellular necrosis, at day four of treatment; therefore providing a plausible mechanistic understanding of how MYBBP1A gene deletion modulates PANC1 cells sensitivity to ASNase treatment and its observed impact on their clonogenic potential.

Of note, *MYBBP1A* gene knockdown was previously shown to impact cellular morphology of HeLa cells, which displayed an abnormal, flattened, and enlarged morphology.[24] The visual examination of the cellular morphology of PANC1 cells following the knockout of *MYBBP1A* also revealed a significant change in their morphology as they became more spindle-shaped and distant from each other. Intriguingly, these features seem to reflect an epithelial-mesenchymal transition (EMT) process through which cells lose their polarity and eventually acquire a fibroblast-like phenotype. These cells also lose cell-cell adhesion showing more intercellular spacing and they also gain migratory and invasive properties.[27, 28] Upon measuring the relative mRNA expression levels of markers associated with the EMT process in MYBBP1A KO and WT cells, a general increase was observed in the levels of the tested markers, with Vimentin, a mesenchymal phenotypic marker, showing a significant 3-fold rise in expression. Of note, following 48 hours of incubation with ASNase, the levels of Vimentin surged significantly in PANC1-Cas9-WT cells, but not in PANC1-Cas9-ΔMYBBP1A cells; further supporting the assumption that they were already at a mesenchymal state. This EMT process, and a very similar phenotype to the one observed in this study were previously documented in PANC1 cells following exposure to transforming growth factor β (TGF-β),[29] or incubation in a hypoxic environment (which was shown to be mediated by NF-kB activation).[27] Interestingly, *MYBBP1A* is known to act as a transcriptional co-repressor of NF-kB and the activation of NF-kB pathway is linked to the development of acute pancreatitis [30] as well as to tumour progression and metastasis.[28]

Our observations regarding the impact of *MYBBP1A* knockout in PANC1 cells are in concordance with results reported in the literature showing that silencing of the expression of this gene had opposing effects on the cells as it was associated with a decreased cellular growth but an increase in cell migration capacity.[10] Altogether, they suggest that *MYBBP1A* gene plays the role of a gatekeeper that controls the balance between cellular proliferation and migration of pancreatic cells. Plausibly, the knockdown of this gene would disrupt the negative feedback loop that regulates the expression of NF-kB and consequently result in an exaggerated activation upon stress response. This could lead to more drug sensitivity and cellular necrosis of normal pancreatic cells, leading to pancreatitis upon treatment with ASNase, or it could induce tumour metastasis of the pancreatic cancer cells, possibly explaining the observed lower survival of patients with low expression of *MYBBP1A* (Supplemental Figure S4).[26]

Understanding how pharmacogenetics influence the response of pancreatic cells to ASNase at a molecular level does not only hold the potential for reducing the risk of ASNase-induced acute pancreatitis and enhancing treatment outcome, but it can also help refine treatment strategies for pancreatic cancers in which asparagine and/or glutamine depletion might be indicated.[31]

By providing new insights into the role of the *MYBBP1A* gene in regulating pancreatic PANC1 cell response to ASNase treatment, these results could further improve our understanding of the pathogenesis of ASNase-induced acute pancreatitis during childhood ALL treatment. However, it should be emphasized that even though this work demonstrates a functional implication of the *MYBBP1A* gene in PANC1 cells' response to

ASNase, pharmacogenetics studies of ASNase-response in other pancreatic cell lines as well as animal models of pancreatitis are still needed in order to improve our understanding of the tissue-specific role of this gene, since several studies have shown that the role of *MYBBP1A* varies across the different tissues. Moreover, in this context, it is worth mentioning that rs3809849, which was previously shown to be associated with multiple ASNase complications, seems to act as an eQTL variant of the *MYBBP1A* gene but that the extent of its effect varies significantly across the different tissues according to data from GTEx database.[32] Notably, the association was strongest in the aorta artery tissue as the expression of the gene was significantly reduced in relation with the number of copies of the minor allele. A similar trend could also be noted in the pancreas tissue, but the association lacked statistical significance. On the other hand, the association had an opposite direction in the Epstein-Barr virus (EBV)-transformed lymphocytes as the minor allele was associated with more gene expression; further consolidating the hypothesis of a context dependent effect of the *MYBBP1A* gene and its rs3809849 SNP.

Finally, further experiments are required to provide a mechanistic model that can explain the involvement of *MYBBP1A* in modulating the risk of ASNase-induced acute pancreatitis and to characterize the differences in the morphology and phenotypes associated with the gene knockout of PANC1 cells at a molecular level. Moreover, a knock-in experiment is needed to introduce the SNPs of interest into the knock-out cells in order to assess the functional impact of the risk allele of rs3809849 on ASNase treatment outcome and toxicity.

5.6. Authorship Contributions

M.K. & R.A. designed the study; R.A., V.R., I.B., P.K., V.G., and R.G. performed the experiments; R.A. performed the data analysis; C.B. supervised the gene editing process in the platform; M.K. provided the funding; R.A. drafted the article; All authors revised the manuscript.

5.7. Disclosure of Conflicts of Interest

The authors declare no competing financial interests.

5.8. Figure Legends

Figure 1. Production of Cas9 expressing, & MYBBP1A gene knock-out PANC1 cell lines and evaluation of cell proliferation capacity.

a) PANC1 cells were transduced with LentiCas9-Blast virus to constitutively express Cas9 protein. The success of viral transduction was assessed by a Western Blot against Cas9. **b)** Cas9-expressing PANC1 cells were used to produce *MYBBP1A* knock-out PANC1 cells by CRISPR-Cas9 genome editing technique. The efficiency of gene deletion was confirmed in several clones by the absence of a signal on Western Blot using anti-MYBBP1A antibodies. **c)** Proliferation of *MYBBP1A* gene knockout PANC1-Cas9 cells (PANC1-Cas9-ΔMYBBP1A, red triangles) and wild-type PANC1-Cas9 control (PANC1-Cas9-WT, green squares) was measured using the WST-1 cell proliferation assay. Absorbance was measured at 435 nm on days 0, 1, 2, 3, 4 and 5 post-plating. The quantitative data shown are the mean absorbance ± SEM from at least three separate experiments per condition. P values were evaluated by Student's t-test (* = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$).

Figure 2. In-vitro sensitivity to asparaginase and vincristine in relation to *MYBBP1A gene Knock-out*.

The distribution of IC_{50} values of **a)** asparaginase (ASNase) or **b)** vincristine (VCR) is plotted. IC_{50} values were calculated using WST-1 viability assay 48 hours post incubation with several concentrations of the drugs. The experiment was repeated 6 times and IC_{50} values were calculated for each experiment separately. The vertical lines represent the mean IC_{50} value of each group. The coloured shapes represent independent values calculated for WT PANC1 cells (green squares) or MYBBP1A knock-out PANC1 cells (red triangles). The p value obtained by the Unpaired Student's t-test represents the difference in drug sensitivity between the two cell lines and is provided on the top of the graph.

Figure 3. Effect of MYBBP1A gene knock-out from PANC1 cells on their colony formation capacity and response to treatment with asparaginase.

Cells were plated in a 6 wells-plate and the number of colonies formed (with and without asparaginase) was counted 14 days after plating. **a)** Representative example of a plate showing the difference in the number and density of colonies formed starting from the same quantity of PANC1 cells (with and without *MYBBP1A* gene deletion) along with magnification of one of the colonies from each plate to demonstrate the change in the morphology of the cells and inter-cellular spacing. **b)** The quantification of the colony forming unit (CFU) data showing the average number of colonies from all experiments along with the error bars. The difference between the two cell lines was evaluated using the Student's t-test and p values are provided.

Figure 4. Impact of MYBBP1A gene deletion on PANC1 cellular functions: a) Cell-Cycle and b) Apoptosis/Necrosis, and response to asparaginase exposure.

a) Percentage of cell in each of the 3 main cell-cycle phases. Cells were plated in a 6 wells-plate and incubated for 48 hours in culture medium without (-) or with (48H) asparaginase, respectively. Each bar represents 100% of cells in each condition indicated at the bottom of the graph (WT stands for PANC1-Cas9-WT cell line and KO for PANC1-Cas9-ΔMYBBP1A). Coloured sections of the bar represent the percentage of cells per respective phase as indicated on the top of the graph. The red-dotted line highlights the additive trend of reduction in percentage of cells entering the G2/M phase. **b)** Percentage of early apoptotic, late apoptotic and necrotic cells. Cells were plated in a 6 wells-plate and incubated in culture medium without asparaginase (-), or with asparaginase for 48 hours (48H) or 96

hours (96H). Each bar represents the fraction of non-healthy cells in each condition indicated at the bottom of the graph (WT stands for PANC1-Cas9-WT cell line and KO for PANC1-Cas9-ΔMYBBP1A). Coloured sections of the bar represent the percentage of cells per respective state as indicated on the top of the graph. The asterisks at the top of the bars indicate the significance level of the difference in non-healthy cell percentage as compared to the first condition (untreated WT cells). The p-values were calculated using Student's t-test (* = $p<0.05$, ** = $p<0.01$).

Figure 5. Relative expression of markers associated with epithelial-mesenchymal transition in PANC1 cells a) in response to MYBBP1A gene deletion and b) asparaginase exposure.

Coloured bards represent the relative mRNA expression levels of epithelial-mesenchymal transition markers in PANC1-Cas9-WT wild-type cells (WT) and PANC1-Cas9-ΔMYBBP1A cells (KO) following 48 hours of incubation. A significant difference in the marker's level is indicated by the presence of asterisks on the top of the bars (* = $p<0.05$, ** = $p<0.01$).

a) Expression of N-cadherin, Vimentin and ZEB without asparaginase exposure presented as fold-change in edited vs unedited cell line. p-values of the differences between the two cell lines were evaluated using Unpaired Student's t-test. **b)** Impact of 48 hours asparaginase treatment on expression levels of Vimentin for each of the cell lines. p-values of the difference within the same cell line in the presence and absence of asparaginase were evaluated using Paired Student's t-test.

5.9. Figures

Figure 1.

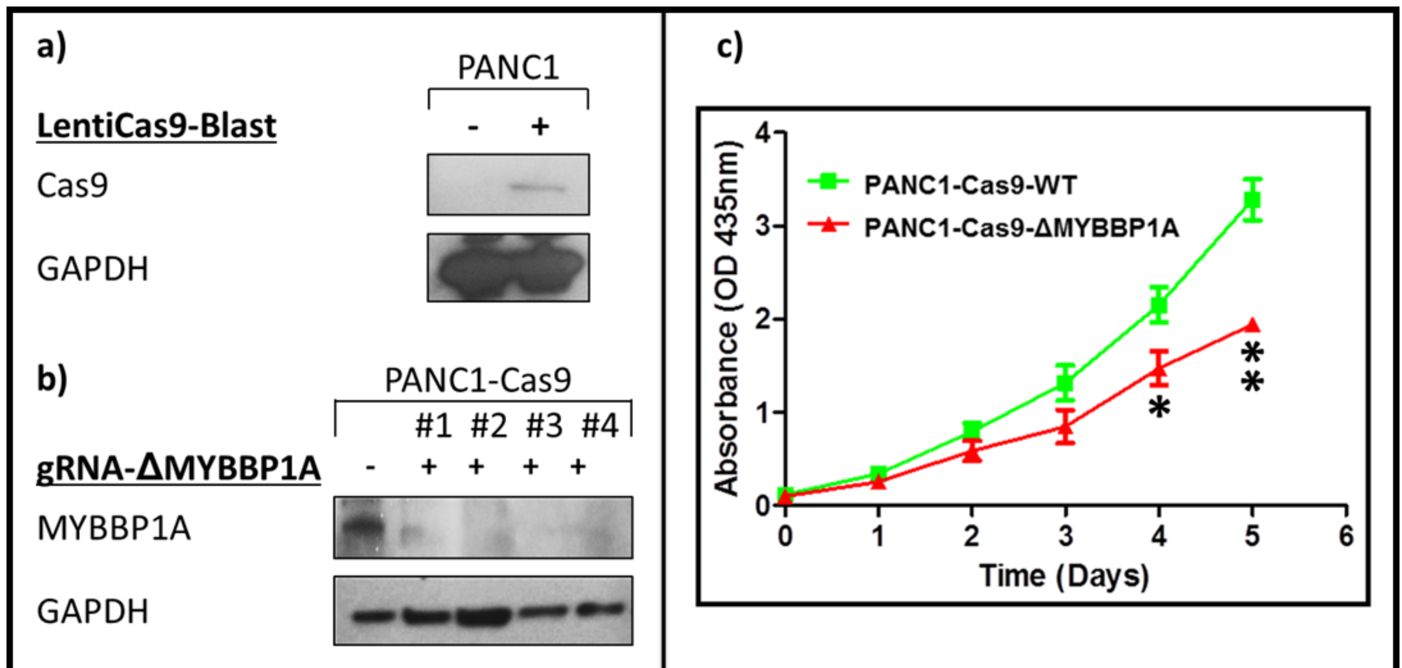


Figure 1. Production of Cas9 expressing, & MYBBP1A gene knock-out PANC1 cell lines and evaluation of cell proliferation capacity.

Figure 2.

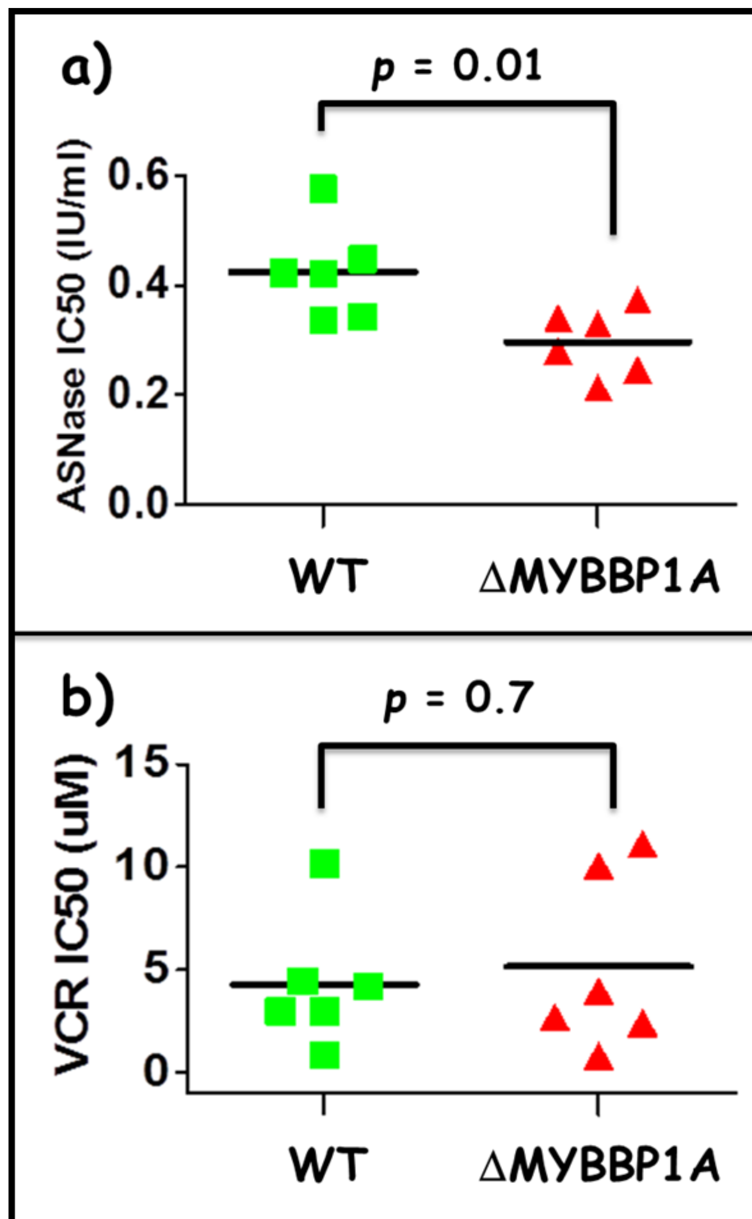


Figure 2. ***In-vitro*** **sensitivity to asparaginase and vincristine in relation to** ***MYBBP1A gene Knock-out*****.**

Figure 3.

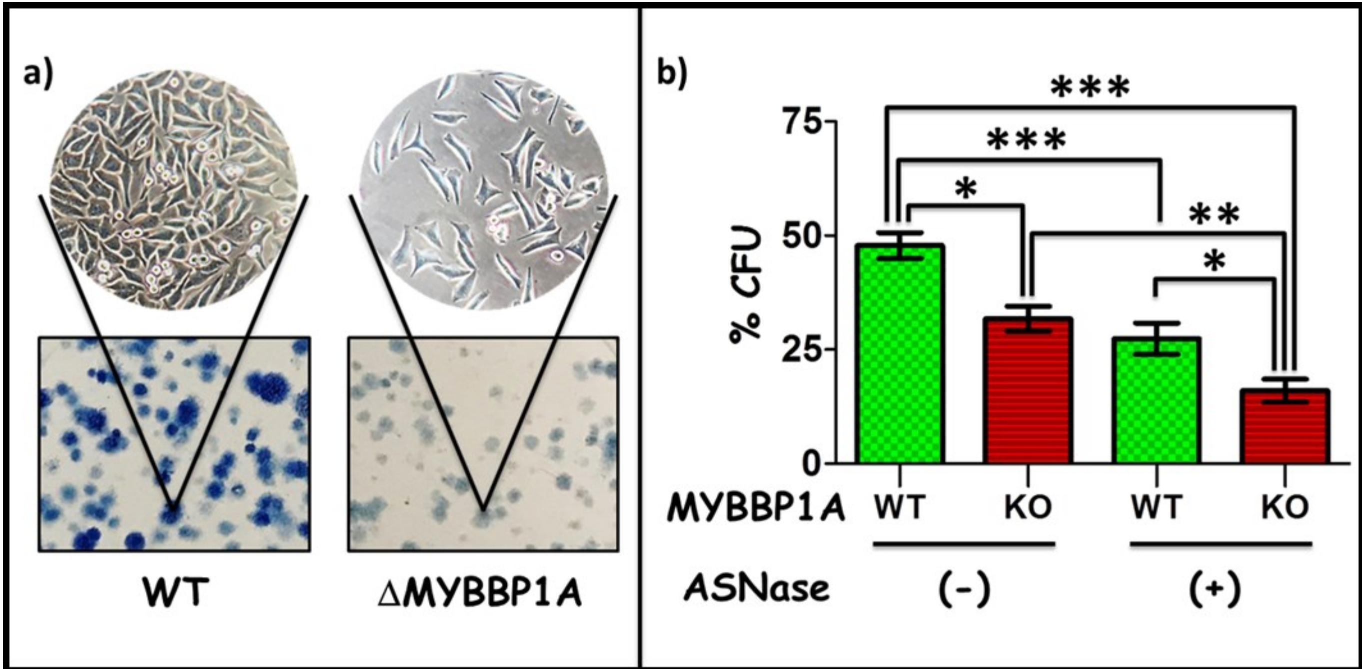


Figure 3. Effect of MYBBP1A gene knock-out from PANC1 cells on their colony formation capacity and response to treatment with asparaginase.

Figure 4.

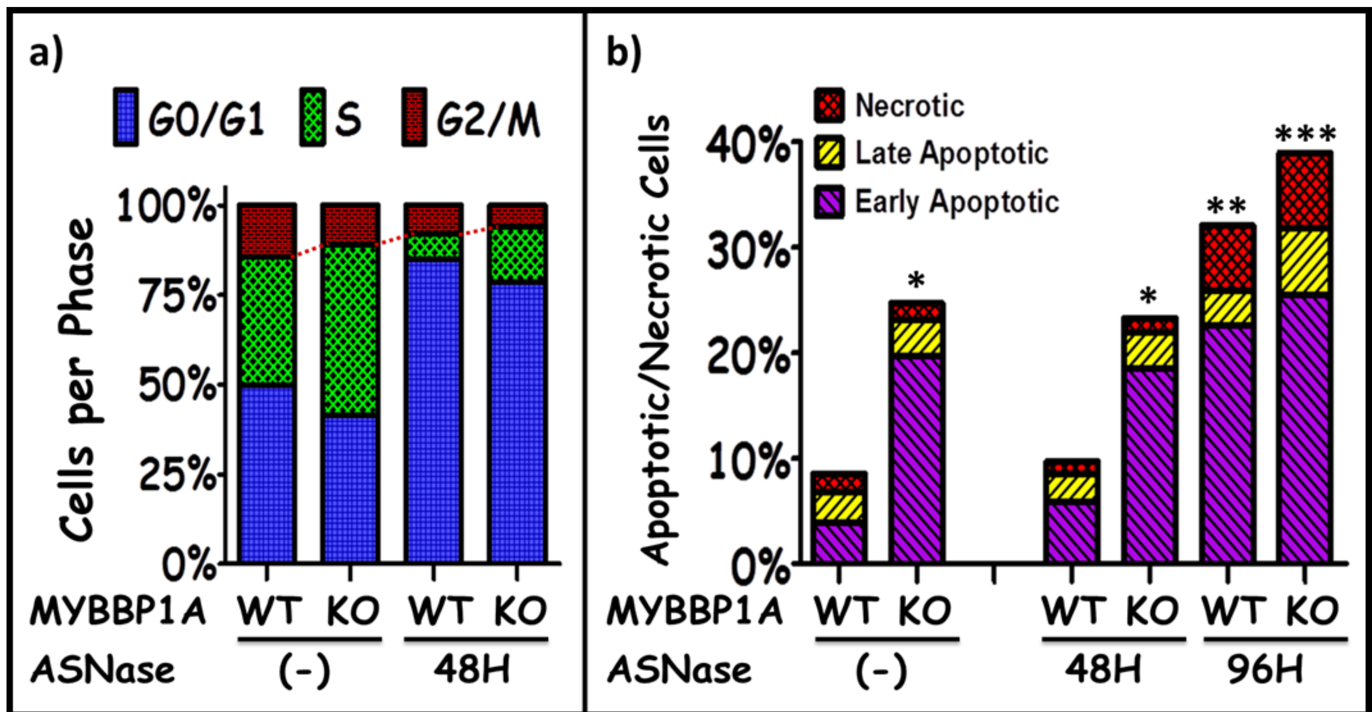


Figure 4. Impact of MYBBP1A gene deletion on PANC1 cellular functions: a) Cell-Cycle and b) Apoptosis/Necrosis, and response to asparaginase exposure.

Figure 5.

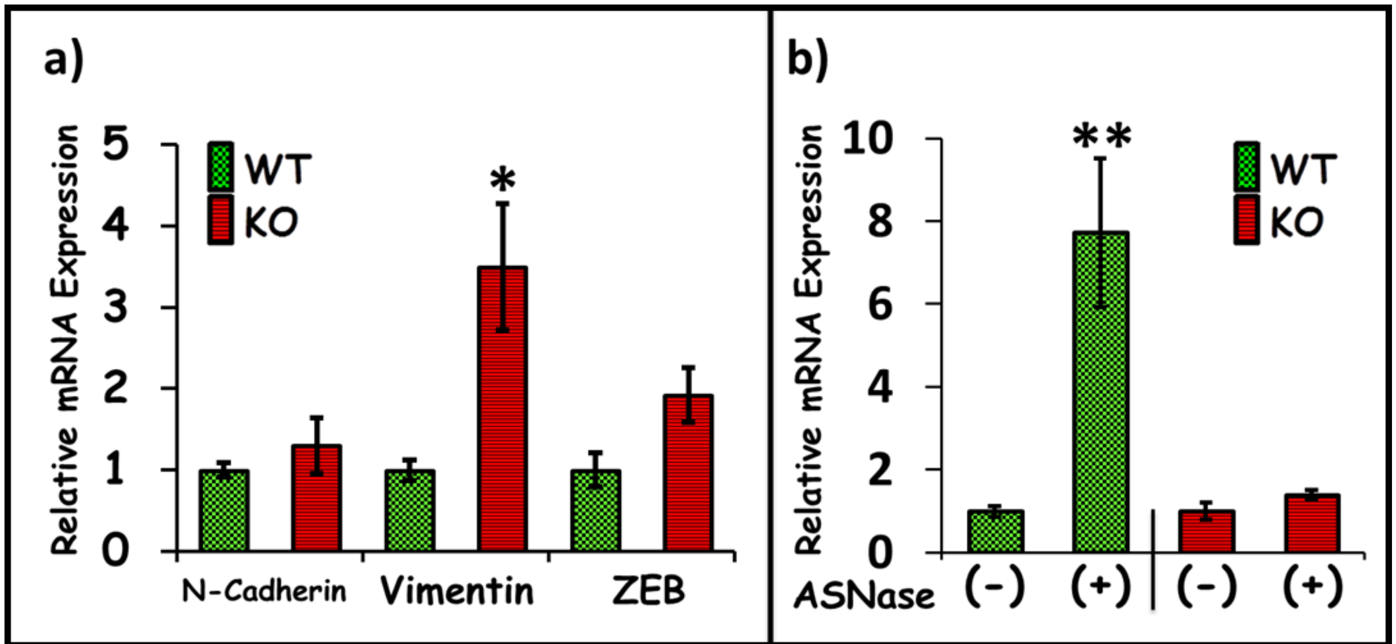


Figure 5. Relative expression of markers associated with epithelial-mesenchymal transition in PANC1 cells a) in response to MYBBP1A gene deletion and b) asparaginase exposure.

5.10. References

1. Ben Tanfous M, Sharif-Askari B, Ceppi F *et al*. Polymorphisms of asparaginase pathway and asparaginase-related complications in children with acute lymphoblastic leukemia. *Clin Cancer Res* 21(2), 329-334 (2015).

2. Rousseau J, Gagne V, Labuda M *et al*. ATF5 polymorphisms influence ATF function and response to treatment in children with childhood acute lymphoblastic leukemia. *Blood* 118(22), 5883-5890 (2011).

3. Schmiegelow K, Muller K, Mogensen SS *et al*. Non-infectious chemotherapy-associated acute toxicities during childhood acute lymphoblastic leukemia therapy. *F1000Res* 6 444 (2017).

4. Wolthers BO, Frandsen TL, Abrahamsson J *et al*. Asparaginase-associated pancreatitis: a study on phenotype and genotype in the NOPHO ALL2008 protocol. *Leukemia* 31(2), 325-332 (2017).

5. Kumar S, Ooi CY, Werlin S *et al*. Risk Factors Associated With Pediatric Acute Recurrent and Chronic Pancreatitis: Lessons From INSPPIRE. *JAMA Pediatr* 170(6), 562-569 (2016).

6. George B, Horn D, Bayo P *et al*. Regulation and function of Myb-binding protein 1A (MYBBP1A) in cellular senescence and pathogenesis of head and neck cancer. *Cancer Lett* 358(2), 191-199 (2015).

7. Favier D, Gonda TJ. Detection of proteins that bind to the leucine zipper motif of c-Myb. *Oncogene* 9(1), 305-311 (1994).

8. Tavner FJ, Simpson R, Tashiro S *et al*. Molecular cloning reveals that the p160 Myb-binding protein is a novel, predominantly nucleolar protein which may play a role in transactivation by Myb. *Mol Cell Biol* 18(2), 989-1002 (1998).

9. Mori S, Bernardi R, Laurent A *et al*. Myb-binding protein 1A (MYBBP1A) is essential for early embryonic development, controls cell cycle and mitosis, and acts as a tumor suppressor. *PLoS One* 7(10), e39723 (2012).

10. Acuna Sanhueza GA, Faller L, George B *et al*. Opposing function of MYBBP1A in proliferation and migration of head and neck squamous cell carcinoma cells. *BMC Cancer* 12 72 (2012).

11. Akaogi K, Ono W, Hayashi Y, Kishimoto H, Yanagisawa J. MYBBP1A suppresses breast cancer tumorigenesis by enhancing the p53 dependent anoikis. *BMC Cancer* 13 65 (2013).

12. Kumazawa T, Nishimura K, Katagiri N, Hashimoto S, Hayashi Y, Kimura K. Gradual reduction in rRNA transcription triggers p53 acetylation and apoptosis via MYBBP1A. *Sci Rep* 5 10854 (2015).

13. Ono W, Hayashi Y, Yokoyama W *et al*. The nucleolar protein Myb-binding protein 1A (MYBBP1A) enhances p53 tetramerization and acetylation in response to nucleolar disruption. *J Biol Chem* 289(8), 4928-4940 (2014).

14. Owen HR, Elser M, Cheung E, Gersbach M, Kraus WL, Hottiger MO. MYBBP1a is a novel repressor of NF-kappaB. *J Mol Biol* 366(3), 725-736 (2007).

15. Perrera C, Colombo R, Valsasina B *et al*. Identification of Myb-binding protein 1A (MYBBP1A) as a novel substrate for aurora B kinase. *J Biol Chem* 285(16), 11775-11785 (2010).

16. Tan BC, Yang CC, Hsieh CL *et al*. Epigeneitc silencing of ribosomal RNA genes by Mybbp1a. *J Biomed Sci* 19 57 (2012).

17. Tang Y, Zhao W, Chen Y, Zhao Y, Gu W. Acetylation is indispensable for p53 activation. *Cell* 133(4), 612-626 (2008).

18. Fan M, Rhee J, St-Pierre J *et al*. Suppression of mitochondrial respiration through recruitment of p160 myb binding protein to PGC-1alpha: modulation by p38 MAPK. *Genes Dev* 18(3), 278-289 (2004).

19. Jones LC, Okino ST, Gonda TJ, Whitlock JP, Jr. Myb-binding protein 1a augments AhR-dependent gene expression. *J Biol Chem* 277(25), 22515-22519 (2002).

20. Yang CC, Liu H, Chen SL *et al*. Epigenetic silencing of myogenic gene program by Myb-binding protein 1a suppresses myogenesis. *EMBO J* 31(7), 1739-1751 (2012).

21. Karim MF, Yoshizawa T, Sato Y *et al*. Inhibition of H3K18 deacetylation of Sirt7 by Myb-binding protein 1a (Mybbp1a). *Biochem Biophys Res Commun* 441(1), 157-163 (2013).

22. Vernier M, Bourdeau V, Gaumont-Leclerc MF *et al*. Regulation of E2Fs and senescence by PML nuclear bodies. *Genes Dev* 25(1), 41-50 (2011).

23. Abaji R, Gagne V, Xu CJ *et al*. Whole-exome sequencing identified genetic risk factors for asparaginase-related complications in childhood ALL patients. *Oncotarget* 8(27), 43752-43767 (2017).

24. Hochstatter J, Holzel M, Rohrmoser M *et al*. Myb-binding protein 1a (Mybbp1a) regulates levels and processing of pre-ribosomal RNA. *J Biol Chem* 287(29), 24365-24377 (2012).

25. Weng X, Wu J, Lv Z *et al*. Targeting Mybbp1a suppresses HCC progression via inhibiting IGF1/AKT pathway by CpG islands hypo-methylation dependent promotion of IGFBP5. *EBioMedicine* 44 225-236 (2019).

26. The-Human-Protein-Atlas. MYBBP1A-Pathology Atlas-Proggnostic Sammary. 2018(18-12),

27. Cheng ZX, Wang DW, Liu T *et al*. Effects of the HIF-1alpha and NF-kappaB loop on epithelial-mesenchymal transition and chemoresistance induced by hypoxia in pancreatic cancer cells. *Oncol Rep* 31(4), 1891-1898 (2014).

28. Julien S, Puig I, Caretti E *et al*. Activation of NF-kappaB by Akt upregulates Snail expression and induces epithelium mesenchyme transition. *Oncogene* 26(53), 7445-7456 (2007).

29. Wang H, Wu J, Zhang Y *et al*. Transforming growth factor beta-induced epithelial-mesenchymal transition increases cancer stem-like cells in the PANC-1 cell line. *Oncol Lett* 3(1), 229-233 (2012).

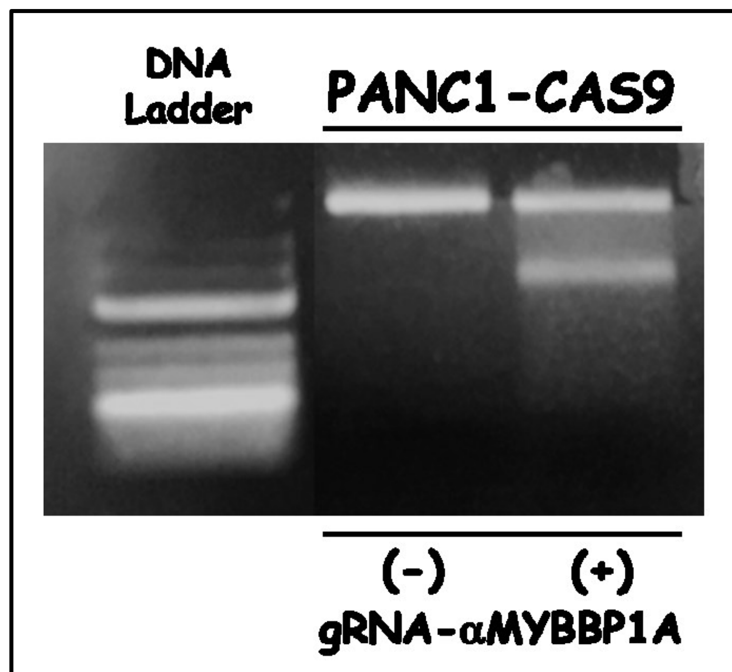
30. Rakonczay Z, Jr., Hegyi P, Takacs T, Mccarroll J, Saluja AK. The role of NF-kappaB activation in the pathogenesis of acute pancreatitis. *Gut* 57(2), 259-267 (2008).

31. Cui H, Darmanin S, Natsuisaka M *et al*. Enhanced expression of asparagine synthetase under glucose-deprived conditions protects pancreatic cancer cells from apoptosis induced by glucose deprivation and cisplatin. *Cancer Res* 67(7), 3345-3355 (2007).

32. Gtex-Portal. MYBBP1A-rs3809849. 2019(11-12), (2019).

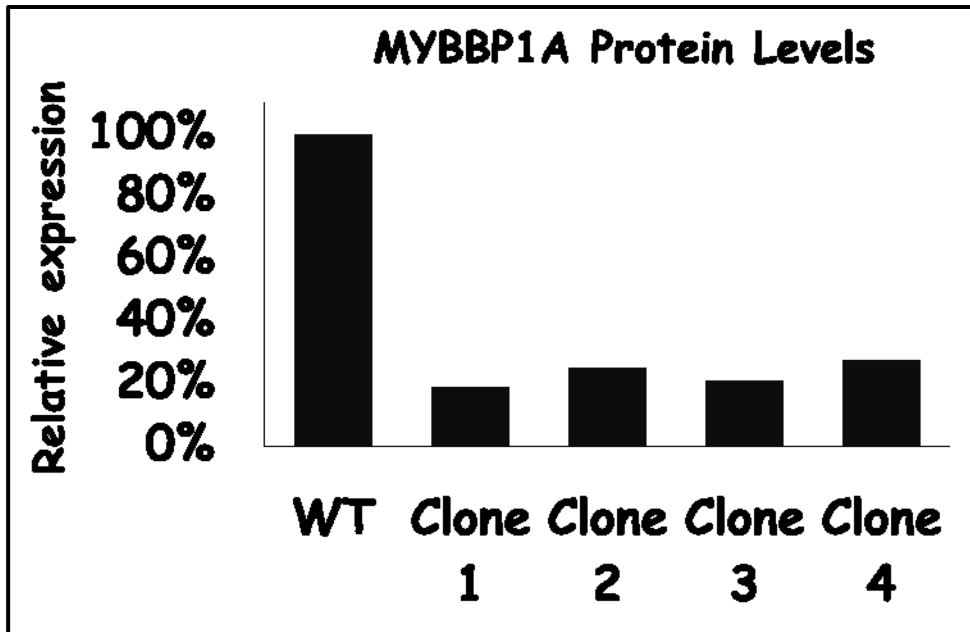
5.11. Supplemental Material

5.11.1. Supplemental Figures



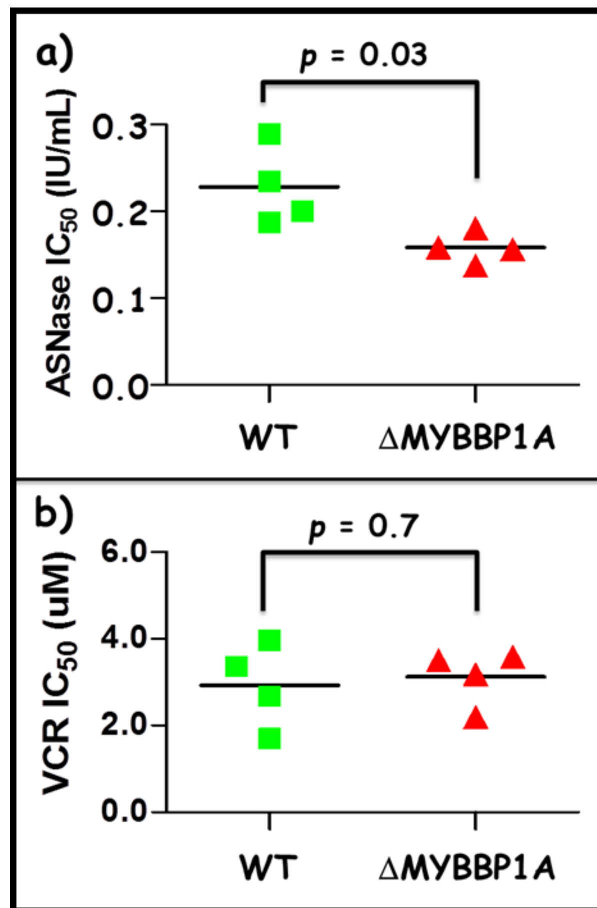
Supplemental Figure S1. Result of the mismatch cleavage assay performed on the cell population used for clonal selection.

The result of the mismatch assay indicates an approximate cleavage efficiency of 30% within the cell population used for clonal selection. This roughly corresponds to the percentage of cells where a gene-editing event occurred.



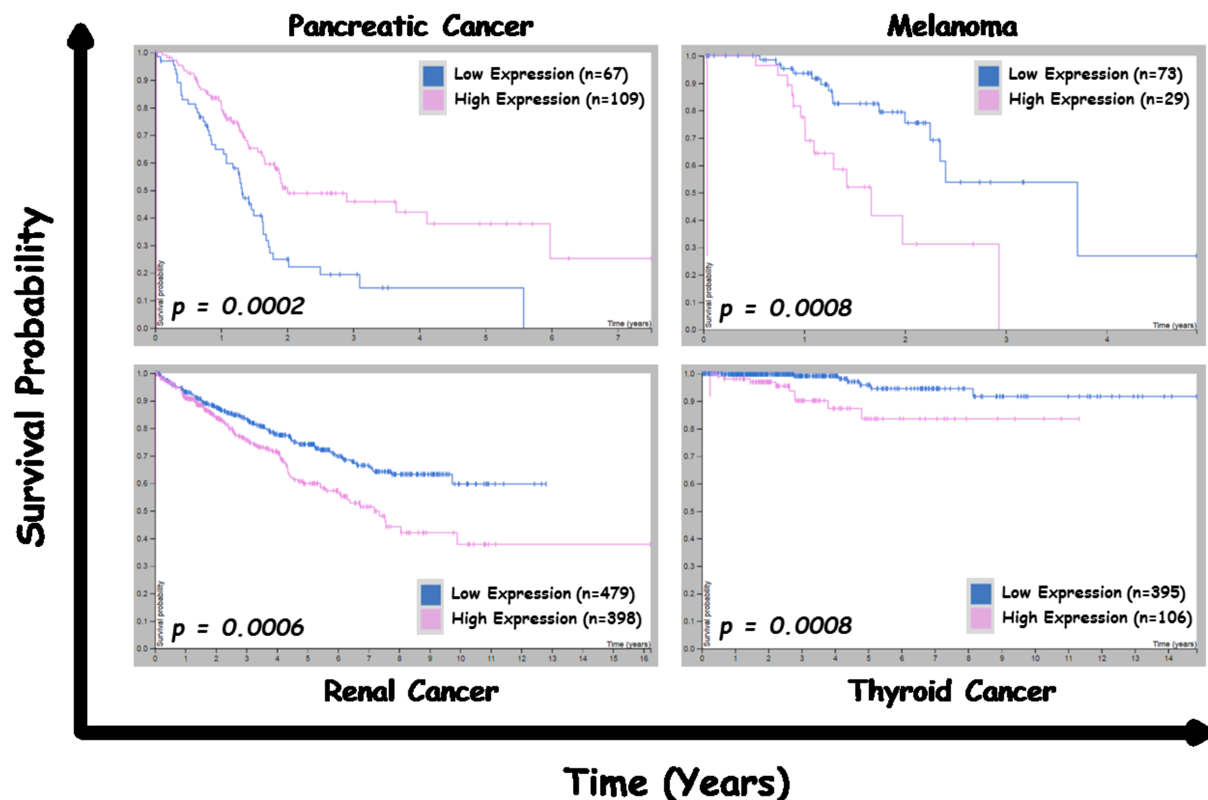
Supplemental Figure S2. Relative protein expression levels of MYBBP1A protein.

Protein levels of MYBBP1A protein in the four ∆MYBBP1A Knockout clones indicate a relative reduction in protein expression compared to the wild-type (WT) control population.



Supplemental Figure S3. *In-vitro* sensitivity to asparaginase and vincristine in relation to *MYBBP1A gene Knock-out 96 hours post incubation with* a) asparaginase (ASNase) or b) vincristine (VCR).

IC50 values were calculated using WST-1 viability assay 96 hours post incubation with several concentrations of the drugs. The experiment was repeated 4 times and IC_{50} values were calculated for each experiment separately. The vertical lines represent the mean IC_{50} value of each group. The coloured shapes represent independent values calculated for WT PANC1 cells (green squares) or MYBBP1A knock-out PANC1 cells (red triangles). The p value obtained by the Unpaired Student's t-test represents the difference in drug sensitivity between the two cell lines and is provided on the top of the graph.



Supplemental Figure S4. Association of MYBBP1A gene expression levels with survival probability in different types of cancer.

The expression of MYBBP1A is prognostic for outcome in several types of cancers but the direction of the effect is dependent on the type of cancer. A high level of expression is significantly associated with favourable outcome in pancreatic cancer, while it portends a worsening prognosis in renal cancer, melanoma and thyroid cancer. High and low expression levels are represented by pink and blue colours, respectively. The p-values obtained by the log rank test for the difference across expression levels are provided on each plot.

This figure was created based on data from publicly available on the Human Protein Atlas website: https://www.proteinatlas.org/ENSG00000132382-MYBBP1A/pathology

5.11.2. Supplemental Methods

Virus production

HEK293 cells were plated in a 6-well plate ($9x10^5$) coated with amine (Fisher Scientific) and were covered with 2ml per well of DMEM + 10% FBS without antibiotic. The next day, the DMEM medium was replaced with RPMI +10% FBS without antibiotic. For transfection, two mixes were prepared: MIX A was composed of 100µL of OPTI-MEM, 300ng pREV, 390ng pVSVG, 750ng pMDL and 450ng of the vector construct of interest. MIX B: contained 100µL of OPTI-MEM and 4µL of lipofectamine 2000. After 5 to 20 minutes of separate incubation at room temperature, the two mixes were combined and left at room temperature for one hour. This new mix (200uL) was then added to the prepared cells for a total volume of 1.2mL and incubated overnight at 37°C. The following day, the medium was removed and replaced with 1.2mL DMEM + 10% FBS + antibiotic (Penicillin + Streptomycin) and the transfected cells were incubated for 30 hours at 37°C. After the incubation period, the culture medium was collected, spinning was done at 3000 rpm for 5 minutes, and the virus-containing supernatant was collected for later use.

MYBBP1A single-guide RNA expression vector construction

Briefly, 5 µg of plasmid was digested with 3 µl FastDigest *BsmBI* (Fermentas) for 30 min at 37°C in the presence of 3 µl FastAP (Fermentas) and 6 µl 10X FastDigest Buffer in a total reaction volume of 60 µl. Digested plasmid was gel purified using QIAquick Gel Extraction Kit. One hundred micromolar of each pair of oligos was phosphorylated and annealed using T4 polynucleotide kinase (New England Biolabs (NEB) M0201S) and 1 µl 10× T4 Ligation Buffer (NEB) in a total volume of 10 µl in a thermal cycler (Applied Biosystems). The cycling conditions were 37°C for 30 min, then 95°C for 5 min, followed by a ramp to 25°C at 5°C/min. The annealed oligo duplex was ligated into the *BbsI*-digested pLentiGuide vector using 5 µl of 5X T4 DNA Ligase Buffer (NEB 15224-041) and 1 µl T4 DNA Ligase (NEB 15224-041) in a total reaction of 11 µl. The ligation mixture was then transformed into STBL3 bacteria and incubated overnight on Luria agar + ampicillin 100µg/mL at 37°C. Mini-prep was then performed for several clones and sent to the McGill platform for sequencing.

Section C

Chapter 6

Genetic risk factors for VIPN in childhood acute lymphoblastic leukemia patients identified using whole-exome sequencing

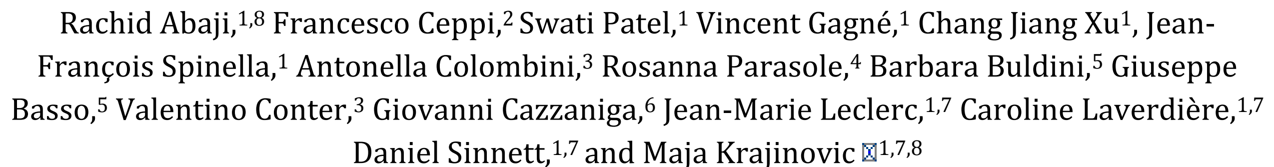
This chapter reports the results of a study that focused on the identification of genetic variants that have the potential to modulate the risk of developing high-grade vincristine-induced peripheral neuropathy and puts forward three genes that have relevant functions in the context of this complication, which merit further investigation.

My involvement in this work spanned the entire project from the production of genotype libraries of the variants of interest that were identified following the initial analysis, to performing the association studies with clinical response parameters in the discovery group as well as testing their reproducibility in the validation cohort. I also analysed the extent of individual contributions of validated variants to the overall combined-effect in modulating the response and constructed the risk prediction model for high-grade vincristine-induced peripheral neuropathy. I drafted the manuscript under the supervision of Dr. Krajinovic, which was then revised by all authors. My contribution to this work can be estimated as 70% of the total input.

This work was published in the peer-reviewed journal, *Pharmacogenomics*.

Genetic risk factors for VIPN in childhood acute lymphoblastic leukemia patients identified using whole-exome sequencing

Running title: WES-identified genetic markers for VIPN


Rachid Abaji,[1,8] Francesco Ceppi,[2] Swati Patel,[1] Vincent Gagné,[1] Chang Jiang Xu[1], Jean-François Spinella,[1] Antonella Colombini,[3] Rosanna Parasole,[4] Barbara Buldini,[5] Giuseppe Basso,[5] Valentino Conter,[3] Giovanni Cazzaniga,[6] Jean-Marie Leclerc,[1,7] Caroline Laverdière,[1,7] Daniel Sinnett,[1,7] and Maja Krajinovic ✉[1,7,8]

[1]Charles-Bruneau Cancer Center, CHU Sainte-Justine Research Center, Montreal, QC, H3T1C5, Canada

[2]Pediatric Hematology-Oncology Unit & Pediatric Hematology-Oncology Research Laboratory, Division of Pediatrics, Department ofWoman-Mother-Child, University Hospital of Lausanne, 1004 Lausanne, Switzerland

[3]Department of Pediatrics, University of Milano-Bicocca, Ospedale S Gerardo, 20835 Monza, Italy

[4]Department of Pediatric Hemato-Oncology, Santobono-Pausilipon Hospital, 80129 Naples, Italy

[5]Department of Woman & Child Health, Laboratory of Haematology-Oncology, University of Padova, 35128 Padova, Italy

[6]Centro Ricerca Tettamanti, Department of Pediatrics, University Milano Bicocca, 20835 Monza, Italy

[7]Department of Pediatrics, Faculty of Medicine, University of Montreal, Montreal, QC, H4A 3J1, Canada

[8]Department of Pharmacology& Physiology, Faculty of Medicine, University of Montreal, Montreal, QC, H3C 3J7, Canada

✉Corresponding author.

Corresponding author:

Name: Maja Krajinovic

Address: Centre de Recherche, CHU Sainte-Justine, 3175 chemin de la Côte-Ste-Catherine, Montréal, QC, H3T 1C5 Canada

E-mail address: maja.krajinovic@umontreal.ca

Phone Number: 514 345-4931 #6259

6.1. Abstract

Aims: To identify genetic markers associated with Vincristine-induced peripheral neuropathy (VIPN) in childhood acute lymphoblastic leukemia (ALL).

Patients & Methods: Whole-exome sequencing data was combined with exome-wide association study to identify predicted-functional germline variants associated with high-grade VIPN. Genotyping was then performed for top-ranked signals (N=237), followed by validation in independent replication group (N=405).

Results: Minor alleles of rs2781377/*SYNE2* (p=0.01) and rs10513762/*MRPL47* (p=0.01) showed increased risk whereas that of rs3803357/*BAHD1* had a protective effect (p=0.007). Using a genetic model based on weighted genetic risk scores, an additive-effect of combining these loci was observed (p=0.003). The addition of rs1135989/*ACTG1* further enhanced model performance (p=0.0001).

Conclusion: Variants in *SYNE2*, *MRPL47* and *BAHD1* genes are putative new risk factors for VIPN in childhood ALL.

<u>Key words:</u>

Acute lymphoblastic leukemia; Vincristine-induced peripheral neuropathy; association study; pharmacogenetics; whole-exome sequencing; adverse drug reactions; polymorphism; genetics; cancer.

6.2. Introduction

Childhood Acute lymphoblastic leukemia (ALL) is the most common pediatric malignancy, accounting for 25% of all childhood cancers [1, 2]. Collaborative efforts have paved the way to modern treatment protocols that achieved cure rates of more than 90% in favourable settings [1-3]. This is attributed to the progressive optimization of anti-leukemia protocols, the use of improved multi-agent therapeutic regimens, and the personalization of treatment through better risk stratification. However, therapy resistance and treatment-induced adverse drug reactions (ADRs) continue to pose important challenges. While enhancing the efficacy of a given treatment is of utmost importance to successful therapy, another aspect of significant importance is to be able to predict the risk of short and long-term side-effects that could be provoked by such a treatment in order to minimize or prevent their occurrence [1].

The vinca alkaloid vincristine (VCR) is a cornerstone medication in most pediatric ALL treatment protocols. Vincristine-induced peripheral neuropathy (VIPN) is a dose-dependent, mostly reversible, ADR and is thought to arise from the impairment of axonal transport secondary to axonal degeneration. It has been associated with doses of 2 to 6 mg/m^2, which is the reason why most protocols cap the dose at 2 mg/m^2 [4-8]. VIPN is associated with debilitating symptoms that may continue to worsen even after stopping the treatment (coasting effect) and can be long-term for many of the survivors, subjecting them to other comorbidities and affecting their quality of life [5, 7-9]. Moreover, depending on their severity, such neuropathic symptoms might necessitate dose reduction, and in certain

cases, treatment interruption or early cessation, which might in turn influence the efficacy of treatment and patient survival [5].

The main form of VIPN is sensory neuropathy which can affect 30%-45% of patients [5, 7, 8, 10]. It is characterized by a symmetrical, length-dependent, glove-and-stocking distribution primarily affecting the large sensory nerve fibers in the extremities of the upper and lower limbs [5, 8, 10]. Other forms of VIPN, which are less common, are autonomic and motor neuropathies; the latter affecting 5-10% of patients [10] and influencing their ability to walk and perform fine motor tasks [5, 6, 8, 10]. Sensory and motor VIPN are thought to affect older children more than younger ones, whereas the impact of sex remains largely debatable with female gender being more often associated with an increased risk of toxicity [11, 12]. The incidence of the more serious, high-grade VIPN (grades III and IV), is reported in up to 37% of cases [11].

Earlier studies concluded that variability in the risk of VIPN could not be directly predicted via pharmacokinetics [12, 13], suggesting the role of other contributing factors, including genetic predisposition [14]. Indeed, genetic role in VIPN has been well documented, in Charcot-Marie-Tooth disease (CMT), a hereditary sensorimotor neuropathy associated with severe forms of neurotoxicity when treated with vincristine [5, 8, 15-17]. Several polymorphisms associated with VIPN in childhood ALL have been previously identified. A widely studied example is the *CYP3A5* gene whose CAP3A5 *3 variant showed inconsistent results across the literature [13, 14, 18-20]. Additional studies identified SNPs with protective effect against the toxicity {e.g. those in the *ABCB1* and *CAPG* genes [4]} whereas others increase the risk of VIPN {e.g. *ACTG1* [4], *CEP72* [21-23], *ABCC1* [23] *ABCC2*

[24] *VDR* [19] and several SNPs in genes of the glucocorticoid pathway [19]}. Particularly interesting is the association derived through genome-wide association study of rs924607 in the *CEP72* gene encoding a centrosomal protein involved in microtubule formation [21]. It was associated with increased risk of VIPN (grades II-IV) and CEP72 mRNA levels, which in turn affected cellular sensitivity to VCR in functional assays [21]. Other research groups targeted polymorphisms in micro-RNAs which could regulate VCR-related genes, with one study reporting positive associations (miR-3117, miR-4481 and miR-6076) [25] while another reporting negative results [24]. Polymorphisms in other genes, alone or in combination, could further contribute to VIPN and might help explain the remaining variability in VCR responses [3, 26].

Since there is currently no preventative, neuroprotective or curative treatment for VIPN [5, 27], it could be beneficial to develop early-detection strategies based on a comprehensive understanding of the pathophysiological mechanisms implicated in the development of this toxicity as well as on patient-specific risk factors such as clinical and genetic predisposition.

Here we report on combination of whole-exome sequencing data and association study that led to the identification of additional SNPs associated with VIPN as well as their additive effect when multiple risk loci were combined in weighted genetic risk score (wGRS) model.

6.3. Patients and Methods

6.3.1. Study population and endpoints in the analysis

Discovery set was composed of 237 French-Canadian patients of European origin who belong to the well-characterized Quebec childhood ALL (QcALL) cohort [28-30] and for whom data regarding the presence/absence of high-grade neurotoxicity during the treatment were available (Table 1) [4]. These patients were diagnosed with ALL and treated at the University Health Centre Sainte-Justine (UHCSJ), Montreal, QC, Canada, between January 1989 and July 2005 [4, 30-32]. All patients received VCR as part of the Dana-Farber Cancer Institute (DFCI) ALL Consortium protocols 87-01, 91-01, 95-01, or 00-01. Specific details of the administration schedule and dose intensity can be found elsewhere [4, 31, 32]. Briefly, the induction phase of all four protocols involves the administration of a standard weekly dose of VCR (1.5 mg/m^2) for four doses. DFCI 91-01 includes an extra dose of (1.5 mg/m^2), while DFCI 95-01 and 00-01 each involve a similar fifth dose but capped at 2 mg. As for the consolidation and continuation phases, the same dose of VCR in all four protocols is administered every 3 weeks (2 mg/m^2 for a maximum of 2 mg and a total of 100 weeks of treatment) [4, 31, 32].

The VIPN data in the QcALL cohort were obtained previously [4] through patients` medical charts evaluation which included clinical signs and symptoms graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE, version 3.0) and was confirmed by a documented subsequent dose reduction for all cases.

Information on all grades of neurotoxicity was available, but the exploratory association analysis focused only on patients with higher-grade neurotoxicity (i.e. grade III or IV vs. grade 0) which is deemed to be more clinically important. Essentially, grade III neurotoxicity is defined as limiting the performance of basic activities of self-care while grade IV is considered life-threatening.

The replication cohort consisted of 405 childhood ALL patients that share similar characteristics with the discovery cohort (Table 1) but who were treated in the context of the Italian Association of Pediatric Haematology and Oncology (*Associazione Italiana di Ematologiae Oncologia Pediatrica*, AIEOP) arm of the AIEOP and the Berlin-Frankfurt-Münster Acute Lymphoblastic Leukemia (AIEOP-BFM ALL 2000) study protocol [33]. Data on high-grade VIPN were available for all patients (also following the CTCAE, version 3.0) and targeted genotyping was performed to obtain genetic data on the polymorphisms identified as significantly associated in the discovery cohort. The AIEOP-BFM ALL 2000 protocol incorporates the minimum residual disease levels into the stratification algorithm in order to establish 3 distinct risk groups defined as standards, intermediate and high risk. Like with the DFCI protocols, all patients receive four weekly doses of VCR (1.5 mg/m^2) during the induction phase; however, following a consolidation phase where no VCR is administered, the high-risk (HR) group of patients uniquely receive 3 consecutive treatment blocks (HR1, HR2 and HR3) in which 4 more doses of VCR (1.5 mg/m^2) are administered. During the re-induction phase, patients receive between 2 and 8 VCR doses (1.5 mg/m^2) depending on their risk group and the randomization arm, while no VCR is administered during the maintenance phase [33].

6.3.2. *Methods*

Whole exome sequencing (WES) data of tumor and germinal genome were obtained through the mutation ALL screening of the QcALL cohort [28-30]. Information on constitutive genetic variants were available for 179 patients along with VIPN data through our initiative at Cancer Research Center UHCSJ to catalogue somatic mutations and germline polymorphisms using whole-exome sequencing. Exome-wide association study (EWAS) with high-grade VIPN (grades III/IV) vs. no-toxicity (grade 0) was performed. Similar design, as described previously [29], aimed to reduce the complexity of analysis was used, whereby EWAS focused on functionally predicted common variants. Specific details on sequencing, variant calling, and association analyses [29, 34] are provided in the supplemental material (Supplemental Methods). Top-ranking EWAS signals with false discovery rate (FDR) <15% were further confirmed via genotyping in the entire discovery cohort followed by a replication analysis in a validation group. Univariate c*omparison* between genotypes and the frequency of VIPN was performed using χ^2 or Fisher-Exact tests. The genetic model that was most representative of the effect of a given variant (i.e. additive, dominant, or recessive) was also tested. The genotype-associated risk was expressed as odds ratio (OR) with 95% confidence interval (CI). Multivariate analysis was performed using logistic regression and included the categorical variables which had p-values ≤ 0.1 in univariate analysis: i.e. genotype, age (< 10 years or ≥ 10 years), risk (standard/high in QcALL; or standard, intermediate and high in the AIEOP) and DFCI protocol (in the QcALL cohort only). The potential of a combined-effect for the significantly associated variants was

investigated by calculating the weighted genetic risk scores (wGRS). Statistical analyses were performed with IBM SPSS Statistics for Windows, Version 24.0. (IBM Corp. Armonk, NY).

6.3.2.1. *Risk Prediction using* weighted Genetic Risk Score

Briefly, weighted Genetic Risk Scores (wGRS) was estimated from the number of risk alleles present for each patient by calculating the sum of weighted natural logarithm of OR, ln(OR), for each allele as explained elsewhere [29, 35]. The models included either all 3 SNPs found to be significantly associated in this study with VCR neurotoxicity or additionally included another associated variant (rs1135989 in ACTG1) identified previously by candidate gene approach in the same discovery cohort [4]. Area under the receiver operator characteristic (ROC) curve was used to assess the classification efficiency of the model in the discovery group and the prediction efficiency in the replication set.

6.4. Results

6.4.1. Vincristine-induced peripheral neuropathy

Out of the 237 patients included in the neurotoxicity study, thirty-five patients (14.8%) had high-grade toxicity (i.e. grades III/IV) (Table 1). There was no difference in patients' characteristics or in VIPN between the entire discovery cohort (N=237) and the subgroup of patients (N=179) for which WES data were available (Table 1). The observed

frequencies of high-grade VIPN in the discovery and replication cohorts were both within the range reported in literature [11]. However, patients in the replication cohort were at the lower end of this range (3.2%) (Table 1).

6.4.2. Association Study

A total of 5527 common variants with minor allele frequency of ≥ 5% (according to 1000 genomes database) and a predicted functional impact were recovered from the WES data. Of these, 4543 SNPS (located in 3802 genes) satisfied the Hardy-Weinberg equilibrium. The association analysis of these variants with VIPN led to the identification of 21 SNPs (distributed across 19 genes) significant at FDR <15%, (Supplemental Table S1). Seven variants were excluded from further analyses (due to linkage-disequilibrium, MAF lower than 5% in the QcALL cohort, or located in a gene coding for an olfactory receptor, Supplemental Table 1). This resulted in 14 top-ranking SNPs that were subjected to a confirmatory step by genotyping (Figure 1 and Supplemental Table S1). Three SNPs were found significantly associated in entire cohort upon genotyping (Figure 1; Table 2). Carriers of the minor allele of rs2781377 in the *SYNE2* gene had an increased risk of VIPN which was proportional to the number of copies of the risk allele A (OR = 2.5; 95% CI, 1.2–5.2; p= 0.01) whereas the effect of rs10513762 minor allele T in the *MRPL47* gene followed the dominant model (OR = 3.3; 95% CI, 1.4–7.7; p= 0.01). In contrast, the minor allele of rs3803357 in the *BAHD1* gene had a protective effect against high grade VIPN (Table 2) as the presence of the polymorphism was associated with a lower incidence of toxicity in the dominant model (OR = 0.35; 95% CI, 0.2–0.7; p= 0.007). All of the identified associations remained significant in

the multivariate analysis in their respective models (Table 2). Moreover, there was a significant correlation between the presence of any of the studied polymorphisms and the reduction in the cumulative VCR dose received, as depicted by the average dose administered -expressed as a percentage of the planned cumulative full dose (Supplemental Table S2). An association was also noted between the number of episodes of high-grade VIPN and the SNPs in *BAHD1* and *MRPL47* genes (Supplemental Table S3). We also verified whether identified SNPs were associated with grades I/II; None of the SNPs showed a significant association (Supplemental Table S4).

6.4.3. *Combined Effect Model*

To evaluate the potential of an additive effect of combining risk loci on the risk of VIPN, we applied the weighted genetic risk score (wGRS) method [35]. The individual contribution of each SNP (i.e. rs2781377, rs10513762 and rs3803357) was calculated from the ln(OR) derived from the multivariate logistic regression model and multiplied by the number of variant alleles carried by each patient. We tested the classification capacity of the model using the area under the ROC (AUC) curve to determine the efficiency of the model, which was: AUC= 0.68 ± 0.05; 95% CI, 0.58-0.79 (p= 0.0005; Figure 2A). Next, patients were divided into 2 groups using the distribution above/below the median (calculated for control patients exhibiting no VIPN) thus giving rise to high/low risk groups, respectively. We then tested the difference in the frequency of VIPN between the two groups and found a significant increase in the number of patients with VIPN in the high risk group as compared to the low risk group (p= 0.002; Figure 2B).

6.4.4. Replication Analysis

None of the significant associations with high-grade VIPN identified in the discovery cohort reached the significance level in the replication cohort. However, when performing a meta-analysis of the two cohorts, all three associations maintained their significance (Supplemental Figure S1). Moreover, when using wGRS to calculate the individual risk in the replication set using the same ORs and cut-off values derived from the discovery cohort, the same pattern of distribution for patients with VIPN was observed (Figure 2B); the difference in the frequency of patients with VIPN between the two risk groups was of borderline-significance (p= 0.09; Figure 2B).

6.4.5. Risk Prediction

In an attempt to increase the discrimination ability of the combined genetic effect model, rs1135989 in the *Actin Gamma 1* (*ACTG1*) gene, was incorporated into the model. This variant was previously found to be significantly associated with an increased risk of high-grade VIPN in the same QcALL cohort using candidate gene approach which investigated components of VCR pathway and was the only risk variant to maintain a significant association in the multivariate model in that study [4]. Accordingly, the new wGRS calculation was based on the sum of the weighted individual contribution of each of the 4 SNPs which substantially enhanced the performance of the model (AUC = 0.70 ± 0.05; 95% CI, 0.60-0.81; *p*= 0.0001). Patients were then divided into 3 groups based on their

genetic risk scores. Those with (wGRS < 0) were considered to have a low risk for toxicity while those who had scores (0 ≥ wGRS ≤ 0.474) were grouped into one intermediate risk group and those with (wGRS > 0.474) were assigned to the high risk group. A linear association was observed across the risk groups (OR = 2.4; 95% CI, 1.5–3.9; *p*= 0.0001; Figure 3A) with significantly higher frequency of patients developing VIPN in the high risk group as compared with the low risk group (OR = 5.7; 95% CI, 2.2–14.5; *p*= 0.0002; Figure 3A).

To assess the performance of the new genetic model in predicting the risk of VIPN, a risk score was assigned to each patient in the validation cohort by calculating the wGRS across the 4 loci using the same OR values derived from the discovery cohort. Patients were grouped into risk groups using the same cut-off values for the low, intermediate and high genetic-risk groups of the discovery cohort (Figure 3A). Area under the ROC curve (AUC) in the replication cohort was significantly above the random prediction line and was identical to that obtained in discovery cohorts (Figure 3B, AUC = 0.68). A linear increase in the risk was observed across the risk groups in the replication cohort (OR = 2.2; 95% CI, 1.1–4.4; p= 0.01; Figure 3B). Patients predicted to have the highest risk of VIPN had significantly higher frequency of individuals who actually developed VIPN and the observed OR of this group was significantly greater than that of the low risk group (OR = 5.1; 95% CI, 1.3–21.3; p= 0.02; Figure 3B).

6.5. Discussion

Recent advances in pharmacogenetics have identified genetic polymorphisms that could contribute to the observed inter-individual variability in susceptibility to VIPN [4, 13, 18, 19, 21, 22, 24, 25]. However, there is still no consensus on the actionable associations between genetic variants and the risk of VIPN [5, 17], since not all SNPs are necessarily applicable to all protocols or treatment phases [36, 37].

By combining whole-exome sequencing and an exome-wide association study strategy, we identified two common variants significantly associated with an increased risk of high-grade VIPN, rs2781377 in the *SYNE2* gene and rs10513762 in the *MRPL47* gene, whereas the one of rs3803357 in the *BAHD1* gene played a protective role. The multivariate logistic regression analysis indicates that the associations are independent of other non-genetic factors for patients treated according to the DFCI protocol.

These variants were also associated with the number of high-grade episodes and with the reduction in VCR dose. Moreover, it is important to note *that* since VCR was withheld (or the dose was reduced) whenever patients experienced VIPN, there was, a correlation in the discovery cohort between the development of VIPN and the inability to complete the full cumulative VCR dose; which precluded any meaningful adjustment based on the total VCR dose received in the multivariate analysis. Furthermore, there was no significant association between the survival outcome -in terms of event-free survival and overall survival- and the presence or absence of VIPN, or with the dose of VCR administered.

MRPL47 belongs to the MRP family of genes which encode the mitochondrial ribosomal proteins essential for mitochondrial protein synthesis. They play an important role in the oxidative phosphorylation system suggesting that mutations in these genes could be linked to neuropathies, myopathies and developmental disorders due to a reduced capacity for ATP production [38]. Indeed, many of these genes were mapped to regions associated with sensorineural disorders [38, 39]. Taken together, this could suggest a state of genetic predisposition to the toxicity of vincristine similar to the association observed with CMT syndrome. Another explanation of the observed association is based on the chromosomal localisation of *MRPL47* relative to the *BAF53A* gene. The latter belongs to a chromatin remodeling complex required for proliferation and differentiation of neural stem-cells and neuronal development [40]. This hypothesis stems from the fact that the two genes are found adjacent to each other in an antisense, tail-to-tail orientation which raises the possibility of a regulated alternate expression [41].

The SNP in the *SYNE2* gene is a G-to-A polymorphism resulting in a stop-gain mutation. *SYNE2* (Spectrin repeats containing nuclear envelope 2) also known as nesprin-2 (Nuclear envelope spectrin-repeat protein-2) codes for a multi-isomeric nuclear-envelope anchored protein which serves as a linker within the cellular cytoskeleton. It interacts with the nuclear lamina and plays an important role in various cellular and nuclear functions including DNA damage repair, chromatin organization, chromosome movement, organelle positions, cells signalling and cell polarity/migration [42, 43]. Nesprins have been linked to neurological diseases and thought to play a critical role in neurogenesis and neuronal migration. [42-44]. Interestingly, this same variant was previously found to be affiliated to Emery-Dreifuss muscular dystrophy, which was reported to be associated with axonal neuropathy in several cases [45-49].

Bromo adjacent homology domain containing protein 1 (*BAHD1*) is an important regulator of gene silencing through heterochromatin formation. Previous studies have linked *BAHD1* to tumor suppression and inflammation, and identified its important role in the repression of proliferative and survival genes {such as the insulin-like growth factor II, *IGF2* [50], the control of steroid and lipid metabolism [51]}, or acting as an inflammation regulatory factor through the TNF signalling pathway [52]. It was also suggested that *BAHD1* has a crucial role in controlling the spatial architecture of the genome and that a dysfunctional BAHD1 complex could be the cause of many diseases due to an aberrant epigenetic signature, which in itself, was linked to sensory and autonomic neuropathy [51, 53, 54].

It is worth mentioning that, while each of these three exonic polymorphisms is associated with a modification in the amino acid sequence and could exert its impact by changing the protein function, another possibility lies in the ability of these SNPs to alter the expression of their respective genes. Indeed, expression data from the GTEx database [55] suggest that all of the identified variants are associated, to varying extents, with a differential expression of their genes (Supplemental Figure S2). Moreover, any of these polymorphisms have the potential to alter the expression of nearby or distant genes through which it could be altering the risk of VIPN. An interesting example is the very strong association between rs3803357 and *C15orf57* ($p=8x10^{-30}$ in the Tibial-Nerve) which merits further investigation [55].

The associations failed to replicate in the validation cohort, which is likely due to differences in the dose, intensity, frequency and the duration of treatment with VCR. Patients in the AIEOP cohort generally receive lower quantities of VCR per dose that is also administered less frequently and for shorter duration of time. This can also explain lower frequency of high-grade VIPN in AIEOP cohort, which could have reduced the power needed to detect single SNP associations. The differences in the definition of adverse events or the strategies for their identification and reporting cannot be ruled out [1, 5].

However, a similar genetic effect was noted between the discovery and replication cohorts when the risk of VIPN was analyzed in relation to the number of risk-loci carried by each patient, suggesting the presence of a synergistic effect. In fact, patients of the QcALL cohort who experienced VIPN had significantly higher risk scores when compared to the ones with no VIPN. Using the distribution of scores around the median, patients were successfully divided into two risk groups that differed significantly in their risk of developing VIPN. A similar trend of borderline significance was seen in the validation cohort. Moreover, when performing a meta-analysis of the two cohorts, all three associations maintained their significance which could be indicative of the stability of the associations in a larger cohort (Supplemental Figure S1). However, with the exception of a small gain in the significance for rs10513762 in the *MRPL47* gene, the effects in the combined cohort seem to be driven by the strong associations in the discovery cohort.

Given the observed strength of the combined-effect model, we introduced rs1135989 in the *Actin Gamma 1* (*ACTG1) gene*, which encodes for the major cytoskeletal protein ACTG1 [4] and which we previously identified through candidate gene approach as

significantly associated with high grade VIPN in the same discovery cohort. The new scores were significantly better at classifying patients into risk groups as shown by the increase in the AUC of the ROC curve and successfully predicted the risk of VIPN in the validation cohort with almost identical values between predicted and observed ORs.

We acknowledge that there are some limitations to this study. For example, the distribution of treatment protocols and risk groups varied significantly between the two cohorts, which could have introduced variability as patients might have received considerably different VCR doses. In addition, while this study only included patients with European origins, it is important to note that ethnicity can play an important role due to inter-population differences in SNPs prevalence. Also, the sample size of the discovery cohort was relatively small and the selected FDR threshold of <15% was relaxed, which might have increased the number of false-positives. Moreover, this study did not adjust for the intake of other drugs that have the potential to impact the pharmacokinetics/pharmacodynamics of VCR [56], or for other genetically predisposing factors, or the presence/absence of CMT syndrome associated with the risk of neurotoxicity; which could have modified the magnitude of the observed effect [5, 8, 15-17]. Furthermore, the design of this study, using whole-exome data, did not allow for testing or validating the effect of polymorphisms located in the introns or promoter regions of other genes previously reported in GWAS or candidate-gene studies [21, 23] and thus their potential influence on the risk of VIPN cannot be ruled out. However, the fact that the individually observed associations with VIPN were not replicated whereas the much stronger combined-effect of these associations was, could be used to support the idea that the replication failure was not due to a lack of an individual SNP effect but rather to a reduced power of detection.

6.6. Conclusion

In conclusion, the identification of patients who are at high risk of developing VIPN remains important and might help clinicians with the individualization of treatment to reduce the frequency and intensity of VIPN. This can be particularly relevant for patients who are already considered to be at higher risk of peripheral neuropathy like patients with CMT syndrome or diabetes mellitus. This study identified three genetic markers associated with modulation of the risk of VCR related neurotoxicity and whose functions have the potential to explain the observed effect. All of these SNPs merit being further investigated in replication analysis with standardized objective measures of neuropathy and larger number of patients as well as in functional assays. We have also shown that while single associations might have mild effects and thus be difficult to validate in replication cohorts (especially when following different protocols), the strong effect of combining those variants might prove useful in enhancing the detection power and predicting the risk of VIPN in a model that incorporates information on multiple loci.

6.7. Summary Points

- Vincristine-induced peripheral neuropathy (VIPN) is a common adverse-event to Vincristine for which there is currently no preventative, neuroprotective or curative treatment.

- There is no consensus on actionable genetic markers that can predict and/or influence the risk of VIPN which can be used for treatment individualization.

- Using whole-exome sequencing (WES) data in the context of association study, we identified three SNPs associated with modulation of the risk of VIPN in *SYNE2*, *MRPL47* and *BAHD1* genes.

- All three genes have relevant functions in the context of VIPN and merit further investigation.

- Minor alleles of rs2781377 in *SYNE2* and rs10513762 in *MRPL47* showed increased risk (OR=2.5; 95%CI, 1.2–5.2; p=0.01 and OR=3.3; 95%CI, 1.4–7.7; p=0.01, respectively), whereas the minor allele of rs3803357 in *BAHD1* had a protective effect (OR=0.35; 95%CI, 0.2–0.7; p=0.007).

- All three polymorphisms were also association with vincristine dose reduction, number of neurotoxicity episodes and all-grades VIPN (grade I-IV).

- The combined-effect genetic model using weighted genetic risk score shows an additive effect for identified risk alleles (p=0.003).

- We also present a prediction model combining multiple loci identified through WES or candidate gene approach to improve the prediction of VIPN in childhood ALL treatment with good efficiency (p=0.0001 and p=0.01, in the discovery and replication cohorts, respectively).

6.8. Acknowledgements

We thank all patients and their parents who consented to participate in genetics studies related to leukemia. Leukemia Lymphoma Society of Canada, Canadian Institutes of Health Research, Charles Bruneau Foundation, Cancer Research Society and Associazione Italiana per la Ricerca sul Cancro (AIEOP) supported this study. RA and JFS are recipients of fellowships from Cole Foundation and Network of Applied Medical Genetics. The sequencing data were obtained through TCF grant supported by the Institute of Cancer Research (ICR) of the Canadian Institutes of Health Research (CIHR), in collaboration with C17 Council, Canadian Cancer Society (CCS), Cancer Research Society (CRS), Garron Family Cancer Centre at the Hospital for Sick Children, Ontario Institute for Cancer Research (OICR) and Pediatric Oncology Group of Ontario (POGO).

6.9. Authorship Contributions

M.K. designed the study; R.A., S.P. and V.G. performed experiments; F.C. performed medical chart reviews; C.L. JM.L., A.C., R.P., B.B, G.B., V.C., G.C. and D.S. contributed to sample and clinical data collection and interpretation; D.S. supervised whole exome sequencing; R.A, J-F.S., CJ.X, V.G. and M.K. performed the data analysis; R.A. drafted the article; All authors revised the manuscript.

6.10. Disclosure of Conflicts of Interest

The authors declare no competing financial interests.

6.11. Ethical conduct of research

This study was approved by the Research Ethics Committee of the CHU Sainte-Justine Research Center and had the following file number: 1664. The AIEOP-BFM ALL 2000 study was approved by the Ethics committees of all participating institutions and by IRB of CHU Sainte Justine. Written informed consent was obtained in accordance with the Declaration of Helsinki from all participants and/or their parents or legal guardians.

*****Supplementary information is available at the *Pharmacogenomics* journal's website*****

6.12. References

1. Schmiegelow K, Attarbaschi A, Barzilai S *et al*. Consensus definitions of 14 severe acute toxic effects for childhood lymphoblastic leukaemia treatment: a Delphi consensus. *The Lancet Oncology* 17(6), e231-e239 (2016).

2. Hough R, Rowntree C, Goulden N *et al*. Efficacy and toxicity of a paediatric protocol in teenagers and young adults with Philadelphia chromosome negative acute lymphoblastic leukaemia: results from UKALL 2003. *Br J Haematol* 172(3), 439-451 (2016).

3. Mei L, Ontiveros EP, Griffiths EA, Thompson JE, Wang ES, Wetzler M. Pharmacogenetics predictive of response and toxicity in acute lymphoblastic leukemia therapy. *Blood Rev* 29(4), 243-249 (2015).

4. Ceppi F, Langlois-Pelletier C, Gagne V *et al*. Polymorphisms of the vincristine pathway and response to treatment in children with childhood acute lymphoblastic leukemia. *Pharmacogenomics* 15(8), 1105-1116 (2014).

5. Park SB, Goldstein D, Krishnan AV *et al*. Chemotherapy-induced peripheral neurotoxicity: a critical analysis. *CA Cancer J Clin* 63(6), 419-437 (2013).

6. Haim N, Epelbaum R, Ben-Shahar M, Yarnitsky D, Simri W, Robinson E. Full dose vincristine (without 2-mg dose limit) in the treatment of lymphomas. *Cancer* 73(10), 2515-2519 (1994).

7. Verstappen CC, Koeppen S, Heimans JJ *et al*. Dose-related vincristine-induced peripheral neuropathy with unexpected off-therapy worsening. *Neurology* 64(6), 1076-1077 (2005).

8. Kerckhove N, Collin A, Conde S, Chaleteix C, Pezet D, Balayssac D. Long-Term Effects, Pathophysiological Mechanisms, and Risk Factors of Chemotherapy-Induced Peripheral Neuropathies: A Comprehensive Literature Review. *Front Pharmacol* 8 86 (2017).

9. Kim BJ, Park HR, Roh HJ *et al*. Chemotherapy-related polyneuropathy may deteriorate quality of life in patients with B-cell lymphoma. *Qual Life Res* 19(8), 1097-1103 (2010).

10. Windebank AJ, Grisold W. Chemotherapy-induced neuropathy. *J Peripher Nerv Syst* 13(1), 27-46 (2008).

11. Lavoie Smith EM, Li L, Chiang C *et al*. Patterns and severity of vincristine-induced peripheral neuropathy in children with acute lymphoblastic leukemia*. J Peripher Nerv Syst* 20(1), 37-46 (2015).

12. Crom WR, De Graaf SS, Synold T *et al*. Pharmacokinetics of vincristine in children and adolescents with acute lymphocytic leukemia*. J Pediatr* 125(4), 642-649 (1994).

13. Moore AS, Norris R, Price G *et al*. Vincristine pharmacodynamics and pharmacogenetics in children with cancer: a limited-sampling, population modelling approach*. J Paediatr Child Health* 47(12), 875-882 (2011).

14. Egbelakin A, Ferguson MJ, Macgill EA *et al*. Increased risk of vincristine neurotoxicity associated with low CYP3A5 expression genotype in children with acute lymphoblastic leukemia*. Pediatr Blood Cancer* 56(3), 361-367 (2011).

15. Naumann R, Mohm J, Reuner U, Kroschinsky F, Rautenstrauss B, Ehninger G. Early recognition of hereditary motor and sensory neuropathy type 1 can avoid life-threatening vincristine neurotoxicity*. Br J Haematol* 115(2), 323-325 (2001).

16. Orejana-Garcia AM, Pascual-Huerta J, Perez-Melero A. Charcot-Marie-Tooth disease and vincristine*. J Am Podiatr Med Assoc* 93(3), 229-233 (2003).

17. Nishikawa T, Kawakami K, Kumamoto T *et al*. Severe neurotoxicities in a case of Charcot-Marie-Tooth disease type 2 caused by vincristine for acute lymphoblastic leukemia*. J Pediatr Hematol Oncol* 30(7), 519-521 (2008).

18. Aplenc R, Glatfelter W, Han P *et al*. CYP3A genotypes and treatment response in paediatric acute lymphoblastic leukaemia*. Br J Haematol* 122(2), 240-244 (2003).

19. Kishi S, Cheng C, French D *et al*. Ancestry and pharmacogenetics of antileukemic drug toxicity*. Blood* 109(10), 4151-4157 (2007).

20. Hartman A, Van Schaik RH, Van Der Heiden IP *et al*. Polymorphisms in genes involved in vincristine pharmacokinetics or pharmacodynamics are not related to impaired motor performance in children with leukemia*. Leuk Res* 34(2), 154-159 (2010).

21. Diouf B, Crews KR, Lew G *et al*. Association of an inherited genetic variant with vincristine-related peripheral neuropathy in children with acute lymphoblastic leukemia*. JAMA* 313(8), 815-823 (2015).

22. Stock W, Diouf B, Crews KR *et al*. An Inherited Genetic Variant in CEP72 Promoter Predisposes to Vincristine-Induced Peripheral Neuropathy in Adults With Acute Lymphoblastic Leukemia*. Clin Pharmacol Ther* 101(3), 391-395 (2017).

23. Wright GEB, Amstutz U, Drogemoller BI *et al*. Pharmacogenomics of vincristine-induced peripheral neuropathy implicates pharmacokinetic and inherited neuropathy genes*. Clin Pharmacol Ther* doi:10.1002/cpt.1179 (2018).

24. Lopez-Lopez E, Gutierrez-Camino A, Astigarraga I *et al*. Vincristine pharmacokinetics pathway and neurotoxicity during early phases of treatment in pediatric acute lymphoblastic leukemia*. Pharmacogenomics* 17(7), 731-741 (2016).

25. Gutierrez-Camino A, Umerez M, Martin-Guerrero I *et al*. Mir-pharmacogenetics of Vincristine and peripheral neurotoxicity in childhood B-cell acute lymphoblastic leukemia*. Pharmacogenomics J* doi:10.1038/s41397-017-0003-3 (2017).

26. Van De Velde ME, Kaspers GL, Abbink FCH, Wilhelm AJ, Ket JCF, Van Den Berg MH. Vincristine-induced peripheral neuropathy in children with cancer: A systematic review*. Crit Rev Oncol Hematol* 114 114-130 (2017).

27. Albers J, Chaudhry V, Cavaletti G, Donehower R. Interventions for preventing neuropathy caused by cisplatin and related compounds*. Cochrane Database Syst Rev* doi:10.1002/14651858.CD005228.pub2(1), CD005228 (2007).

28. Healy J, Richer C, Bourgey M, Kritikou EA, Sinnett D. Replication analysis confirms the association of ARID5B with childhood B-cell acute lymphoblastic leukemia*. Haematologica* 95(9), 1608-1611 (2010).

29. Abaji R, Gagne V, Xu CJ *et al*. Whole-exome sequencing identified genetic risk factors for asparaginase-related complications in childhood ALL patients*. Oncotarget* 8(27), 43752-43767 (2017).

30. Rousseau J, Gagne V, Labuda M *et al*. ATF5 polymorphisms influence ATF function and response to treatment in children with childhood acute lymphoblastic leukemia*. Blood* 118(22), 5883-5890 (2011).

31. Silverman LB, Gelber RD, Dalton VK *et al*. Improved outcome for children with acute lymphoblastic leukemia: results of Dana-Farber Consortium Protocol 91-01*. Blood* 97(5), 1211-1218 (2001).

32. Silverman LB, Stevenson KE, O'brien JE *et al*. Long-term results of Dana-Farber Cancer Institute ALL Consortium protocols for children with newly diagnosed acute lymphoblastic leukemia (1985-2000)*. Leukemia* 24(2), 320-334 (2010).

33. Conter V, Bartram CR, Valsecchi MG *et al*. Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study*. Blood* 115(16), 3206-3214 (2010).

34. Spinella JF, Mehanna P, Vidal R *et al*. SNooPer: a machine learning-based method for somatic variant identification from low-pass next-generation sequencing. *BMC Genomics* 17(1), 912 (2016).

35. De Jager PL, Chibnik LB, Cui J *et al*. Integrating genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility. *Lancet neurology* 8(12), 1111-1119 (2009).

36. Gutierrez-Camino A, Martin-Guerrero I, Lopez-Lopez E *et al*. Lack of association of the CEP72 rs924607 TT genotype with vincristine-related peripheral neuropathy during the early phase of pediatric acute lymphoblastic leukemia treatment in a Spanish population. *Pharmacogenet Genomics* 26(2), 100-102 (2016).

37. Diouf B, Crews KR, Evans WE. Vincristine pharmacogenomics: 'winner's curse' or a different phenotype? *Pharmacogenet Genomics* 26(2), 51-52 (2016).

38. O'brien TW. Evolution of a protein-rich mitochondrial ribosome: implications for human genetic disease. *Gene* 286(1), 73-79 (2002).

39. Kenmochi N, Suzuki T, Uechi T *et al*. The human mitochondrial ribosomal protein genes: mapping of 54 genes to the chromosomes and implications for human disorders. *Genomics* 77(1-2), 65-70 (2001).

40. Lessard J, Wu JI, Ranish JA *et al*. An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron* 55(2), 201-215 (2007).

41. Genecards. http://www.genecards.org/cgi-bin/carddisp.pl?gene=MRPL47&keywords=MRPL47. (Accessed: 01-09-2017).

42. Rajgor D, Shanahan CM. Nesprins: from the nuclear envelope and beyond. *Expert Rev Mol Med* 15 e5 (2013).

43. Cartwright S, Karakesisoglou I. Nesprins in health and disease. *Semin Cell Dev Biol* 29 169-179 (2014).

44. Zhang X, Lei K, Yuan X *et al*. SUN1/2 and Syne/Nesprin-1/2 complexes connect centrosome to the nucleus during neurogenesis and neuronal migration in mice. *Neuron* 64(2), 173-187 (2009).

45. Madej-Pilarczyk A, Kotruchow K, Kabzinska D, Cegielska J, Kochanski A, Hausmanowa-Petrusewicz I. Emery-Dreifuss muscular dystrophy type 2 associated (?) with mild peripheral polyneuropathy. *Folia Neuropathol* 53(3), 270-274 (2015).

46. Benedetti S, Bertini E, Iannaccone S *et al*. Dominant LMNA mutations can cause combined muscular dystrophy and peripheral neuropathy. *J Neurol Neurosurg Psychiatry* 76(7), 1019-1021 (2005).

47. Duparc A, Cintas P, Somody E *et al*. A cardio-neurological form of laminopathy: dilated cardiomyopathy with permanent partial atrial standstill and axonal neuropathy. *Pacing Clin Electrophysiol* 32(3), 410-415 (2009).

48. Goizet C, Yaou RB, Demay L *et al*. A new mutation of the lamin A/C gene leading to autosomal dominant axonal neuropathy, muscular dystrophy, cardiac disease, and leuconychia. *J Med Genet* 41(3), e29 (2004).

49. Clinvar. https://www.ncbi.nlm.nih.gov/clinvar/RCV000366551/. (Accessed: 01-09-2017).

50. Bierne H, Tham TN, Batsche E *et al*. Human BAHD1 promotes heterochromatic gene silencing. *Proc Natl Acad Sci U S A* 106(33), 13826-13831 (2009).

51. Lakisic G, Lebreton A, Pourpre R *et al*. Role of the BAHD1 Chromatin-Repressive Complex in Placental Development and Regulation of Steroid Metabolism. *PLoS Genet* 12(3), e1005898 (2016).

52. Zhu H, Wan X, Li J *et al*. Computational Prediction and Validation of BAHD1 as a Novel Molecule for Ulcerative Colitis. *Sci Rep* 5 12227 (2015).

53. Libertini E, Lebreton A, Lakisic G *et al*. Overexpression of the Heterochromatinization Factor BAHD1 in HEK293 Cells Differentially Reshapes the DNA Methylome on Autosomes and X Chromosome. *Front Genet* 6 339 (2015).

54. Sun Z, Wu Y, Ordog T *et al*. Aberrant signature methylome by DNMT1 hot spot mutation in hereditary sensory and autonomic neuropathy 1E. *Epigenetics* 9(8), 1184-1193 (2014).

55. Gtex-Portal. The Genotype-Tissue Expression (GTEx) Project.

56. Moriyama B, Henning SA, Leung J *et al*. Adverse interactions between antifungal azoles and vincristine: review and analysis of cases. *Mycoses* 55(4), 290-297 (2012).

6.13. Reference Annotations

***Reference 8:** Kerckhove N, Collin A, Conde S, Chaleteix C, Pezet D, Balayssac D. Long-Term Effects, Pathophysiological Mechanisms, and Risk Factors of Chemotherapy-Induced Peripheral Neuropathies: A Comprehensive Literature Review. Front Pharmacol. 2017;8:86. **This reference is of interest because it contains essential and comprehensive information on VIPN.**

****Reference 4:** Ceppi F, Langlois-Pelletier C, Gagne V, Rousseau J, Ciolino C, De Lorenzo S, et al. Polymorphisms of the vincristine pathway and response to treatment in children with childhood acute lymphoblastic leukemia. Pharmacogenomics. 2014;15(8):1105-16. **This reference is of considerable interest because it reports a previous work on VIPN done in the same discovery cohort.**

***Reference 5:** Park SB, Goldstein D, Krishnan AV, Lin CS, Friedlander ML, Cassidy J, et al. Chemotherapy-induced peripheral neurotoxicity: a critical analysis. CA Cancer J Clin. 2013;63(6):419-37. **This reference is of interest because it contains essential and comprehensive information on VIPN.**

***Reference 21:** Diouf B, Crews KR, Lew G, Pei D, Cheng C, Bao J, et al. Association of an inherited genetic variant with vincristine-related peripheral neuropathy in children with acute lymphoblastic leukemia. JAMA. 2015;313(8):815-23. **This reference is of interest**

because it reports a genome-wide association study of VIPN in patients from a prospective clinical trials for childhood ALL.

****Reference 29:** Abaji R, Gagne V, Xu CJ, Spinella JF, Ceppi F, Laverdiere C, et al. Whole-exome sequencing identified genetic risk factors for asparaginase-related complications in childhood ALL patients. Oncotarget. 2017;8(27):43752-67. **This reference is of considerable interest because it reports a previous work using the same methodology and study design.**

****Reference 35:** De Jager PL, Chibnik LB, Cui J, Reischl J, Lehr S, Simon KC, et al. Integrating genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility. Lancet neurology. 2009;8(12):1111-9. **This reference is of considerable interest because reports the results of a study using wGRS method for building the genetic combined-effect model.**

6.14. Tables

Table 1. Characteristics of the discovery and the replication cohorts.

<table>
<tr><th colspan="2">Cohort Characteristics</th><th colspan="2">EWAS</th><th>QcALL †</th><th>AIEOP</th></tr>
<tr><th colspan="2">Total Included</th><th colspan="2">179</th><th>237</th><th>405</th></tr>
<tr><td rowspan="2">Sex</td><td>Female</td><td colspan="2">78 (43.6%)</td><td>107 (45.1%)</td><td>190 (46.9%)</td></tr>
<tr><td>Male</td><td colspan="2">101 (56.4%)</td><td>130 (54.9%)</td><td>215 (53.1%)</td></tr>
<tr><td rowspan="2">Age</td><td>< 10 years</td><td colspan="2">148 (82.7%)</td><td>196 (82.7%)</td><td>337 (83.2%)</td></tr>
<tr><td>≥ 10 years</td><td colspan="2">31 (17.3%)</td><td>41 (17.3%)</td><td>68 (16.8%)</td></tr>
<tr><td rowspan="3">Risk</td><td>Standard</td><td colspan="2">93 (52%)</td><td>127 (53.6%)</td><td>116 (28.6%)</td></tr>
<tr><td>Intermediate</td><td colspan="2">-</td><td>-</td><td>253 (62.5%)</td></tr>
<tr><td>High</td><td colspan="2">86 (48%)</td><td>110 (46.4%)</td><td>36 (8.9%)</td></tr>
<tr><td rowspan="4">Protocol</td><td rowspan="4"></td><td>DFCI 00-01</td><td>64 (35.7%)</td><td>89 (37.6%)</td><td rowspan="4">405 (100%) BFM-AIEOP-2000</td></tr>
<tr><td>DFCI 95-01</td><td>91 (50.8%)</td><td>100 (42.2%)</td></tr>
<tr><td>DFCI 91-01</td><td>18 (10.1%)</td><td>35 (14.8%)</td></tr>
<tr><td>DFCI 87-01</td><td>6 (3.4%)</td><td>13 (5.5%)</td></tr>
<tr><td rowspan="2">High Grade Peripheral Neuropathy</td><td>-</td><td colspan="2">155 (86.6%)</td><td>202 (85.2%)</td><td>392 (96.8%)</td></tr>
<tr><td>+</td><td colspan="2">24 (13.4%)</td><td>35 (14.8%)</td><td>13 (3.2%)</td></tr>
</table>

EWAS, Exome Wide Association Study cohort; QcALL, Quebec Childhood ALL cohort; DFCI, Dana-Farber Cancer Institute ALL Consortium cohort; AIEOP, Associazione Italiana di Ematologia Oncologia Pediatrica; BFM, Berlin-Frankfurt-Münster; VIPN, Vincristine-induced peripheral neuropathy.

† Whole-exome sequencing data were available for 179 patients of QcALL cohort for whom EWAS analysis was performed

Table 2. Top-ranking signals from the exome-wide association study confirmed by genotyping.

SNP	Univariate Analysis									Multivariate Analysis	
	Neurotoxicity		OR (95%-CI)	P (Fisher)	Model	Neurotoxicity		OR (95%-CI)	P	OR (95%-CI)	P
	+	-				+	-				
SYNE2_rs2781377: G > A											
GG	26 (74.3%)	176 (87.6%)	1	*Ref.*				1	*Ref.*	2.7 (1.2-6.0)	0.02
GA	7 (20%)	24 (11.9%)	2 (0.8-5.0)	0.2				2.5 (1.2-5.2)[a]	0.01		
AA	2 (5.7%)	1 (0.5%)	13.5 (1.2-154)	0.05							
MRPL47_rs10513762: C > T											
CC	25 (71.4%)	180 (89.1%)	1	*Ref.*	CC	25 (71.4%)	180 (89.1%)	1	*Ref.*	3.9 (1.5-10)	0.004
CT	10 (28.6%)	21 (10.4%)	3.4 (1.4-8.1)	0.006	CT+TT	10 (28.6%)	22 (10.9%)	3.3 (1.4-7.7)[d]	0.01		
TT	0 (0%)	1 (0.5%)	-	-							
BAHD1_rs3803357: C > A											
CC	20 (57.2%)	64 (31.8%)	1	*Ref.*	CC	20 (57.1%)	64 (31.8%)	1	*Ref.*	0.3 (0.2-0.8)	0.009
CA	11 (31.4%)	96 (47.8%)	0.37 (0.2-0.8)	0.01	CA+AA	15 (42.9%)	137 (68.2%)	0.35 (0.2-0.7)[d]	0.007		
AA	4 (11.4%)	41 (20.4%)	0.3 (0.1-1)	0.06							

The SNPs are presented as a change from major to minor alleles. OR, odds ratio; CI, confidence interval. Analysis in both co-dominant model and a model that best fits the data are presented. The final univariate models are either additive (a) or dominant (d). The regression model in the multivariate analysis included genotypes coded according to the indicated model, and as covariates, age, risk and DFCI protocol.

6.15. Figure Legends

Figure 1. Confirmatory step following the exome-wide association study.

Each triangle contains all the SNPs that are inside of it, including the ones in the smaller triangles. The largest triangle represents the top-ranking signals (N=21) associated with high-grade VIPN in the exome-wide association study. The middle triangle represents variants with minor allele frequencies > 5% in the QcALL cohort (N=16) which were subjected to confirmation through genotyping. The top triangle represents significant associations with high-grade VIPN (N=3) retained for analysis in replication cohort.

* rs35432946 in the *TRIM4* gene was eliminated from further analysis since in linkage disequilibrium with rs33998596 also in the TRIM4 gene.

** rs9323693 in the *OR11H6* gene codes for an olfactory receptor and was not considered for further analysis.

Figure 2. Performance of the wGRS based combined-effect model in the discovery cohort and classification efficiency in both the discovery and replication cohorts.

A) Performance of the model (depicted by the area under the curve, AUC, of receiver-operating characteristics, ROC curves) in discriminating between patients with and without high-grade VIPN based on their genetic profiles. Weighted genetic risk scores (wGRS) were used to estimates the risk of toxicity and were derived from the In(OR) obtained in the logistic regression model for each of the three SNPs identified in the EWAS. **B)** Frequency of patients with and without VIPN (represented by bars) between the two risk groups.

Number of patients per group is provided on the top of each bar. Patients were divided into two equal groups which were assigned a low risk or a high risk status based on the distribution of wGRS values relative to the median. Results are displayed for the discovery (QcALL) and replication (AIEOP) cohorts.

Figure 3. Performance of the comprehensive combined-effect model in predicting the risk of VIPN and classifying patients into risk groups.

A) Distribution of patients with VIPN across risk groups in discovery and replication cohort; **B)** The discrimination capacity of the model in discovery and replication group.

Patients in the QcALL cohort were assigned to risk groups based on their individual wGRS calculated from the genetic model combining the 4 loci associated with the risk of VIPN in this cohort. Three groups of risk were identified: Low (wGRS < 0), Intermediate (0 ≥ wGRS ≤ 0.474) and High (wGRS > 0.474). Patients in the AIEOP replication cohort were assigned to the same risk groups based on a predicted risk score derived using the same algorithm for calculating wGRS in the discovery cohort. In **A)** distribution of patients with high-grade VIPN across risk groups is represented by bars. The number of cases in each category is indicated on the top of each column. Mean wGRS is provided for each risk group as a white circle ± 2SE. The p-value of the association across groups is provided at the bottom of the graph while the difference between the highest and the lowest risk groups is displayed on top of the graph together with the odds-ratio (OR) and the 95% confidence-intervals (CI) in brackets. In **B)** the discrimination capacity of the model was assessed using the AUC of ROC curves obtained after classifying patients with and without VIPN into risk groups.

6.16. Figures

Figure 1.

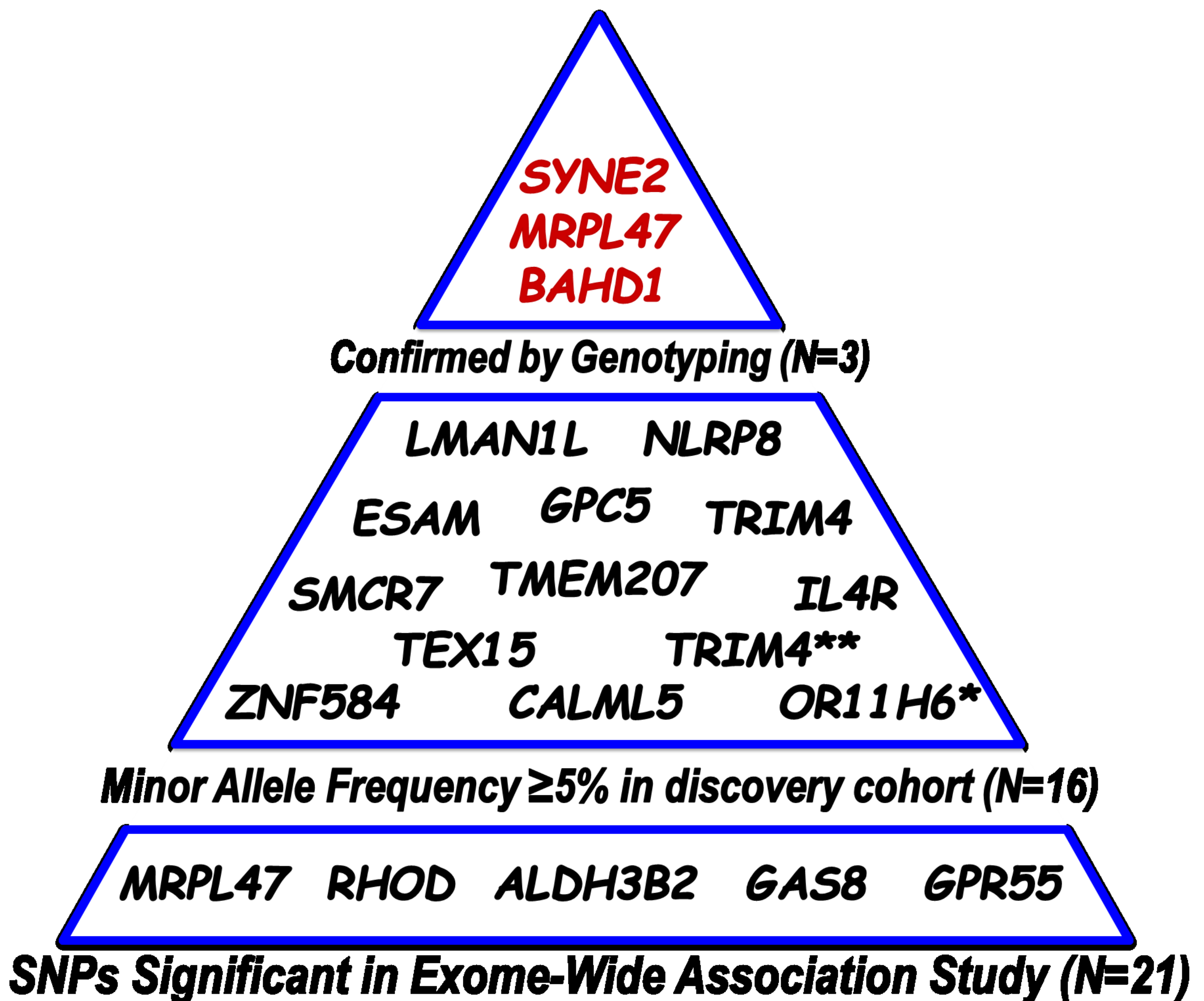


Figure 1. Confirmatory step following the exome-wide association study.

Figure 2.

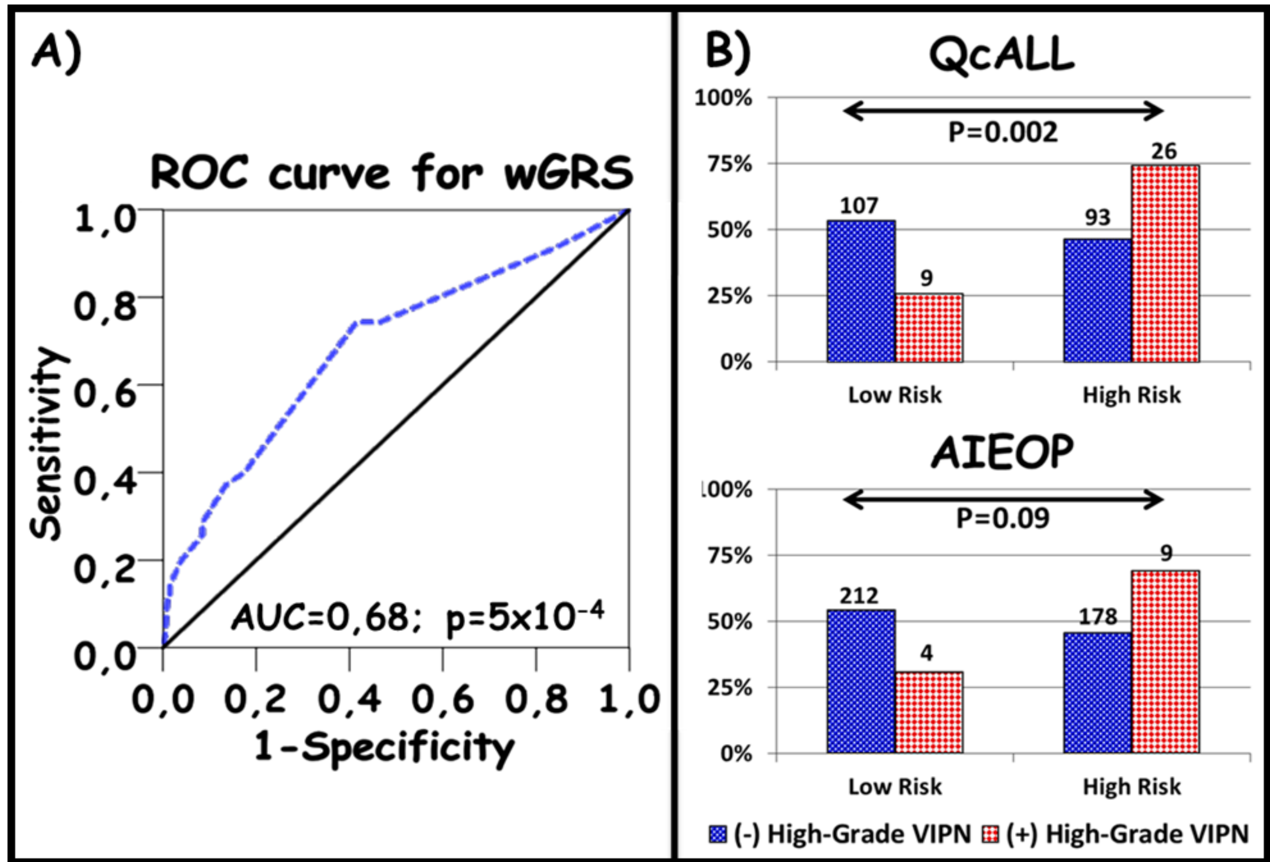


Figure 2. Performance of the wGRS based combined-effect model in the discovery cohort and classification efficiency in both the discovery and replication cohorts.

Figure 3.

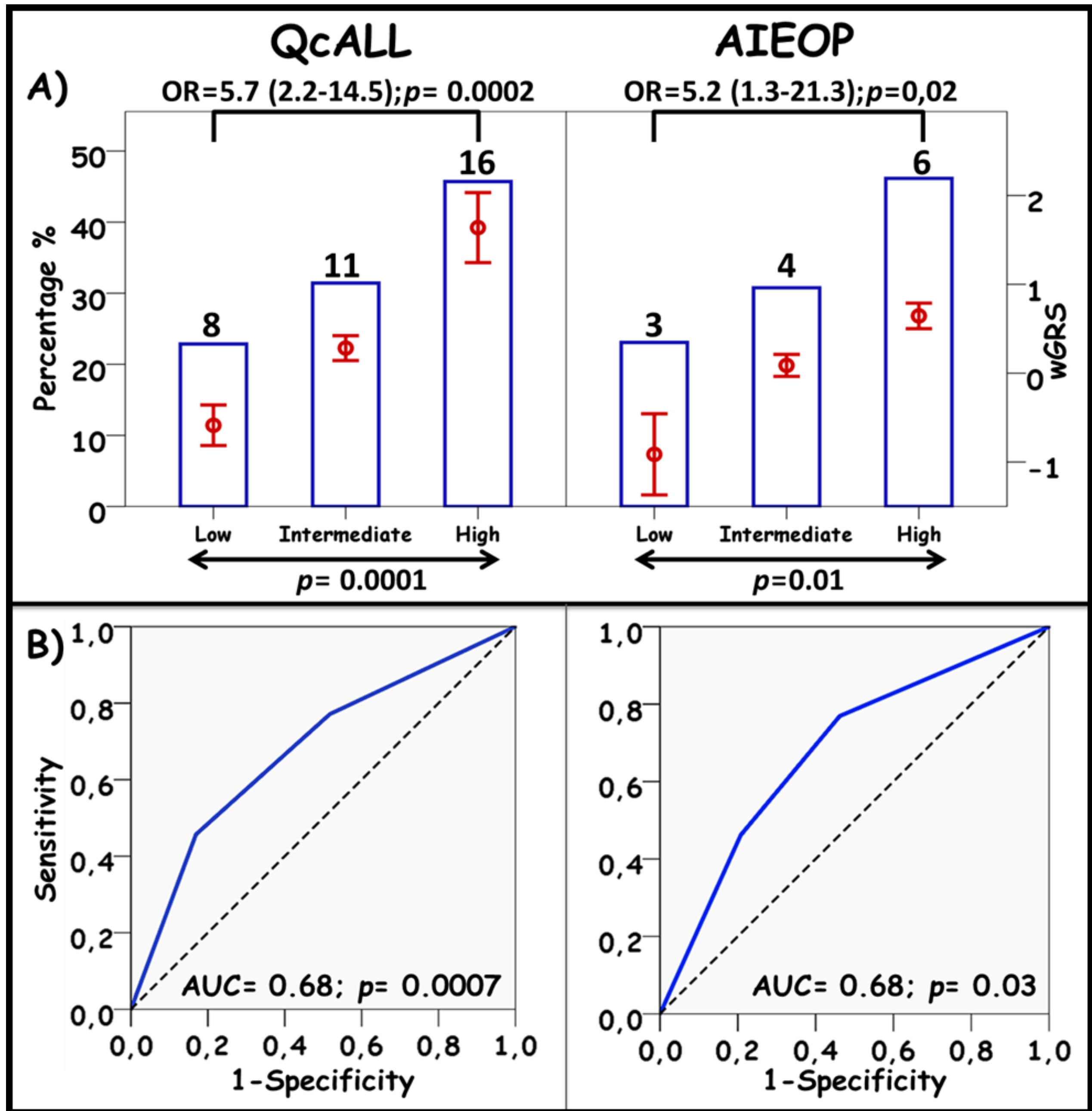


Figure 3. Performance of the comprehensive combined-effect model in predicting the risk of VIPN and classifying patients into risk groups.

6.17. Supplemental Material

Supplemental Table S1. Association of WES data with vincristine-induced peripheral neuropathy.

Gene_SNP	Minor Allele Frequency %	P-value Allelic Association (Sequencing Data)	FDR (Q-Value)	P-value Association by Genotype (Sequencing Data)	P-value Association by Genotype (Genotyping Data)
RHOD_rs4930409: T > C	1%†	2.1E-07	0.05%	1.8E-01	-
GAS8_rs17178299: G > A	1%†	7.7E-06	0.91%	1.3E-04	-
ALDH3B2_rs17856219: G > A	1%†	1.8E-03	8.94%	2.0E-02	-
GPR55_rs3749073: C > A	3%†	2.8E-03	12.59%	9.3E-03	-
MRPL47_rs2339844: A > C	4%†	1.6E-03	7.99%	4.6E-03	-
OR11H6_rs9323693: C > G	7%	4.2E-04	2.58%	8.0E-04	§
TRIM4_rs35432946: G > A	7%	3.6E-03	14.99%	4.2E-03‡	‡
TRIM4_rs33998596: G > C	**5%**	**8.5E-04**	**4.75%**	**2.4E-03**	**3.4E-01**
TEX15_rs61732457: A > C	**5%**	**3.0E-03**	**13.06%**	**6.3E-03**	**1.2E-01**
SMCR7_rs12603700: G > A	**5%**	**3.5E-03**	**14.87%**	**1.4E-02**	**1.1E-01**
NLRP8_rs7259764: A > G	**6%**	**4.5E-04**	**2.71%**	**1.0E-03**	**4.9E-01**
IL4R_rs1805012: T > C	**6%**	**8.7E-04**	**4.75%**	**1.8E-03**	**4.9E-01**
ESAM_rs12792040: G > A	**7%**	**1.5E-03**	**7.75%**	**2.5E-03**	**5.9E-01**
TMEM207_rs35161724: G > C	**9%**	**1.8E-03**	**9.02%**	**8.9E-03**	**8.6E-02**
GPC5_rs553717: C > T	**11%**	**2.9E-03**	**13.03%**	**3.6E-04**	**4.1E-01**
LMAN1L_rs79217743: G > T	**14%**	**5.0E-04**	**2.94%**	**3.2E-03**	**3.8E-01**
ZNF584_rs11668789: C > T	**16%**	**1.7E-04**	**1.71%**	**1.1E-04**	**1.0E+00**
CALML5_rs11546426: T > C	**19%**	**2.4E-03**	**11.59%**	**1.7E-02**	**1.0E+00**
MRPL47_rs10513762: C > T	**7%**	**2.3E-04**	**1.71%**	**4.7E-04**	**1.2E-02**
SYNE2_rs2781377: G > A	**8%**	**2.7E-03**	**12.59%**	**1.4E-02**	***1.1E-02***
BAHD1_rs3803357: C > A	**37%**	**8.6E-04**	**4.75%**	**7.6E-03**	**7.1E-03**

The SNPs selected for validation through genotyping are highlighted in grey while those that remained significant are depicted in white font and highlighted in black.

P-values reflect the difference across genotype groups regardless of genetic model. Further analysis in accordance to appropriate models is presented in Table 2.

† Variants with minor allele frequencies lower than 5% in the QcALL cohort were removed from further analysis.

‡ rs35432946 was not considered further since in linkage disequilibrium with rs33998596.

§ rs9323693 in the *OR11H6* gene codes for an olfactory receptor and was not considered for further analysis.

Supplemental Table S2. Association of the risk alleles with the reduction in the VCR dose administered.

Group	N	Mean Rank	Average Percentage of Cumulative VCR Dose Administered	P-Value
SYNE2_rs2781377: G > A				
GG	202	120.8	96.3%	0.02
GA	31	109.2	93.3%	
AA	3	58.3	84.6%	
MRPL47_rs10513762: C > T				
CC	205	122.1	96.6%	0.004
CT+TT	32	99	90.6%	
BAHD1_rs3803357: C > A				
CC	84	108.3	93.8%	0.006
CA+AA	152	124.1	96.8%	

Each variant was tested for the association with the percentage of the cumulative VCR dose administered by applying the non-parametric Kruskal-Wallis. Results are shown following the same model (additive or dominant) used in the manuscript file.

Supplemental Table S3. Association of the risk alleles with the number of episodes of high-grade VIPN.

Group/ N of patients (%)	N of episodes of high-grade VIPN			P-Value
	0	1	≥ 2	
SYNE2_rs2781377: G > A				
GG	176 (87.1%)	6 (3%)	20 (9.9)	0.09
GA	24 (77.4%)	2 (6.5%)	5 (16.1%)	
AA	1 (33.3%)	2 (66.7%)	0 (0%)	
MRPL47_rs10513762: C > T				
CC	180 (87.8%)	8 (3.9%)	17 (8.3%)	0.002
CT+TT	22 (68.8%)	2 (6.3%)	8 (25.0%)	
BAHD1_rs3803357: C > A				
CC	64 (76.2%)	5 (6%)	15 (17.9%)	0.003
CA+AA	137 (90.1%)	5 (3.3%)	10 (6.6%)	

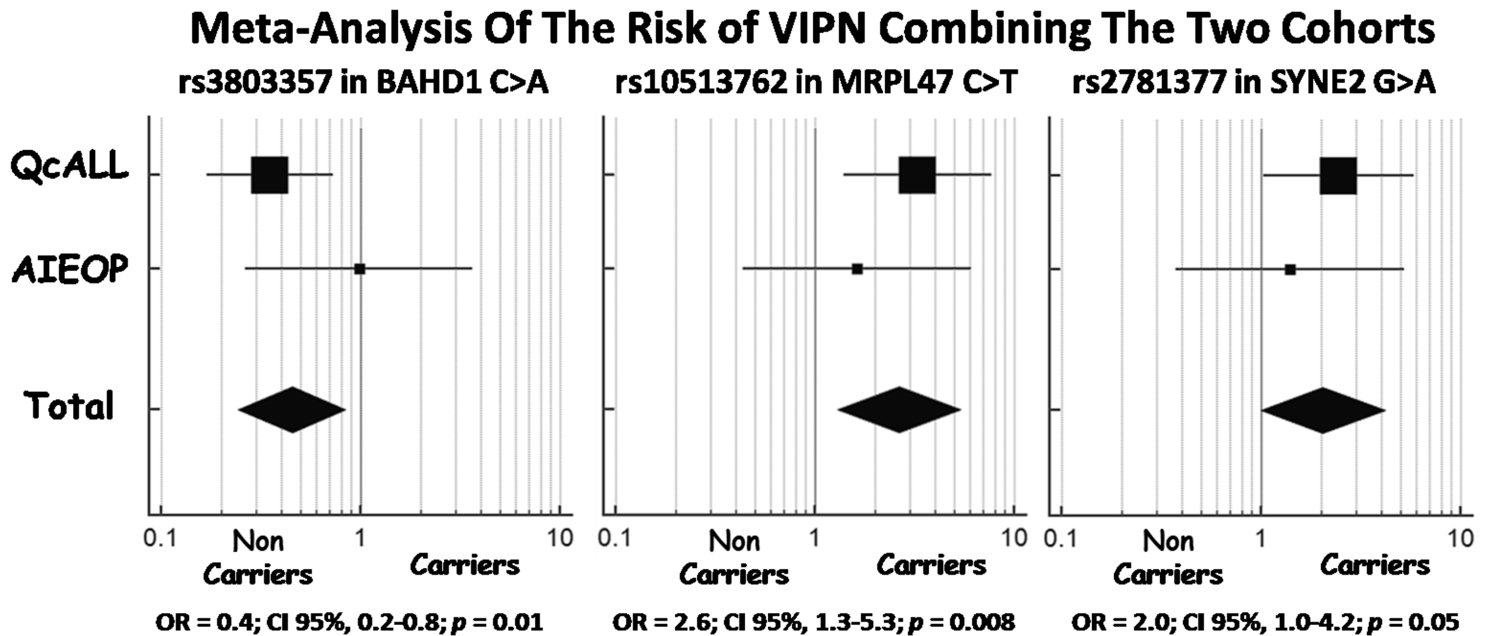
P-values were estimated via the Fisher`s exact test and results are shown following the same model (additive or dominant) used in the manuscript file.

Supplemental Table S4. Analysis of the association between the risk alleles and lower-grade (I/II), higher-grade (III/IV) and all-grades (I-IV) VIPN.

Gene_SNP	P-Value			
	No VIPN vs.			(No VIPN + Grade I/II) vs. Grade III/IV
	Grade I/II	Grade III/IV	All Grades (I-IV)	
SYNE2_rs2781377: G > A	0.09	0.01	0.01	0.03
MRPL47_rs10513762: C > T	0.2	0.01	0.02	0.02
BAHD1_rs3803357: C > A	0.6	0.007	0.06	0.008

Results are shown following the same model (additive or dominant) used in the manuscript file. Columns 2, 3 and 4 provide the p-values of the association between SNPs and VIPN by comparing patients with no toxicity to those with lower-grade, higher grade and all grades toxicity, respectively. Column 5 shows the p-values comparing patients with higher-grade toxicity to patients with either lower-grade toxicity or no toxicity at all.

Supplemental Figures:

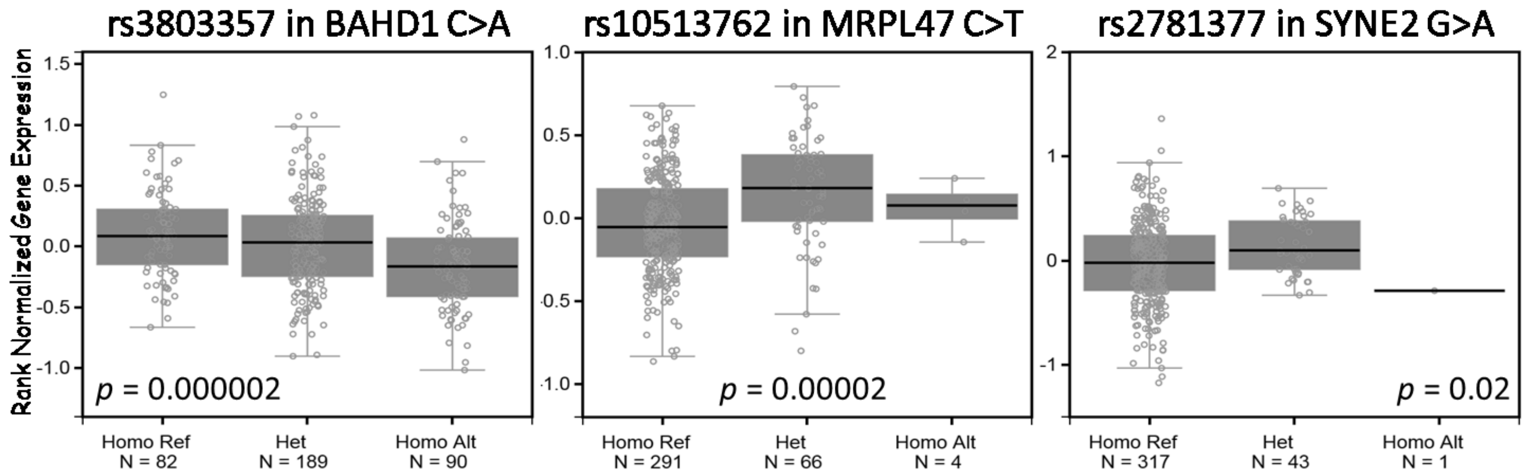


Supplemental Figure S1

Supplemental Figure S1. Meta-analysis of the top-ranking associations combining both cohorts.

Each plot represents the association of a polymorphism with VIPN (reported at the top of the graph) as tested in the discovery cohort (QcALL), the validation cohort (AIEOP) and the cohort combining them both (Total). Odd-Ratios (OR) comparing carriers to non-carriers, along with the 95% confidance intervals (CI 95%) and the p-values of the associations are provided at the bottom of each graph. The Meta-Analysis was performed using Mantel-Haenszel method implemented in MedCalc software and assuming a a fixed-effect model.

Supplemental Figure S2



Supplemental Figure S2. Association between the top-ranking hits and the expression of their respective genes based on the genotype.

Each plot represents an association between a polymorphism and the expression of its gene as indicated on the top of the respective graph. P-value of the associations across the genotype groups is provided inside of the graph. Homo Ref refers to the major allele and Homo Alt refers to the minor allele of each polymorphism. Plots are downloaded from www.gtexportal.org/home/ and edited to incorporate essential data.

Supplemental methods

Whole Exome Sequencing (WES)

Whole exomes from peripheral blood or bone marrow samples obtained after remission of QcALL cohort patients (1, 2) were captured in solution with Agilent's SureSelect Human All Exon 50Mb kits, and sequenced on the Life Technologies SOLiD System (patients mean coverage ~35X). Reads were aligned to the hg19 reference genome using SOLiD LifeScope software. PCR duplicates were removed using Picard.(3) Base quality score recalibration was performed using the Genome Analysis ToolKit (GATK)(4) and QC Failure reads were removed. Cleaned BAM files were used to create pileup files using SAMtool.(5)

Germline variants have been called using SNooPer(6) a variant caller based on a machine learning algorithm that uses a subset of variant positions from the sequencing output for which the class is known, either actual variation or sequencing error, to train a data-specific model.

The annotation of the identified germline variants was performed using ANNOVAR.(7) Only missense, nonsense and variations in splicing sites were conserved. The predicted effect of missense variants on the protein function was assessed in silico using Sift (≤0.05) (8) and Polyphen2 (≥0.5).(9) Minor allele frequencies (MAF) higher than 5% were derived from the 1000 Genomes (European population) (10) and the NHLBI GO Exome Sequencing Project (European population, ESP).(11)

Sequencing data and clinical information on higher grade neurotoxicity data was available for 176 patients. Fisher's Exact test (allelic association) and Cochran-Armitage trend test, implemented in PLINK(12), were used for an association study. Adjustment for multiple testing was performed by bootstrap false discovery rate (FDR)(13) method; the SNPs retained for further analysis had FDR lower than 15%.

Validation of top-ranking EWAS signals by Genotyping

Genotyping of top ranking EWAS signals was either performed at the McGill University and Génome Québec Innovation Centre through Sequenom genotyping platform or by allele-specific oligonucleotides (ASOs) hybridization as described elsewhere.(14)

Supplemental References

1. Healy J, Richer C, Bourgey M, Kritikou EA, Sinnett D. Replication analysis confirms the association of ARID5B with childhood B-cell acute lymphoblastic leukemia. Haematologica. 2010 Sep;95(9):1608-11.

2. Baccichet A, Qualman SK, Sinnett D. Allelic loss in childhood acute lymphoblastic leukemia. Leuk Res. 1997 Sep;21(9):817-23.

3. BroadInstitiute. GITHUB-PICARD. [Website] [cited; Available from: http://broadinstitute.github.io/picard.

4. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010 Sep;20(9):1297-303.

5. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078-9.

6. Spinella JF, Mehanna P, Vidal R, Saillour V, Cassart P, Richer C, et al. SNooPer: a machine learning-based method for somatic variant identification from low-pass next-generation sequencing. BMC Genomics. 2016 Nov 14;17(1):912.

7. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010 Sep;38(16):e164.

8. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4(7):1073-81.

9. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet. 2013 Jan;Chapter 7:Unit7 20.

10. Genomes Project C, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, et al. A map of human genome variation from population-scale sequencing. Nature. 2010 Oct 28;467(7319):1061-73.

11. Server EV. [Website] [cited; Available from: http://evs.gs.washington.edu/EVS/

12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 Sep;81(3):559-75.

13. Storey JD, Taylor JE, Siegmund D. Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2004;66(1):187-205.

14. Labuda D, Krajinovic M, Richer C, Skoll A, Sinnett H, Yotova V, et al. Rapid detection of CYP1A1, CYP2D6, and NAT variants by multiplex polymerase chain reaction and allele-specific oligonucleotide assay. Anal Biochem. 1999 Nov 1;275(1):84-92.

Section D

Chapter 7

General Discussion

In this final chapter, I summarize the most important ideas to be derived from the different sections presented throughout this thesis. I discuss the relevant information currently available in the literature to support the validity of the findings and present a hypothetical model of the possible synergistic effect between the two most supported genetic polymorphisms in modulating the risk of pancreatitis. I also discuss some interesting additional associations from unpublished data which can further support the importance of the concerned variants in modulating the treatment response. Moreover, I address the limitations that could have influenced the results and propose experiments for future studies that can provide more information on the identified associations. Finally, I conclude by summarizing the most important points to be retained from the work presented in this thesis.

7. General Discussion

7.1. Discussion of Section A

In acute lymphoblastic leukemia (ALL) treatment protocols, as with most cancers, chemotherapy regiments rely on the administration of multi-drug combinations to potentiate the anti-cancer effect and to reduce the risk of resistance. Nonetheless, treatment-related toxicity can endanger patients' lives and is classified among the principal causes of treatment interruption or cessation in childhood cancers. Toxic-effects of anti-leukemia drugs can range from mild and transient organ damage to more serious, life-threatening and permanent outcomes impeding the survivors' ability to lead a normal adult life. Accordingly, and given the gravity of the consequence of treatment failure, which is frequently life-threatening, having the ability to predict the specific response to a particular treatment prior to its administration can be highly valuable and lies in the core interest of personalized oncology. This would empower clinicians to better calculate the overall efficacy/toxicity ratio of a given treatment, which remains a major challenge in the vulnerable pediatric population. The risk of relapse and treatment toxicity can be modulated by multiple factors and differences in genetic composition between patients have recently driven considerable attention.[1,2]

Pharmacogentics (PGx) is the study of how the variability in the genetic component between individuals can influence the observed variability in treatment efficacy and toxicity. In that sense, genetic polymorphisms in genes that can affect the pharmacokinetics or the pharmacodynamics of chemotherapeutic agents (coding for drug-metabolizing enzymes, transporters, or drug targets.) have naturally been the first targets to be

explored.[3] Indeed, success stories in which pharmacogenetics discoveries have restructured the medical practice are numerous and one classical example is the genotyping of TPMT gene to guide the dosing of 6-mercaptopurine which is almost considered mandatory in most recent ALL treatment guideline.[4] The influence of polymorphisms of this gene on treatment outcome is well-documented in scientific and clinical literature and was replicated in many studies and clinical trials. However, the adoption of TPMT pharmacogenetics testing was not a straight forward process. It rather underwent a long path of scientific scrutiny and clinical validation and overcame many pitfalls before it eventually evolved into an important pharmacogene through cumulative knowledge that is worth decades of experience. Indeed, recent prospective studies have demonstrated the importance of pre-emptive *TPMT* genetic screening and subsequent dose adjustment in mitigating the hematotoxicity associated with thiopurine administration such as myelosuppression, among others, while maintaining treatment efficacy and favorable long-term outcomes. This allowed the development of dosing recommendations and treatment strategies to optimize and individualize the prescribing of thiopurines based on the pharmacogenetics of *TPMT*.

However, the effect of most genes on treatment response phenotypes remains largely unknown and a lot of times, unexpected. Since it is rather long, or even unrealistic, to individually test all associations of the genetic alterations in the entire genome against the whole spectrum of possible response phenotypes to known drugs, the implementation of unbiased association techniques in large scale exome-wide or genome-wide association studies can bring forward interesting genes and thus offer new insights on their implication in the mechanisms of pathogenesis and drug response.

7.2. Discussion of Section B

Following an initial exome-wide association study that employed whole-exome sequencing data available from the Sainte-Justine Hospital and Research Centre, we identified a list of top-ranking SNPs associated with adverse drug reactions during childhood ALL treatment. In order to control for the quality of the sequencing data, we sought to confirm the identified top-ranking SNPs by genotyping. Our confirmatory analysis was able to identify 12 SNPs associated with major adverse drug events attributable to the administration of asparaginase (ASNase); of which, 3 were associated with allergies, 3 with pancreatitis and 6 with thrombosis. Interestingly, rs3809849 in the *MYBBP1A* gene was associated with allergy, pancreatitis, thrombosis, event-free survival (EFS) and overall survival (OS); while each of rs11556218 in *IL16* and rs34708521 in *SPEF2* were associated with both thrombosis and pancreatitis. We also identified strong additive effects of harbouring multiple risk alleles on the possibility of developing the respective side-effect.

In order to assess the reproducibility of our findings, we carried a validation step in which we tested the identified associations in an independent validation cohort that, similar to QcALL, followed the Dana Farber Cancer Institute (DFCI) treatment protocols and we were able to replicate the results for 3 of the SNPs associated with pancreatitis. Moreover, we used our results to derive a prediction model, which was able to efficiently predict the risk of developing pancreatitis based on the weighted genetic risk score of individual patient. Of note, the prediction efficiency of this model was confirmed in the validation cohort.

However, since these genes, and their polymorphism, were not previously reported to be involved with the studied toxicities, the possibility of a random association cannot be ruled out (and their causality cannot be confirmed) unless functional data proving their involvement at a molecular level is provided. Accordingly, in follow-up analyses, we aimed to confirm the active role of the *MYBBP1A* and *IL16* genes in modulating the risk of ASNase complications through cell-based functional analyses. We also developed hypotheses that could explain the link between the identified polymorphisms, the gene function and the associated risks.

7.2.1. MYBBP1A & Pancreatitis

The expression of this transcriptional regulator gene was previously found to be enriched in endoderm during specific stages of endocrine pancreas development.[5] We selected PANC1 pancreatic cancer cell line to produce MYBBP1A knockout cells and studied changes in their viability, drug sensitivity and morphology. The deletion of the gene was associated with a significant reduction in cell viability (represented by a slower proliferation rate and a reduced clonogenic potential) as well as a selective increase in sensitivity to ASNase.

MYBBP1A knockout cells also exhibited changes in their morphology and marker expression profile, such as Vimentin and ZEB, suggestive of an Epithelial-to Mesenchymal transition (EMT). This potential involvement of *MYBBP1A* in regulating EMT process is quite intriguing as many findings in the literature support the hypothesis that abolishing

MYBBP1A expression can induce EMT. For example, the MYBBP1A protein was shown to have an inverse regulatory relationship with AKT phosphorylation at the (Ser473) residue [6] while AKT regulates various cellular mechanisms including EMT mediated by NF-kB. Indeed, NF-kB acts as a regulator of other mediators of the EMT process (e.g. Snail & E-cadherin) and its activation was demonstrated to trigger a signaling cascade leading both to acute pancreatitis [7] and EMT.[8] This observation merits further investigation at the molecular level. Moreover, in human carcinomas, the activation of AKT kinase is considered to be a frequent and recurrent event [6,8] and a histological staining pattern that shows a low expression of MYBBP1A and a high expression of pAKT(Ser473) was correlated with shorter progression-free and overall survival in patients with primary oropharyngeal squamous cell carcinoma, suggesting the use of this staining pattern as an independent prognosticator for high risk of treatment failure.[6]

An aberrant *MYBBP1A* gene expression resulting in significantly reduced MYBBP1A protein levels was reported in recurrent head and neck squamous cell carcinoma compared to samples from patients with primary tumors.[9] *MYBBP1A* was also suggested to have a critical role in the regulation of senescence under genotoxic stress, since silencing its expression was associated with a significantly increased relative abundance of senescent cells after DNA damage, although not sufficient to induce senescence on its own.[6] Indeed, its downregulation was shown to result in abolishing local DNA methylation and histone marks associated with gene silencing, consequently leading to elevated ribosomal RNA expression as a resulted of altered promoter occupancy of various epigenetic factors.[10] Of note, DNA damage caused by treatment of tumor cells with etoposide was associated with a

significant decrease in MYBBP1A protein levels [6] and cases of acute pancreatitis induced by etoposide-containing drug combinations have been reported;[11,12] thus arguing for the involvement of this gene with pancreatitis.

The down-regulation of *MYBBP1A* was associated with a reduction in the proliferation capacity of human HeLa cells, where it can also promote apoptosis, cell-cycle arrest at G2/M, or delayed and anomalous mitosis.[13] We have provided evidence that these changes could possibly be a result of a specific cell cycle blockage at the S-phase, along with an induction of apoptosis in the KO cells. Moreover, we showed that combining the gene deletion with ASNase exposure leads to an additional cell cycle arrest at G0/G1, as well as inducing stronger apoptotic reactions and provoking cellular necrosis; therefore providing a plausible mechanistic understanding of how *MYBBP1A* gene deletion modulates PANC1 cells sensitivity to ASNase treatment and its observed impact on their clonogenic potential.

Moreover, the MYBBP1A protein was shown to enhance the activity of p53 through promoting its tetramerization followed by its acetylation, a crucial process for p53 to exert its biological activity since it prevents MDM2-dependent degradation;[14-16] thereby determining cell fate between cell cycle arrest and apoptosis.[14] Furthermore, it was shown to play a role in tumor prevention through p53 activation during anoikis, defined as detachment-induced apoptosis, and it is involved in suppressing tumorigenesis and colony formation of breast cancer cells.[17] Hence, investigation of the relationship between *MYBBP1A* and p53 in pancreatic cells upon ASNase challenge could provide a valuable mechanistic understanding of the role of this gene in the development of ASNase induced acute pancreatitis.

MYBBP1A also acts as a modulator of many transcriptional factors that play a role in development and organogenesis. Notably, it was identified as a regulator of Prep1-Pbx1 transcriptional activity through physical interaction as it competes with Pbx1 for binding to *Prep1*.[18] On one hand, studies demonstrated that *Pbx1* is important for pancreas organogenesis [19] and that *Pbx1*-deficient mice are associated with deficient pancreas development.[20] On the other hand, prep1 deficiency in mouse models was associated with protection from diabetes and increased insulin sensitivity; an effect that was mediated by MYBBP1A protein.[21]

Other possible mechanisms through which *MYBBP1A* might regulate pancreatic response can be extrapolated from known mechanisms involved in increasing the risk of diabetes. A relationship between acute pancreatitis and the risk of diabetes has been previously reported whereby a diagnosis of acute pancreatitis was shown to increase the risk of diabetes by more than two-fold over a period of 5 years.[22] One of these mechanisms is based on the role of *MYBBP1A* gene in regulating the activity of *PPAR-gamma coactivator-1α* (*PGC-1α*); a key regulator of glucose and energy metabolism and other metabolic processes.[23] Notably, its overexpression in mice was associated with a reduction in β-cell mass and size, as well as pancreatic dysfunction resulting in decreased insulin secretion.[24] It was also demonstrated that a reduction in the expression of MYBBP1A protein resulted in hyper-activation of *PGC-1α*, which was associated with an increased sensitivity to insulin, whereas the overexpression of *MYBBP1A* was associated with a reduction in *PGC-1α* expression. However, this effect was noted in a myoblast cell line but not in the liver, which could further support the hypothesis that the role of *MYBBP1A* is cell-type dependent.[21]

Interestingly, on the contrary to the inhibitory role of *MYBBP1A* gene in suppressing all of the transcriptional factors mentioned above, it exerts a stimulatory activity when it comes to the aryl hydrocarbon receptor (AhR).[25] Importantly, this receptor is a ligand-dependent transcription factor that mediates the interactions between pancreatic leukocytes and epithelial cells by regulating the expression of IL-22; an interleukin that modulates the immune response in acute pancreatitis and determines its severity and progression through binding to IL-22RA1 receptor expressed by pancreatic epithelial cells. Of note, AhR inactivation was shown to decrease the levels of pancreatic IL-22 and to worsen acute pancreatitis response while its activation protects from acute pancreatitis by inducing expression of IL-22.[26]

7.2.2. IL16 & Pancreatitis

Interleukin-16 (IL-16) is a pleiotropic cytokine that acts as a lymphocyte chemoattractant and a modulator of the activation of T-cells, as well as monocytes, eosinophils, maturing macrophages and dendritic cells by binding the CD4 receptor. In this manner, it stimulates the secretion of various inflammatory cytokines, including TNF-α, IL1β, IL6, and IL15, and therefore initiates and sustains the inflammatory response.[27,28] It has also been shown to act in concert with IL-2 and/or IL-15 by priming CD4+ T-cells for IL-2 responsiveness, thus promoting their proliferation.[29]

This cytokine is produced by activated CD8+ T cells, B-cells and mast cells as a precursor protein, pro–IL-16,[30] which is then transformed into its bioactive form through

caspase-3 mediated cleavage,[31] and has been implicated in the pathogenesis of various inflammatory diseases, as well as in the development and progression of tumors such as colorectal cancer, osteosarcoma and multiple myeloma.[27,32-34] The serum levels of IL16 were demonstrated to be significantly increased in advanced tumour stages and a worsening outcome in different types of cancer.[30,34] It also plays a key role in autoimmune diseases like asthma, [35,36] allergy [37,38] and acquired immune deficiency syndrome (AIDS).[39]

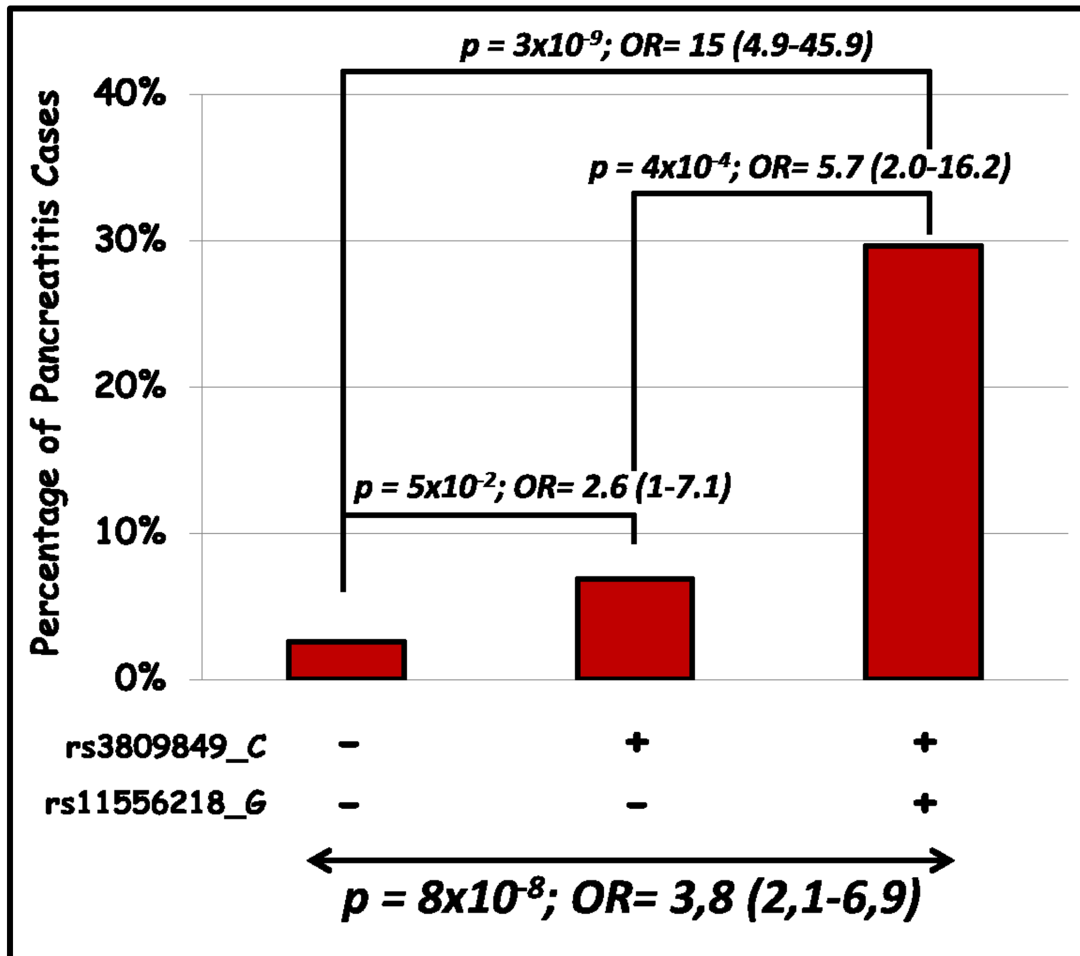
The identified rs11556218 SNP in IL16 gene is of particular interest because this same polymorphism has been previously associated with a wide range of conditions such as endometriosis,[40] sporadic Alzheimer's Disease,[41] emphysema,[42] coronary artery disease, [43] ischemic stroke,[44] systemic lupus erythematous,[45] chronic hepatitis B infection,[46] osteoarthritis,[40,47] overall cancer risk, as well as particular cancer types.[48] While some studies linked carrier-state of the minor allele of this polymorphisms with higher levels of IL16 in the plasma,[33] others found no association.[32,47] One interesting observation though, is the strong eQTL effect of this polymorphism in monocytes as the variant allele was associated with a significant increase in IL16 expression (2.2×10^{-21}).[49]

Studies suggest that IL16 is not expressed in intact β-cells islets, and that it seems to be rather produced by the immune cells upon their infiltration into islets lesions following inflammatory response.[50] It was found to be produced by several types of mononuclear autoimmune cells in islet lesions, consequently promoting the infiltration of additional CD4+ T-cells into the lesion site and exacerbating the inflammatory response. Additionally, elevated IL-16 activity was found to be associated with reduction of β-cells mass.[50]

An intriguing hypothesis that can explain the association of IL16 with pancreatitis stems from the observation that the production of IL16 in the pancreas correlates with T-cells infiltration into the injured pancreatic islets and that the progressive infiltration of these islets by lymphocytes is known to be involved in the mechanism of pancreatic β-cell destruction resulting in diabetes.[51,52] Moreover, the depletion of islet macrophages, [52] or the neutralization of IL-16, [50] were shown to reduce the infiltration of lymphocytes into the islets and consequently protect from autoimmune type 1 diabetes. Another possible mechanism could be based on the inflammatory cascade triggered by IL16, which involves the induction of IL6, an inflammatory cytokines that is suggested as an independent prognostic markers of severe acute pancreatitis.[53]

7.2.3. rs3809849, rs11556218 & Pancreatitis

It is rather tempting to provide a simplistic model that combines the cumulative knowledge on the involvement of *MYBBP1A* and *IL16* genes in increasing the risk of pancreatitis. Briefly, on one hand, the minor allele of rs3809849 reduces the expression of *MYBBP1A* gene and its protein, thus rendering the pancreatic cells more sensitive to the effect of ASNase and less capable of damage repair. On the other hand, the minor allele of rs11556218 increases the secretion of IL16 by monocytes infiltrating the pancreatic lesion thus further exacerbating the response to injury by recruiting more lymphocytes and triggering the production of other inflammatory cytokines. This strong synergistic effect of the two SNPs was demonstrated in (Discussion Figure 1). While carrying the minor allele of rs3809849 is sufficient to significantly increase the risk of pancreatitis by more than double, patients who also carry the minor allele of rs11556218 have a 15-fold increase in the risk of developing this toxicity. However, the entire model is merely hypothetical and requires further investigation.



Discussion Figure 1. Additive effect of carrying the minor alleles of rs11556218 in *IL16* gene and rs3809849 in *MYBBP1A* gene on the risk of pancreatitis.

This association analysis was performed in the combined DFCI cohort described in chapter 4 of this thesis. The p values of the differences between the groups calculated by Pearson's chi-square method are provided, along with the odd-ratio and 95% confidence interval (in brackets). The genotype groups are indicated at the bottom of the graph and the frequency of patients with pancreatitis in each category is represented by red bars.

7.2.4. Other associations worth discussing

7.2.4.1. rs3809849 and chemotherapy-induced osteonecrosis

One of the mechanisms by which ASNase is suggested to exert its toxic side-effects is through its ability to influence the exposure to other drugs. This is most relevant in the context of glucocorticoids as ASNase induced antibodies can decrease the plasma exposure to itself as well as to dexamethasone; which has been associated with a higher risk of relapse.[54,55] On the other hand, it has been reported, both in animal and clinical studies, that the concomitant administration of ASNase with dexamethasone can significantly increase the risk of osteonecrosis, one of the most common side effects to glucocorticoids, plausibly due to decreased clearance and increased exposure of dexamethasone.[56,57]

Osteonecrosis (ON) is one of the most vexing problems associated with contemporary therapy for ALL [58-60] and is majorly attributed to the use of corticosteroids, like prednisone and dexamethasone.[58,61-69] They exert their anti-cancer effect by inducing apoptosis of leukemia cells, whereas their undesired effect comes from their influence on the number or function of osteoclasts/osteoblasts, eventually promoting bone loss by increased bone resorption and ultimately leading to osteoporosis and vertebral fractures.[70-72] Osteonecrosis occurs in 5 to 10% of patients; [69,73] and while in some patients it may remain asymptomatic and cause no disabilities, in others, it can be serious and debilitating. It can manifest in severe pain, joint damage or articular collapse, particularly affecting the hips, knees, shoulders, and ankles, and often requires surgical management including joint

replacement.[69,73] It is a dose-limiting toxicity that can prompt early withdrawal of CS from therapy for ALL [74] and modifications to glucocorticoid administration schedules can decrease the risk of osteonecrosis.[75]

Several pharmacogenetics studies have been performed in order to identify genetic variations that can influence the risk of ON with the hope of being able to better understand the mechanisms underlying the predisposing factors to this toxicity and consequently implementing personalized treatment or prophylactic options. Many polymorphisms have been linked to an altered risk of developing ON [1,57,74,76-78] such as variants near the glutamate receptor *GRIN3A* locus,[74] the bone morphogenic protein 7 (*BMP7)*[76] and within *Acid Phosphatase 1* (*ACP1*) gene [57] and the *BCL2-Like 11* (BCL2L11) gene encoding Bim protein, among many others.[78]

Following the notable observation of the association of the rs3809849 variant allele in *MYBBP1A* gene with several major toxicities of ASNase treatment, and given the possibility that ASNase can affect the response to GC treatment and consequently alter the risk of osteonecrosis, we explored the possibility of an association between this allele and the risk of osteonecrosis. Intriguingly, the analysis demonstrated a significant association that suggests an additive effect of harboring the variant allele that results in an increased risk of osteonecrosis (Unpublished Data Figure U1 – Chapter 4). A possible explanation lies in the fact that MYBBP1A protein interacts with *c-MYB* product and suppresses its transactivation activity, [79,80] and that *MYB* gene expression was previously shown to modulate dexamethasone-responsiveness and BCL2 mediated apoptosis.[81]

Moreover, since several clinical factors are known to influence the risk of this toxicity, we performed a stratified analysis to test the observed association in clinical subgroups (Unpublished Data Table U1 – Chapter 4). It is generally known that a higher risk of ON is associated with increasing age [58] (as the prevalence in teenagers is between 15 to 20%), gender (as girls are affected more commonly than boys),[59,69,73] higher body mass index,[82] and concomitant drugs used in ALL therapy.[56] In our analysis, the association was significant in female patients, but not in males, suggesting the possibility of an interaction between gender and genotype. Nonetheless, the more interesting finding remains the strong association in younger children but not in the older group, which goes opposite to the general direction of age-related risk of osteonecrosis and suggests an interaction between age and genotype and merits further investigation.

Treatment with dexamethasone has been associated with an improved CNS penetration and decreased risk of relapse than seen with prednisone, but this comes at the expense of an increased incidence of toxicities, including avascular necrosis, infection, and reduction in linear growth.[83] However, in our analysis, there was no significant difference in the risk of osteonecrosis relative to genotype among patients treated with prednisone or dexamethasone (data not provided).

Furthermore, the association of rs3809849 genotype with the risk of osteonecrosis maintained its significance when tested in a multi-variant model that incorporated all other factors that can possibly alter this outcome, which can indicate that the observed association is independent from these factors (Unpublished Data Table U2 – Chapter 4). Of note, age and type of corticosteroids used also remained significant in this model.

Another interesting finding in this stratification analysis was the positive association of the genotype with ON in the group of patients who did not manifest ASNase-induced allergies, but not in the group that did. This could be related to an elevated dexamethasone clearance due to possible silent-inactivation and ASNase induced antibodies, thus reducing the total exposure of dexamethasone and resulting in lower toxicity. However, this remains a mere speculation and the results could be biased by the low number of patients in the group that did experience allergies. This can be tested in future observational studies that would combine data on rs3809849 genotype and *MYBBP1A* expression, with clinical parameters measuring the levels of anti-ASNase antibodies, ASNase and dexamethasone clearance, as well as the development of allergic reactions versus osteonecrosis, among other complications.

7.2.4.2. rs11556218 and drug sensitivity, and ALL prognosis

Studies have reported that treatment with anti–IL-16 results in an increased apoptosis of CD4+ T cells [50] and that the deficiency of its precursor (pro-IL-16) is a common observation in human T-cell leukemia and lymphoma cell lines probably implicating this cytokine in the development of T-cell malignancy. Indeed, the Introduction of pro-interleukin-16 can inhibit T-lymphoblastic leukemia growth and induces cells to become quiescent.[84] CD4+ T lymphocytes were shown to have a distinct response to IL-16 compared to monocytes and macrophages.[27]

Given the essential role of the IL16 protein and its precursor in regulating the immune cells mediated inflammatory response, and since the carrier status of the variant allele of rs11556218 was previously shown to be significantly associated with lymphocyte count (5.5×10^{-7}),[85] we investigated a possible association between this variant and the initial number of white blood cells at the time of the diagnosis. Of note, carriers of this variant allele were more likely to have high white blood cell count at presentation (Unpublished Data Figure U2 – Chapter 4). While such result is usually indicative of a poor outcome, it was rather intriguing to find that this particular group of patients who present high white blood cell count, but also harbour the variant allele, had a significantly better event-free survival than none-carriers in the same category. Therefore, we tested a possible association between the genotype of rs11556218 and cell sensitivity to different drugs in lymphoblastoid cell lines (LCLs) (Unpublished Data Figure U3 – Chapter 4). Interestingly, the screening suggested that carrier status of the variant allele can render LCLs more sensitive to treatment with ASNase as demonstrated by the significant reduction in IC_{50} (Unpublished Data Figure U1a – Chapter 4). This effect was not present upon treatment with prednisone (Unpublished Data Figure U1b – Chapter 4); thereby suggesting that the effect of this polymorphism might specific to ASNase treatment.

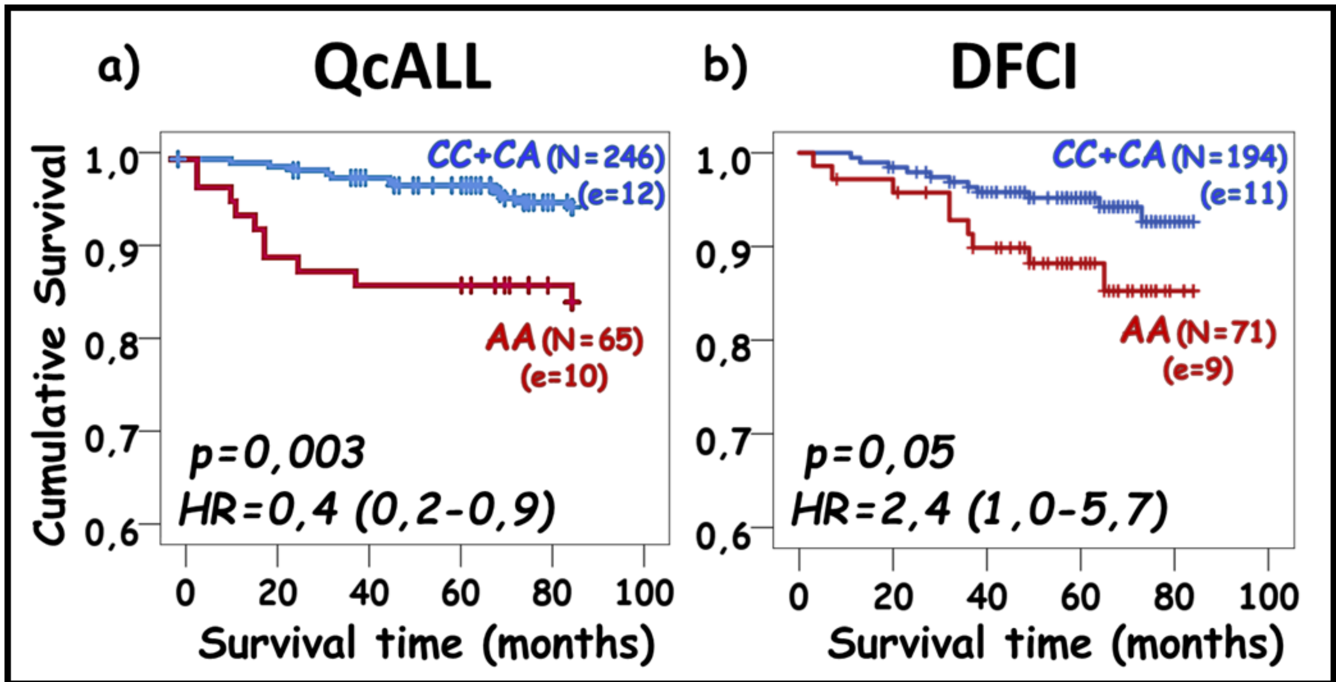
7.3. Discussion of Section C

In a similar fashion to the first EWAS project presented earlier, we carried another analysis that focused on the identification of SNPs that modulate the risk of developing high-grade vincristine-induced peripheral neuropathy (VIPN). We identified risk alleles for *rs2781377* in *SYNE2* and rs10513762 in *MRPL47* that were associated with an increase in the possibility of developing this form of neurotoxicity, as well as another SNP, rs3803357 in the *BAHD1* gene, that had a protective effect by reducing the risk of VIPN. The potential mechanisms through which these identified genes could exert their functional roles were elaborated in Chapter-6.

Once again, we demonstrated a strong combined effect of having more than one risk allele on the development of VIPN. This finding is of particular importance because, despite the lack of replication of the individual associations in the independent validation cohort, which could argue against their validity, the association of their combined effect with VIPN had a borderline significance upon replication. This further supports the utility of incorporating genetic data on multiple risk-associated SNPs into a comprehensive polygenic model to better estimate the extent of the combined contribution of individual risk alleles on the overall risk of developing the toxicity and its severity. Thus, we used this information, combined with data from a previous study, to build a prediction model that was able to classify patients into different risk groups based on their genetic profile. This model was efficient and reproducible when tested in the validation cohort.

The discrepancy of the individual association results between the two groups could be attributed to the considerable difference between the two treatment protocols in regards to the dosing of vincristine, as the AIEOP group received significantly lower cumulative vincristine dose than patients in the QcALL group, which could have played a role in modulating the effect of the genes and their polymorphisms. Moreover, the frequency of VIPN was significantly lower in the AIEOP group, which was to be expected since this toxicity is dose-dependent. Consequently, the lower number of affected patients could have reduced the power to detect the association in this cohort. This scenario reinforces the need for testing the reproducibility of signals identified through association studies in independent validation cohorts that follow the same treatment protocol of the discovery group, as well as in cohorts following other protocols in order to better understand the magnitude of the effects and determine its universality.

Unfortunately, we did not have clinical data on VIPN for patients from the DFCI cohort described in Chapter-4, which involved patients treated following the same DFCI protocols as the discovery cohort, so we were not able to test the impact of the identified polymorphisms on the risk of this toxicity. Interestingly, however, since the protective variant in *BAHD1* gene, rs3803357, was also significantly associated with modulation of the survival outcome in the QcALL cohort, and given the availability of survival data in the DFCI group as explained in Chapter-4, we tested the reproducibility of this association in the DFCI cohort and found the same significant association with overall survival (Discussion Figure 2).



Discussion Figure 2. Association of rs3803357 in *BAHD1* gene with Overall Survival in the a) QcALL and b) DFCI cohorts.

This association analysis was performed in the discovery (QcALL) and replication (DFCI) cohort described in chapter 4 of this thesis. The p-values obtained by the log rank test for the difference across genotypes are provided on each plot. The number of patients represented by each genotype and number of patients with event (in brackets) are indicated next to each curve. Hazard-ratios (HR) obtained through Cox-regression analysis are given with 95% CI.

7.4. Limitations

The results of the work presented in this thesis might have been influenced by several limitations that should be addressed.

It should be noted that due to the low frequency of childhood ALL among the general population, most association studies are performed on relatively small cohorts. This would create an inherent error in association studies which can increase the possibility of finding false-positive associations. While the statistical methods applied in current association studies aim to reduce this possibility by applying stringent adjustment techniques, it remains essential to replicate the findings in independent validation cohorts in order to confirm their statistical and clinical validity. Moreover, since the effect of the associated genes, or their variants, can be modified by patients' clinical factors, ethnicity, or protocol specific characteristics, it strongly calls for the need to also test the newly discovered associations in cohorts that involve patients from various ethnicities and who followed different treatment protocols before attempting to extrapolate data on their clinical utility. Indeed, we can conclude from the results of the EWAS which investigated ASNase-related complications that only few associations were successfully validated in the replication cohort; even though this latter cohort was composed of patients with similar characteristics to the first and who were treated following the same protocol. Furthermore, none of the findings of the second EWAS that investigated the associations with vincristine-induced peripheral neuropathy was replicated individually in the validation cohort that had a considerably different treatment protocol and vincristine dosing regimens compared to the

discovery cohort. Nonetheless, given the relatively pertinent functions of the associated genes to the studied toxicities, all of the associations reported in this work remain quite interesting and merit being investigated in other cohorts as well as in functional studies.

Also, by using whole-exome rather than whole-genome sequencing data to perform the association analysis, we risked missing important variants in non-coding regions of the genome. However, the choice of the study design depended on the type of data available for the association. Also, focusing on non-synonymous polymorphisms in the exonic regions provides a cheaper and less challenging alternative to GWAS and has the advantage of detecting variants with higher probability of functional involvement. Indeed, the initial motive behind the implementation of the whole-exome approach was driven by the prior availability of the sequencing data and was based on the assumption that by concentrating our focus on polymorphisms in the coding region that are predicted to have a functional impact, we would privilege the true-positive associations that can be reproduced later in the validation cohort. Nonetheless, given that each of the analyses performed detected a considerable number of signals, that then needed to be validated through genotyping data and replication in an independent cohort, and due to different constraints (technical and financial), we decided to apply a selective exclusion approach to narrow the list down in order to focus our resources on the polymorphisms that were sufficiently represented in the discovery cohort and had pertinent functions or expression profiles in the context of the studied complication; therefore holding the highest potential to be a true-positive association. We do acknowledge this selective exclusion step as an important limitation to the study design and a deviation from the hypothesis-free concept of whole-exome/genome-approach that might have resulted in overlooking some novel, potentially

meaningful associations. Nevertheless, the success of subsequent validation steps performed on the selected polymorphisms as well as the supporting evidence of functional involvement coming from the cell-based analyses attest to the reliability of the filtration method in prioritizing the signals according to potential importance

While the results of the functional analysis of *MYBBP1A* gene knockout PANC1 cells seem promising and provide a reasonable continuation of the EWAS analysis and further support the involvement of the gene in ASNase-related toxicities, they do not provide justification for the observed associations at the polymorphism level. It is tempting to speculate that this rs3809849 polymorphism is modulating the risk of pancreatitis through altering the expression of the *MYBBP1A* gene (Figure 1 in Chapter 6). Indeed, the variant allele was found to be a strong eQTL in many tissues by reducing the expression of the gene, which goes in a similar direction to silencing and knockout. However, this should be interpreted with caution since the extent of the effect of this polymorphism varies significantly across the different tissues and only had borderline significance in the pancreas.[86] There is also the other possibility that, being in the exonic region, this polymorphism can affect the function of the protein, which cannot be ruled out without further testing, even though computational prediction models seem to suggest that it does not have a deleterious effect on the protein's function.[87,88]

Moreover, the effect of overexpressing the protein of this gene, along with the different alleles of rs3809849 polymorphism, can provide important insight on the role of the over-activation of the *MYBBP1A* gene on different biological processes. One way to test

this can be by comparing the cellular behaviour and drug resistance profile between cells overexpressing the protein product with and without this particular variant.

While this work demonstrated a functional implication of the *MYBBP1A* gene in PANC1 cells, it is important to take in consideration that this is a pancreatic cancer cell line; and thus, caution should be made when interpreting the results in the context of normal pancreatic tissue response to ASNase-induced acute pancreatitis. Moreover, it would be tricky to extrapolate the conclusions from this study and generalize them onto other pancreatic cell lines, since studies have shown that the role of *MYBBP1A* varies across the different tissues. Therefore, pharmacogenetics studies of ASNase-response in other pancreatic cell lines and/or other tissues can be useful to advance our understanding on the cell type selective role of this gene.

Furthermore, if the functional role of MYBBP1A gene in modulating the various activities of the pancreatic tissue is successfully confirmed at the cellular level, the next step should aim to understand how this comes into play at the level of the organ as well as the organism. Several experimental animal models of acute and chronic pancreatitis are available and can provide a clinically relevant platform to understand the impact of perturbations to the activity of MYBBP1A gene and its protein on the development and/or exacerbation of pancreatitis.[89] For instance, the impact of gene deletion or overexpression can be studied in rats or mice using the cerulein-induced acute pancreatitis model, whereby acute pancreatitis is induced by an intravenous or intraperitoneal injection of an overdose of cerulean, a hormone known to induce pancreatic enzyme activation. This model is highly reproducible and the phenotype is reversible upon the withdrawal of cerulean, which

allowed it to become one of the most extensively used models for acute pancreatitis.[89] In this context, changes to the molecular pathways and the frequency and severity of pancreatitis can be compared between MYBBP1A gene knockout animals or those overexpressing it, and control animals with the same genetic background but a normal gene/protein activity. Likewise, an inducible, tissue-specific Cre-loxP recombination system can be applied in a mouse model offering the advantage of temporal control of the gene activity using a pancreas-specific promoter.[90] This model would allow for testing the effect of ANSase on the pancreatic tissue of the same animal at different time pointe before and after the induction of the change in gene activity, thus providing a valuable insight on the role of this gene in the mechanisms leading to ASNase-induced acute pancreatitis. Moreover, this particular model also allows for testing the efficacy of different measures that can help to mitigate this drug-specific toxicity.

In order to understand the specific role of the polymorphism, a knock-in experiment in which the modified sequence containing the allele of interest is inserted into the genome of the knockout cells to replace the wild-type sequence should be performed. Also, it is becoming more recognized that the CRISPR-Cas9 technique itself can induce off-target effects, which would alter the results of the following analysis in a random way; thereby producing clones with variable behaviour which can introduce a source of variability to the analysis. Testing multiple clones might be useful to eliminate the possibility that the results are clone-specific and not related to the intended editing of the gene. However, performing rescue experiments in which the modified sequence is reinserted into the genome in its original form, followed by revaluation of the functional impact should be considered the gold standard, but can be very challenging.

7.5. Prospective Studies

7.5.1. Pancreatitis

Based on previous studies showing that knocking-down the *MYBBP1A* gene can induce apoptosis in different cell lines, it would be interesting to characterizing the change in percentage of apoptotic or dead cells following the deletion of the gene in PANC1 cells, and also upon treatment with ASNase and other drugs. Also, since this gene is involved in cell division and in regulating the cell-cycle, it can be quite useful to determine how the deletion of this gene, as well as ASNase challenge, would affect the percentage of dividing, not-dividing and dead cells, as well as the ratio of anomalous mitotic figures to further understand its role in pancreatic cells proliferation and viability.

Moreover, the observation that MYBBP1A knockout PANC1 cells seem to undergo a change in morphology and adopt a phenotype suggestive of an epithelial-mesenchymal transition merits to be investigated at the molecular level by to characterizing the protein expression levels of different markers of EMT (e.g. the epithelial marker E-cadherin, the mesenchymal marker vimentin and the EMT regulators like SNAIL and NF-kB).

As for the IL16 gene, it would be useful to study the effect of gene deletion, as well as the identified polymorphism, on the behavior of leukemic blood cells, particularly the T-lymphocytes, and to characterize the impact on their sensitivity to treatment.

7.5.2. Vincristine-induced peripheral neuropathy

In a follow up project, we intend to study the functional role of the associated variant, as well as their potentially synergistic combined-effect, in modulating the risk of VIPN. Here, we propose a study design whereby genetic data is integrated into an *in-vitro* model for assessing the sensitivity of iPSC-derived neurons in response to treatment with vincristine. Accordingly, peripheral neurons will be derived from induced pluripotent stem cells (iPSCs) and will be treated with variable concentrations of vincristine followed by phenotype assessment 72 hours post incubation. Phenotype will be characterized based on multiple parameters which will include cellular morphological changes accompanied with cell viability and apoptosis assays, as well as measurement of the neurite outgrowth and mitochondrial migration using high content imaging techniques. [91]

To begin with, we sought to employ and optimize a differentiation protocol based on techniques previously published in literature to derive human neuron cells from iPSCs.[92-96] Since the neurotoxic effect of vincristine is attributed to its toxicity on the peripheral nerves (mostly sensorial, but also motor and autonomous neurons to a lesser extent), we aimed at producing the particular subtype of peripheral sensory neurons. Indeed, we were able to successfully derive peripheral sensory neurons from an in-house sample previously reprogrammed into iPSC, and the specific neuronal subtype was confirmed by immuno-histological techniques testing for the co-expression of the following markers: Peripherin (specific to peripheral neurons) and BRN3A (specific to sensorial neurons). This particular

neuronal subtype was shown to be sensitive to treatment to vincristine and thus could serve as a model to test the interaction of the drug with the genetic background.[91]

Moreover, we acquired a total of 20 iPSCs cell-lines, reprogrammed starting from fibroblast samples of CEU individuals. Genetic sequencing data is available for all patients, allowing for the selection of representative samples for the assessment of the effect of distinct genotypes for each of the SNPs, as well as their possible combinations. This project would provide valuable insight on the role of the identified genes and their variants in modulating the risk of VIPN.

7.6. Conclusion

Inter-individual drug response can be extremely variable which could be attributed to multiple factors contributing to the eventual phenotypic outcome. The observed association between genetic alterations and modulation of the risk of treatment toxicity in clinical setting of ALL may be a result of a change in sensitivity to particular components of the multi-agent treatment protocols driven by genetics. Understanding the role of genetic alterations in modulating drug response could provide valuable insights on how to optimize therapeutic methods in order to ensure maximum benefit with minimum risk.

Pharmacogenomics holds the potential to enhance the efficacy and precision of existing drug dosing regimens by empowering clinicians to better calculate the overall risk/benefit ratio of a given treatment, which remains a big challenge in the vulnerable pediatric population. As next-generation sequencing is evolving into a revolutionary tool in genetics, the ability to sequence the entire human genome holds the promise to improve the pharmacogenomics knowledge. In the meanwhile, whole-genome sequencing is still costly and labor intense. WES provides a cheaper and less challenging alternative which is proving it utility in the research, diagnostic and clinical settings.

Using whole-exome Sequencing data in the context of exome-wide association Studies is a successful ``hypothesis-free`` strategy that can be used to identify significant genetic associations with adverse drug reactions in children treated for ALL without any prior selection for specific regions, genes, or variants.

We were able to identify common genetic variants significantly associated with ASNase complications and vincristine-induced peripheral neuropathy, and the role of *MYBBP1A* gene was further confirmed through follow-up functional studies. However, further cell based and animal based studies need to be done in order to confirm the role of this and other identified genes and their variants in modulating the respective toxicities before they can be considered in a clinical context.

One of the prominent observations in this work is the reproducibility of polygenic models in classifying patients according to the predicted genetic risk score. Therefore, future studies should aim to simultaneously test for the interaction between a wide range of environmental, genetic, and patient specific factors favoring the application of personalized medicine.

This work contributes to the general knowledge of the genetic roots of variability and how it influence treatment response from one patient to the other.

7.7. References

1. Mei L, Ontiveros EP, Griffiths EA, Thompson JE, Wang ES, Wetzler M. Pharmacogenetics predictive of response and toxicity in acute lymphoblastic leukemia therapy. *Blood Rev.* 2015;29(4):243-249.

2. Zhang G, Nebert DW. Personalized medicine: Genetic risk prediction of drug response. *Pharmacol Ther.* 2017;175:75-90.

3. Gervasini G, Vagace JM. Impact of genetic polymorphisms on chemotherapy toxicity in childhood acute lymphoblastic leukemia. *Front Genet.* 2012;3:249.

4. Abaji R, Krajinovic M. Thiopurine S-methyltransferase polymorphisms in acute lymphoblastic leukemia, inflammatory bowel disease and autoimmune disorders: influence on treatment response. *Pharmgenomics Pers Med.* 2017;10:143-156.

5. Gu G, Wells JM, Dombkowski D, Preffer F, Aronow B, Melton DA. Global expression analysis of gene regulatory pathways during endocrine pancreatic development. *Development.* 2004;131(1):165-179.

6. George B, Horn D, Bayo P, et al. Regulation and function of Myb-binding protein 1A (MYBBP1A) in cellular senescence and pathogenesis of head and neck cancer. *Cancer Lett.* 2015;358(2):191-199.

7. Rakonczay Z, Jr., Hegyi P, Takacs T, McCarroll J, Saluja AK. The role of NF-kappaB activation in the pathogenesis of acute pancreatitis. *Gut.* 2008;57(2):259-267.

8. Julien S, Puig I, Caretti E, et al. Activation of NF-kappaB by Akt upregulates Snail expression and induces epithelium mesenchyme transition. *Oncogene.* 2007;26(53):7445-7456.

9. Acuna Sanhueza GA, Faller L, George B, et al. Opposing function of MYBBP1A in proliferation and migration of head and neck squamous cell carcinoma cells. *BMC Cancer.* 2012;12:72.

10. Tan BC, Yang CC, Hsieh CL, et al. Epigeneitc silencing of ribosomal RNA genes by Mybbp1a. *J Biomed Sci.* 2012;19:57.

11. Cao CL, Duan PY, Zhang WJ, et al. Acute pancreatitis induced by etoposide-lobaplatin combination chemotherapy used for the treatment of lung cancer: A case report and literature review. *Medicine (Baltimore).* 2017;96(29):e7601.

12. Nashwan A, Elmalik H, Krishnan Nair S, Yassin M. Acute Pancreatitis-Induced by Cisplatin-Etoposide (CE) Regimen. *Journal of Case Reports and Images in Oncology.* 2016;2:61-65.

13. Mori S, Bernardi R, Laurent A, et al. Myb-binding protein 1A (MYBBP1A) is essential for early embryonic development, controls cell cycle and mitosis, and acts as a tumor suppressor. *PLoS One.* 2012;7(10):e39723.

14. Kumazawa T, Nishimura K, Katagiri N, Hashimoto S, Hayashi Y, Kimura K. Gradual reduction in rRNA transcription triggers p53 acetylation and apoptosis via MYBBP1A. *Sci Rep.* 2015;5:10854.

15. Ono W, Hayashi Y, Yokoyama W, et al. The nucleolar protein Myb-binding protein 1A (MYBBP1A) enhances p53 tetramerization and acetylation in response to nucleolar disruption. *J Biol Chem.* 2014;289(8):4928-4940.

16. Tang Y, Zhao W, Chen Y, Zhao Y, Gu W. Acetylation is indispensable for p53 activation. *Cell.* 2008;133(4):612-626.

17. Akaogi K, Ono W, Hayashi Y, Kishimoto H, Yanagisawa J. MYBBP1A suppresses breast cancer tumorigenesis by enhancing the p53 dependent anoikis. *BMC Cancer.* 2013;13:65.

18. Diaz VM, Mori S, Longobardi E, et al. p160 Myb-binding protein interacts with Prep1 and inhibits its transcriptional activity. *Mol Cell Biol.* 2007;27(22):7981-7990.

19. Kim SK, MacDonald RJ. Signaling and transcriptional control of pancreatic organogenesis. *Curr Opin Genet Dev.* 2002;12(5):540-547.

20. Kim SK, Selleri L, Lee JS, et al. Pbx1 inactivation disrupts pancreas development and in Ipf1-deficient mice promotes diabetes mellitus. *Nat Genet.* 2002;30(4):430-435.

21. Oriente F, Fernandez Diaz LC, Miele C, et al. Prep1 deficiency induces protection from diabetes and increased insulin sensitivity through a p160-mediated mechanism. *Mol Cell Biol.* 2008;28(18):5634-5645.

22. Das SL, Singh Pp Fau - Phillips ARJ, Phillips Ar Fau - Murphy R, Murphy R Fau - Windsor JA, Windsor Ja Fau - Petrov MS, Petrov MS. Newly diagnosed diabetes mellitus after acute pancreatitis: a systematic review and meta-analysis. (1468-3288 (Electronic)).

23. Fan M, Rhee J, St-Pierre J, et al. Suppression of mitochondrial respiration through recruitment of p160 myb binding protein to PGC-1alpha: modulation by p38 MAPK. *Genes Dev.* 2004;18(3):278-289.

24. Valtat B, Riveline JP, Zhang P, et al. Fetal PGC-1alpha overexpression programs adult pancreatic beta-cell dysfunction. *Diabetes.* 2013;62(4):1206-1216.

25. Jones LC, Okino ST, Gonda TJ, Whitlock JP, Jr. Myb-binding protein 1a augments AhR-dependent gene expression. *J Biol Chem.* 2002;277(25):22515-22519.

26. Xue J, Nguyen DT, Habtezion A. Aryl hydrocarbon receptor regulates pancreatic IL-22 production and protects mice from acute pancreatitis. *Gastroenterology.* 2012;143(6):1670-1680.

27. Mathy NL, Scheuer W, Lanzendorfer M, et al. Interleukin-16 stimulates the expression and production of pro-inflammatory cytokines by human monocytes. *Immunology.* 2000;100(1):63-69.

28. Liu C, Zhao P, Yang Y, Xu X, Wang L, Li B. Ampelopsin suppresses TNF-alpha-induced migration and invasion of U2OS osteosarcoma cells. *Mol Med Rep.* 2016;13(6):4729-4736.

29. Parada NA, Center DM, Kornfeld H, et al. Synergistic activation of CD4+ T cells by IL-16 and IL-2. *J Immunol.* 1998;160(5):2115-2120.

30. Kovacs E. The serum levels of IL-12 and IL-16 in cancer patients. Relation to the tumour stage and previous therapy. *Biomed Pharmacother.* 2001;55(2):111-116.

31. Zhang Y, Center DM, Wu DM, et al. Processing and activation of pro-interleukin-16 by caspase-3. *J Biol Chem.* 1998;273(2):1144-1149.

32. Gao LB, Rao L, Wang YY, et al. The association of interleukin-16 polymorphisms with IL-16 serum levels and risk of colorectal and gastric cancer. *Carcinogenesis.* 2009;30(2):295-299.

33. Tang YJ, Wang JL, Xie KG, Lan CG. Association of interleukin 16 gene polymorphisms and plasma IL16 level with osteosarcoma risk. *Sci Rep.* 2016;6:34607.

34. Alexandrakis MG, Passam FH, Kyriakou DS, et al. Serum level of interleukin-16 in multiple myeloma patients and its relationship to disease activity. *Am J Hematol.* 2004;75(2):101-106.

35. Laberge S, Ernst P, Ghaffar O, et al. Increased expression of interleukin-16 in bronchial mucosa of subjects with atopic asthma. *Am J Respir Cell Mol Biol.* 1997;17(2):193-202.

36. Cruikshank WW, Long A, Tarpy RE, et al. Early identification of interleukin-16 (lymphocyte chemoattractant factor) and macrophage inflammatory protein 1 alpha (MIP1 alpha) in bronchoalveolar lavage fluid of antigen-challenged asthmatics. *Am J Respir Cell Mol Biol.* 1995;13(6):738-747.

37. Mashikian MV, Tarpy RE, Saukkonen JJ, et al. Identification of IL-16 as the lymphocyte chemotactic activity in the bronchoalveolar lavage fluid of histamine-challenged asthmatic patients. *J Allergy Clin Immunol.* 1998;101(6 Pt 1):786-792.

38. Laberge S, Durham SR, Ghaffar O, et al. Expression of IL-16 in allergen-induced late-phase nasal responses and relation to topical glucocorticosteroid treatment. *J Allergy Clin Immunol.* 1997;100(4):569-574.

39. Zhou P, Goldstein S, Devadas K, Tewari D, Notkins AL. Human CD4+ cells transfected with IL-16 cDNA are resistant to HIV-1 infection: inhibition of mRNA expression. *Nat Med.* 1997;3(6):659-664.

40. Liu Z, Ma L, Qiu S, Jia T. Genetic polymorphisms of interleukin-16 are associated with susceptibility to primary knee osteoarthritis. *Int J Clin Exp Med.* 2015;8(1):1401-1405.

41. Khoshbakht T, Soosanabadi M, Neishaboury M, et al. An Association Study on IL16 Gene Polymorphisms with the Risk of Sporadic Alzheimer's Disease. *Avicenna J Med Biotechnol.* 2015;7(3):128-132.

42. Bowler RP, Bahr TM, Hughes G, et al. Integrative omics approach identifies interleukin-16 as a biomarker of emphysema. *OMICS.* 2013;17(12):619-626.

43. Tong Z, Li Q, Zhang J, Wei Y, Miao G, Yang X. Association between interleukin 6 and interleukin 16 gene polymorphisms and coronary heart disease risk in a Chinese population. *J Int Med Res.* 2013;41(4):1049-1056.

44. Liu XL, Du JZ, Zhou YM, Shu QF, Li YG. Interleukin-16 polymorphism is associated with an increased risk of ischemic stroke. *Mediators Inflamm.* 2013;2013:564750.

45. Xue H, Gao L, Wu Y, et al. The IL-16 gene polymorphisms and the risk of the systemic lupus erythematosus. *Clin Chim Acta.* 2009;403(1-2):223-225.

46. Romani S, Hosseini SM, Mohebbi SR, et al. Interleukin-16 gene polymorphisms are considerable host genetic factors for patients' susceptibility to chronic hepatitis B infection. *Hepat Res Treat.* 2014;2014:790753.

47. Luo SX, Li S, Zhang XH, et al. Genetic polymorphisms of interleukin-16 and risk of knee osteoarthritis. *PLoS One.* 2015;10(5):e0123442.

48. Mo CJ, Peng QL, He Y, et al. Positive association between IL-16 rs11556218 T/G polymorphism and cancer risk: a meta-analysis. *Asian Pac J Cancer Prev.* 2014;15(11):4697-4703.

49. IMMUNPOP. Expression QTL-rs11556218-IL16. http://132.219.138.157/kim/eQTL/. Accessed 21-12, 2018.

50. Meagher C, Beilke J, Arreaza G, et al. Neutralization of interleukin-16 protects nonobese diabetic mice from autoimmune type 1 diabetes by a CCL4-dependent mechanism. *Diabetes.* 2010;59(11):2862-2871.

51. Burrack AL, Martinov T, Fife BT. T Cell-Mediated Beta Cell Destruction: Autoimmunity and Alloimmunity in the Context of Type 1 Diabetes. *Front Endocrinol (Lausanne).* 2017;8:343.

52. Carrero JA, McCarthy DP, Ferris ST, et al. Resident macrophages of pancreatic islets have a seminal role in the initiation of autoimmune diabetes of NOD mice. *Proc Natl Acad Sci U S A.* 2017;114(48):E10418-E10427.

53. Nieminen A, Maksimow M, Mentula P, et al. Circulating cytokines in predicting development of severe acute pancreatitis. *Crit Care.* 2014;18(3):R104.

54. Fernandez CA, Smith C, Yang W, et al. Genome-wide analysis links NFATC2 with asparaginase hypersensitivity. *Blood.* 2015;126(1):69-75.

55. Kawedia JD, Liu C, Pei D, et al. Dexamethasone exposure and asparaginase antibodies affect relapse risk in acute lymphoblastic leukemia. *Blood.* 2012;119(7):1658-1664.

56. Liu C, Janke LJ, Kawedia JD, et al. Asparaginase Potentiates Glucocorticoid-Induced Osteonecrosis in a Mouse Model. *PLoS One.* 2016;11(3):e0151433.

57. Kawedia JD, Kaste SC, Pei D, et al. Pharmacokinetic, pharmacodynamic, and pharmacogenetic determinants of osteonecrosis in children with acute lymphoblastic leukemia. *Blood.* 2011;117(8):2340-2347; quiz 2556.

58. Patel B, Richards SM, Rowe JM, Goldstone AH, Fielding AK. High incidence of avascular necrosis in adolescents with acute lymphoblastic leukaemia: a UKALL XII analysis. *Leukemia.* 2008;22(2):308-312.

59. Mattano LA, Jr., Sather HN, Trigg ME, Nachman JB. Osteonecrosis as a complication of treating acute lymphoblastic leukemia in children: a report from the Children's Cancer Group. *J Clin Oncol.* 2000;18(18):3262-3272.

60. Arico M, Boccalatte MF, Silvestri D, et al. Osteonecrosis: An emerging complication of intensive chemotherapy for childhood acute lymphoblastic leukemia. *Haematologica.* 2003;88(7):747-753.

61. Silverman LB, Stevenson KE, O'Brien JE, et al. Long-term results of Dana-Farber Cancer Institute ALL Consortium protocols for children with newly diagnosed acute lymphoblastic leukemia (1985-2000). *Leukemia.* 2010;24(2):320-334.

62. Silverman LB, Gelber RD, Dalton VK, et al. Improved outcome for children with acute lymphoblastic leukemia: results of Dana-Farber Consortium Protocol 91-01. *Blood.* 2001;97(5):1211-1218.

63. Moghrabi A, Levy DE, Asselin B, et al. Results of the Dana-Farber Cancer Institute ALL Consortium Protocol 95-01 for children with acute lymphoblastic leukemia. *Blood.* 2007;109(3):896-904.

64. te Winkel ML, Appel IM, Pieters R, van den Heuvel-Eibrink MM. Impaired dexamethasone-related increase of anticoagulants is associated with the development of osteonecrosis in childhood acute lymphoblastic leukemia. *Haematologica.* 2008;93(10):1570-1574.

65. Belgaumi AF, Al-Bakrah M, Al-Mahr M, et al. Dexamethasone-associated toxicity during induction chemotherapy for childhood acute lymphoblastic leukemia is augmented by concurrent use of daunomycin. *Cancer.* 2003;97(11):2898-2903.

66. Chan KL, Mok CC. Glucocorticoid-induced avascular bone necrosis: diagnosis and management. *Open Orthop J.* 2012;6:449-457.

67. Hyakuna N, Shimomura Y, Watanabe A, et al. Assessment of corticosteroid-induced osteonecrosis in children undergoing chemotherapy for acute lymphoblastic leukemia: a report from the Japanese Childhood Cancer and Leukemia Study Group. *J Pediatr Hematol Oncol.* 2014;36(1):22-29.

68. Strauss AJ, Su JT, Dalton VM, Gelber RD, Sallan SE, Silverman LB. Bony morbidity in children treated for acute lymphoblastic leukemia. *J Clin Oncol.* 2001;19(12):3066-3072.

69. Hunger SP, Mullighan CG. Acute Lymphoblastic Leukemia in Children. *N Engl J Med.* 2015;373(16):1541-1552.

70. Mazziotti G, Giustina A, Canalis E, Bilezikian JP. Glucocorticoid-induced osteoporosis: clinical and therapeutic aspects. *Arq Bras Endocrinol Metabol.* 2007;51(8):1404-1412.

71. Alos N, Grant RM, Ramsay T, et al. High incidence of vertebral fractures in children with acute lymphoblastic leukemia 12 months after the initiation of therapy. *J Clin Oncol.* 2012;30(22):2760-2767.

72. Weinstein RS, Jilka RL, Parfitt AM, Manolagas SC. Inhibition of osteoblastogenesis and promotion of apoptosis of osteoblasts and osteocytes by glucocorticoids. Potential mechanisms of their deleterious effects on bone. *J Clin Invest.* 1998;102(2):274-282.

73. Te Winkel ML, Pieters R, Wind EJ, Bessems JH, van den Heuvel-Eibrink MM. Management and treatment of osteonecrosis in children and adolescents with acute lymphoblastic leukemia. *Haematologica.* 2014;99(3):430-436.

74. Karol SE, Yang W, Van Driest SL, et al. Genetics of glucocorticoid-associated osteonecrosis in children with acute lymphoblastic leukemia. *Blood.* 2015;126(15):1770-1776.

75. Mattano LA, Jr., Devidas M, Nachman JB, et al. Effect of alternate-week versus continuous dexamethasone scheduling on the risk of osteonecrosis in paediatric patients with acute lymphoblastic leukaemia: results from the CCG-1961 randomised cohort trial. *Lancet Oncol.* 2012;13(9):906-915.

76. Karol SE, Mattano LA, Jr., Yang W, et al. Genetic risk factors for the development of osteonecrosis in children under age 10 treated for acute lymphoblastic leukemia. *Blood.* 2016;127(5):558-564.

77. Gagne V, Rousseau J, Labuda M, et al. Bim polymorphisms: influence on function and response to treatment in children with acute lymphoblastic leukemia. *Clin Cancer Res.* 2013;19(18):5240-5249.

78. Plesa M, Gagne V, Glisovic S, et al. Influence of BCL2L11 polymorphism on osteonecrosis during treatment of childhood acute lymphoblastic leukemia. *Pharmacogenomics J.* 2017.

79. Favier D, Gonda TJ. Detection of proteins that bind to the leucine zipper motif of c-Myb. *Oncogene.* 1994;9(1):305-311.

80. Tavner FJ, Simpson R, Tashiro S, et al. Molecular cloning reveals that the p160 Myb-binding protein is a novel, predominantly nucleolar protein which may play a role in transactivation by Myb. *Mol Cell Biol.* 1998;18(2):989-1002.

81. Jing D, Bhadri VA, Beck D, et al. Opposing regulation of BIM and BCL2 controls glucocorticoid-induced apoptosis of pediatric acute lymphoblastic leukemia cells. *Blood.* 2015;125(2):273-283.

82. Niinimaki RA, Harila-Saari AH, Jartti AE, et al. High body mass index increases the risk for osteonecrosis in children with acute lymphoblastic leukemia. *J Clin Oncol.* 2007;25(12):1498-1504.

83. Cooper SL, Brown PA. Treatment of pediatric acute lymphoblastic leukemia. *Pediatr Clin North Am.* 2015;62(1):61-73.

84. Richmond J, Finkel M, Studwell A, Little F, Cruikshank W. Introduction of pro-interleukin-16 inhibits T-lymphoblastic leukemia growth in mice. *J Cancer Res Clin Oncol.* 2011;137(10):1581-1585.

85. IL16. University of Cambridge; 2018. http://www.phenoscanner.medschl.cam.ac.uk/?query=IL16&catalogue=GWAS&p=1&proxies=None&r2=0.8&build=37. Accessed 20-12-2018.

86. GTEx-Portal. MYBBP1A-rs3809849. 2019; https://gtexportal.org/home/snp/rs3809849. Accessed 11-12, 2019.

87. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet.* 2013;Chapter 7:Unit7 20.

88. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc.* 2009;4(7):1073-1081.

89. Hyun JJ, Lee HS. Experimental models of pancreatitis. *Clin Endosc.* 2014;47(3):212-216.

90. Kim H, Kim M, Im SK, Fang S. Mouse Cre-LoxP system: general principles to determine tissue-specific roles of target genes. *Lab Anim Res.* 2018;34(4):147-159.

91. Wing C, Komatsu M, Delaney SM, Krause M, Wheeler HE, Dolan ME. Application of stem cell derived neuronal cells to evaluate neurotoxic chemotherapy. *Stem Cell Res.* 2017;22:79-88.

92. Cao L, McDonnell A, Nitzsche A, et al. Pharmacological reversal of a pain phenotype in iPSC-derived sensory neurons and patients with inherited erythromelalgia. *Sci Transl Med.* 2016;8(335):335ra356.

93. Chambers SM, Qi Y, Mica Y, et al. Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. *Nat Biotechnol.* 2012;30(7):715-720.

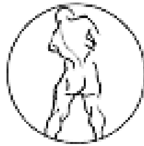
94. Clark AJ, Kaller MS, Galino J, Willison HJ, Rinaldi S, Bennett DLH. Co-cultures with stem cell-derived human sensory neurons reveal regulators of peripheral myelination. *Brain.* 2017;140(4):898-913.

95. Eberhardt E, Havlicek S, Schmidt D, et al. Pattern of Functional TTX-Resistant Sodium Channels Reveals a Developmental Stage of Human iPSC- and ESC-Derived Nociceptors. *Stem Cell Reports.* 2015;5(3):305-313.

96. Young GT, Gutteridge A, Fox H, et al. Characterizing human stem cell-derived sensory neurons at the single-cell level reveals their ion channel expression and utility in pain research. *Mol Ther.* 2014;22(8):1530-1543.

Annex I

Copyright Authorisations



THESIS COPYRIGHT PERMISSION FORM

Title(s) of the Image(s): Terese Winslow LLC owns the copyright to the following image(s):

*Title(s) of illustration(s): **Hematopoietic Tree, Plasma Cell (AV Number: CDR596562)***

Description of the Work: Terese Winslow LLC hereby grants permission to reproduce the above image(s) for use in the work specified:

Thesis title: Using Whole-Exome Sequencing Data In An Exome-Wide Association Study Approach To Identify Genetic Risk Factors Influencing Acute Lymphoblastique Leukemia Response: A Focus on Asparaginase Complications & Vincristine-Induced Peripheral Neuropathy.

*University: **Université de Montréal** (Montréal, Québec, Canada)*

License Granted: Terese Winslow LLC hereby grants limited, non-exclusive worldwide print and electronic rights only for use in the Work specified. Terese Winslow LLC grants such rights "AS IS" without representation or warranty of any kind and shall have no liability in connection with such license.

Restrictions: Reproduction for use in any other work, derivative works, or by any third party by manual or electronic methods is prohibited. Ownership of original artwork, copyright, and all rights not specifically transferred herein remain the exclusive property of Terese Winslow LLC. Additional license(s) are required for ancillary usage(s).

Credit must be placed adjacent to the image(s) in the following format:

For the National Cancer Institute © 2008 Terese Winslow LLC, U.S. Govt. has certain rights

Permission granted to:
*Author name: **Rachid Abaji***

*Mailing address: **Centre de Recherche, CHU Sainte-Justine**, 3175 chemin de la Côte-Ste-Catherine, Montréal, QC, H3T 1C5 Canada, Local : 7.17.027*

Terese Winslow LLC, Medical Illustration
714 South Fairfax Street, Alexandria, Virginia 22314
(703) 836-9121
terese@teresewinslow.com
www.teresewinslow.com

Note: Copyright.com supplies permissions but not the copyrighted content itself.

1 PAYMENT | 2 REVIEW | 3 CONFIRMATION

Step 3: Order Confirmation

Thank you for your order! A confirmation for your order will be sent to your account email address. If you have questions about your order, you can call us 24 hrs/day, M-F at +1.855.239.3415 Toll Free, or write to us at info@copyright.com. This is not an invoice.

Confirmation Number: 11772669
Order Date: 12/09/2018

If you paid by credit card, your order will be finalized and your card will be charged within 24 hours. If you choose to be invoiced, you can change or cancel your order until the invoice is generated.

Payment Information

Rachid Abaji
rasheed3000@homtail.com
+1 (514) 345-4931
Payment Method: n/a

Order Details

JOURNAL OF CLINICAL INVESTIGATION. ONLINE

Order detail ID: 71708993
Order License Id: 4485001053568
ISSN: 1558-8238
Publication Type: e-Journal
Volume:
Issue:
Start page:
Publisher: AMERICAN SOCIETY FOR CLINICAL INVESTIGATION
Author/Editor: American Society for Clinical Investigation

Permission Status: ✔ **Granted**
Permission type: Republish or display content
Type of use: Thesis/Dissertation

View details

Note: This item will be invoiced or charged separately through CCC's **RightsLink** service. More info $ 0.00

Total order items: 1

This is not an invoice.

Order Total: 0.00 USD

New England Journal of Medicine Permissions

Submit Your Permission Request Using RightsLink

If you are seeking permission to copy, reproduce, or republish content from the *New England Journal of Medicine* (NEJM) and you are not the author of the content, you may use the Copyright Clearance Center's RightsLink® service. Simply visit NEJM.org and locate the article or video from which you seek to reuse content.

Confirmation of your permission and related terms and conditions will be sent to you instantly via email for most requests. Some requests may require Publisher review.

If you have questions about using the RightsLink service, please contact:

RightsLink® Customer Support
+1-877-622-5543 (toll free)
+1-978-777-9929
Email: customercare@copyright.com

For general questions about NEJM Permissions, email permissions@nejm.org.

Permission for Authors

If you are the author of an article that has been published in NEJM, see Author Permissions.

Reuse of Content within a Thesis or Dissertation

Content (full-text or portions thereof) may be used in print and electronic versions of a dissertation or thesis without formal permission from the Massachusetts Medical Society (MMS), Publisher of the *New England Journal of Medicine*.

The following credit line must be printed along with the copyrighted material:

Reproduced with permission from (scientific reference citation), Copyright Massachusetts Medical Society.

RightsLink®

Home | Account Info | Help

Title:	Long-term results of Dana-Farber Cancer Institute ALL Consortium protocols for children with newly diagnosed acute lymphoblastic leukemia (1985–2000)
Author:	L B Silverman, K E Stevenson, J E O'Brien, B L Asselin, R D Barr et al.
Publication:	Leukemia
Publisher:	Springer Nature
Date:	Dec 17, 2009

Copyright © 2009, Springer Nature

Logged in as:
Rachid Abaji
Account #:
3001378765
LOGOUT

Order Completed

Thank you for your order.

This Agreement between Rachid Abaji ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

printable details

License Number	4645570179719
License date	Aug 10, 2019
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Leukemia
Licensed Content Title	Long-term results of Dana-Farber Cancer Institute ALL Consortium protocols for children with newly diagnosed acute lymphoblastic leukemia (1985–2000)
Licensed Content Author	L B Silverman, K E Stevenson, J E O'Brien, B L Asselin, R D Barr et al.
Licensed Content Date	Dec 17, 2009
Licensed Content Volume	24
Licensed Content Issue	2
Type of Use	Thesis/Dissertation

Requestor type	academic/university or research institute
Format	electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	no
Title	Using Whole-Exome Sequencing Data in an Exome-Wide Association Study Approach to Identify Genetic Risk Factors Influencing Acute Lymphoblastic Leukemia Response: A Focus on Asparaginase Complications & Vincristine-Induced Peripheral Neuropathy
Institution name	Université de Montréal
Expected presentation date	Aug 2019
Portions	Table 2 Therapy on DFCI ALL Consortium Protocols: 1985–2000
Requestor Location	Rachid Abaji 3175 chemin de la Côte-Ste-Catherine Montreal, QC H3T 1C5 Canada Attn:
Total	0.00 CAD

ORDER MORE CLOSE WINDOW

Copyright © 2019 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement.
Terms and Conditions.
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Welcome, Rachid

Confirmation Number: 11840369
Order Date: 08/10/2019

Customer Information

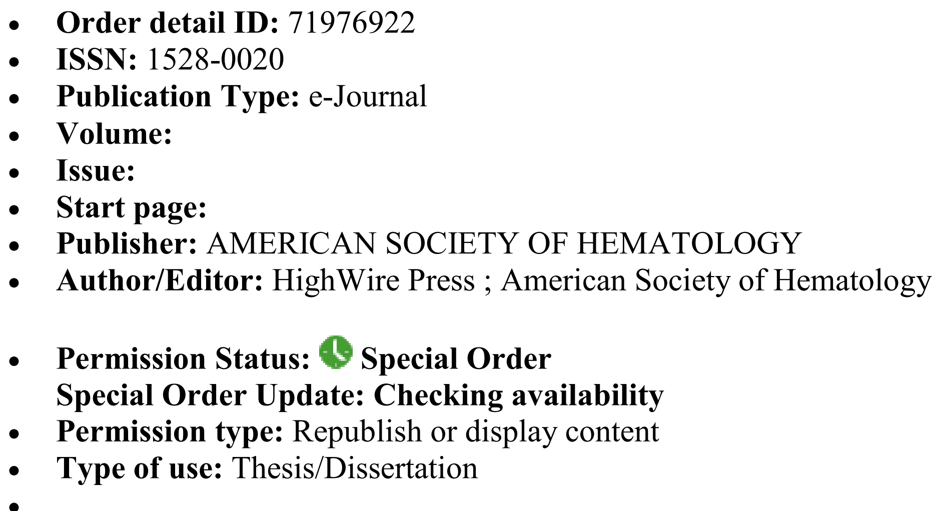
Customer: Rachid Abaji
Account Number: 3001378765
Organization: Université de Montréal
Email: rachid.abaji@umontreal.ca

Order Details

Special Orders

Blood : journal of the American Society of Hematology

Billing Status:
Not Billed



- **Order detail ID:** 71976922
- **ISSN:** 1528-0020
- **Publication Type:** e-Journal
- **Volume:**
- **Issue:**
- **Start page:**
- **Publisher:** AMERICAN SOCIETY OF HEMATOLOGY
- **Author/Editor:** HighWire Press ; American Society of Hematology
- **Permission Status:** **Special Order**
  Special Order Update: Checking availability
- **Permission type:** Republish or display content
- **Type of use:** Thesis/Dissertation
-

Job Ticket: 501506154

American Society of Hematology LICENSE
TERMS AND CONDITIONS
Dec 16, 2019

This is a License Agreement between Rachid Abaji ("You") and American Society of Hematology ("American Society of Hematology") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by American Society of Hematology, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	4654610613410
License date	Aug 10, 2019
Licensed content publisher	American Society of Hematology
Licensed content title	Blood : journal of the American Society of Hematology
Licensed content date	Jan 1, 1946
Type of Use	Thesis/Dissertation
Requestor type	Academic institution
Format	Electronic
Portion	chart/graph/table/figure
Number of charts/graphs/tables/figures	1
The requesting person/organization is:	Rachid Abaji
Title or numeric reference of the portion(s)	Table S1. AIEOP-BFM ALL 2000 treatment phases (see in the foot notes differences of AIEOP Protocol)
Title of the article or chapter the portion is from	Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study
Editor of portion(s)	N/A
Author of portion(s)	N/A
Volume of serial or monograph.	N/A
Page range of the portion	Supplemental appendix, tables, and figures: pages 6, 7 & 8
Publication date of portion	2010
Rights for	Main product

Duration of use	Life of current edition
Creation of copies for the disabled	no
With minor editing privileges	yes
For distribution to	Canada
In the following language(s)	Original language of publication
With incidental promotional use	no
The lifetime unit quantity of new product	Up to 499
Title	Using Whole-Exome Sequencing Data in an Exome-Wide Association Study Approach to Identify Genetic Risk Factors Influencing Acute Lymphoblastic Leukemia Response: A Focus on Asparaginase Complications & Vincristine-Induced Peripheral Neuropathy
Institution name	Université de Montréal
Expected presentation date	Aug 2019
Billing Type	Invoice
Billing Address	Rachid Abaji 3175 chemin de la Côte-Ste-Catherine Montreal, QC H3T 1C5 Canada Attn: Rachid Abaji
Total (may include CCC user fee)	0.00 USD

Terms and Conditions

Annex II

List of other scientific contributions

List of other scientific publications produced during Ph.D. studies not included in the body of the thesis

- V Gagné, A Aubry-Morin, M Plesa, **R Abaji** et al. Genes identified through genome-wide association studies of osteonecrosis in childhood ALL patients. Pharmacogenomics, 2019.

- BO Wolthers, TL Frandsen, CJ Patel, **R Abaji**, et al. Trypsin Encoding PRSS1-PRSS2 Variation Influence the risk of Asparaginase-associated Pancreatitis in Children with Acute Lymphoblastic Leukemia: a Ponte di Legno Toxicity Working Group Report. Haematologica 2018.

- I Goyer, M Iseppon, C Thibault, **R Abaji**, et al. Lactic Acidosis with Chloramphenicol Treatment in a Child with Cystic Fibrosis. Journal of population therapeutics and clinical pharmacology, 2017.

- **R Abaji**, M Krajinovic. Current perspective on pediatric pharmacogenomics. Expert Opinion on Drug Metabolism & Toxicology, 2016.

- DHFR (dihydrofolate reductase) Gene. M Krajinovic, **R Abaji**, B Sharif-Askari. Atlas of Genetics and Cytogenetics in Oncology and Haematology, 2015.

Annex III

List of Scholarships & Prizes

List of Prizes & Scholarships awarded during Ph.D. studies

➢ Prizes

- Best Poster Presentation (33e Congrès des étudiants du CHU Sainte-Justine, 2018)
- Best Poster Presentation (BioConnect, McGill University, 2017)
- Best Poster Presentation (Research Celebration Day of the Cole Foundation, 2017)
- 2nd Best Oral Presentation (32e Congrès des étudiants du CHU Sainte-Justine, 2017)
- Best Poster Presentation (BioConnect, McGill University, 2016)
- Best Poster Presentation (Journée Gabriel L. Plaa, 2016)

➢ Bursaries and Scholarships

- End of Doctoral Studies Scholarship from the Faculty of Higher and Post-Doctoral Education at the University of Montreal (FESP, 2018).
- Fellowship of the Cole Foundation (2016-2018).
- Travel Bursary for the 9th Summer School on Medicine in Barcelona from the University Drug Research Group (DAI-GRUM, 2017).
- Fellowship of the Network of Applied Medical Genetics (RMGA, 2016)
- Recruitment Scholarship of the Faculty of Medicine/University of Montreal (2015).