

Université de Montréal

# **Profilage Moléculaire de la Scoliose Idiopathique de l'Adolescent**

par

Qing Yuan

Département de biochimie

Faculté de médecine, Université de Montréal

Mémoire présenté à la Faculté des études supérieures et postdoctorales  
en vue de l'obtention du grade de maîtrise en Sciences (M.Sc)

Janvier 2013

© Qing Yuan, 2013

Université de Montréal  
Faculté des études supérieures et postdoctorales

Ce mémoire est intitulé :

Profilage Moléculaire de la Scoliose Idiopathique de l'Adolescent (SIA)

Présenté par :  
Qing Yuan

a été évalué par un jury composé des personnes suivantes :

André Tremblay, président-rapporteur  
Stéphane Roy, membre du jury  
Alain Moreau, directeur de recherche

## Résumé

**La scoliose idiopathique de l'adolescent (SIA)** est la déformation la plus fréquente en orthopédie pédiatrique. C'est une maladie qui génère des déformations complexes du rachis, du thorax et du bassin affectant plus sévèrement et majoritairement les filles à la puberté. Malgré de nombreuses années de recherches approfondies dans le domaine de la SIA, la cause n'a toujours pas été résolue.

**OBJECTIFS :** Il a été démontré qu'il existe une dysfonction de la signalisation de la mélatonine par les protéines G inhibitrices (Gi) dans les ostéoblastes isolés de patients atteints de SIA. Ceci a conduit à une stratification des patients en trois groupes fonctionnels. Le but de ce projet de maîtrise est d'établir les profils d'expression moléculaire correspondant à chacun des groupes fonctionnels. Les objectifs spécifiques sont d'identifier les gènes responsables de l'initiation de la SIA et du défaut différentiel observé dans la signalisation des récepteurs couplés aux protéines Gi.

**MÉTHODES :** Des ARNs ont été préparés à partir d'ostéoblastes isolés de patients atteints de SIA sélectionnés pour chaque groupe fonctionnel et comparés aux ARNs obtenus d'ostéoblastes de sujets sains. En plus, des cellules sanguines (PBMCs) isolées d'une paire de jumelles monozygotes discordantes pour la scoliose ont été étudiées pour comparer leur profil d'expression génétique. Nous avons utilisé des biopuces à ADN contenant un ensemble de 54,000 sondes permettant l'analyse du niveau d'expression de plus de 47,000 transcrits représentant 30,000 gènes humains bien caractérisés (Affymetrix, GeneChip® Human gene 1.0 ST array). Les gènes retenus ont par la suite été validés par qPCR sur un plus grand nombre de patients afin de tester la spécificité des profils d'expression pour chaque groupe de patients à partir des cellules osseuses dérivées lors de la chirurgie.

**RÉSULTATS:** Le profilage moléculaire proposé permettra d'établir les fondements moléculaires de la SIA selon le test fonctionnel développé par le Dr Moreau et son équipe et d'identifier de nouveaux gènes associés à la SIA. Ce projet a permis de mettre en évidence des gènes possiblement impliqués dans le défaut de signalisation associé aux protéines Gi communs aux trois groupes, ainsi que des gènes spécifiques à certains groupes, pouvant contribuer au développement de certaines co-morbidités et/ou au risque d'aggravation des déformations rachidiennes. L'étude préliminaire des jumelles monozygotes discordantes pour la scoliose a mis en évidence un nombre limité de gènes possiblement associés au défaut de signalisation observé chez la jumelle scoliotique, gènes dont l'expression pourrait être influencée par des modifications d'ordre épigénétique liées à l'environnement.

**CONCLUSION:** Les données obtenues par des approches de transcriptomiques dans le cadre de ce projet soutiennent notre méthode de stratification fonctionnelle des patients SIA et ont conduit à l'identification de nouveaux gènes associés à la SIA. Ces résultats jettent un éclairage nouveau sur la SIA et contribuent à une meilleure compréhension des mécanismes et facteurs impliqués dans l'étiologie de la SIA. À cet égard, nos résultats permettront éventuellement d'identifier des cibles thérapeutiques potentielles et des traitements personnalisés compte tenu des profils moléculaires distincts observés pour chaque groupe de patients.

**Mots clés :** Scoliose, gènes, endophénotype, microarray, expression, signalisation, protéine Gi, épigénétique, étiopathogenèse.

## **Abstract**

**The adolescent idiopathic scoliosis (AIS)** is the most common deformity in paediatric orthopaedics. It is a disease that generates complex deformations of the spine, thorax and pelvis and severely affecting mainly girls at puberty. Despite many years extensive research into the field of AIS, the cause of this disorder has still not been resolved.

**OBJECTIFS:** We have demonstrated a differential melatonin signaling dysfunction through G inhibitory (Gi) proteins in osteoblasts isolated from AIS patients. This led to a stratification of AIS patients into three distinct functional groups. The goal of this project is to establish molecular expression profiles corresponding to each functional group. The specific objective is to identify genes responsible for the initiation of the AIS and the differential dysfunction in the signaling of Gi protein-coupled receptors.

**METHODES:** Osteoblasts isolated from AIS patients of each functional subgroup were selected and compared to osteoblasts obtained from healthy controls. In addition, blood cells (PBMCs) isolated from a pair of monozygotic discordant twins of scoliosis was studied to compare their gene expression profiles. We used DNA chips with 54,000 probe sets that allow analysis of expression level of over 47,000 transcripts and variants from over 30,000 well-characterized human genes (Affymetrix, GeneChip® Human gene 1.0 ST array). The identified genes were subsequently validated by qPCR in a larger number of patients in order to test the specificity of expression profiles for each group of patients using bone cells derived during surgery.

**RESULTS:** The molecular profiling will establish the molecular basis of AIS according to the functional test developed by Moreau et al., and identify new genes associate with AIS. This project may allow us to identify genes involved in the dysfunction of Gi protein signaling which present in all three groups of patients. Also genes associated specifically with a particular group, can contribute to the development of certain AIS co-morbidities and/or to a risk of aggravation

of spine deformities. The preliminary study of monozygotic discordant twins of scoliosis revealed a limited number of genes potentially associated with the signaling defect observed in the scoliotic twin, the gene expression could be influenced by epigenetic changes related to the environment.

**CONCLUSION:** The data obtained by transcriptomic approaches further support our functional method of AIS patients' stratification, also allowed the identification of novel genes associated with AIS. Such results would provide us a better understanding of the mechanisms and factors involved in the etiology of AIS. In this regard, our results would help to identify potential and specific therapeutic targets and personalized treatments based on the distinct molecular profiles observed in each patient group.

**Keywords:** Scoliosis, gene, endophenotype, Microarray, expression, signalization, Gi protein, epigenetic, etiopathogenesis.

# Table des matières

|   |           |
|---|-----------|
| Résumé .....  | i         |
| Abstract .....  | iii       |
| Table des matières .....  | v         |
| Liste des figures .....   | viii      |
| Liste des abréviations .....  | ix        |
| Remerciements .....   | xiii      |
| <b>CHAPITRE 1 : INTRODUCTION .....</b>  | <b>1</b>  |
| <b>1. Introduction sur la scoliose idiopathique .....</b>                     | <b>2</b>  |
| <b>2. Phénotype de la Scoliose Idiopathique .....</b>                         | <b>2</b>  |
| <b>2.1 Âge.....</b>   | <b>4</b>  |
| <b>2.2 La morphologie et la magnitude des déformations rachidiennes .....</b> | <b>4</b>  |
| <b>2.3 Prévalence.....</b>  | <b>8</b>  |
| <b>2.4 Le traitement et la prévention.....</b>                                | <b>9</b>  |
| <b>3. Étiopathogenèse .....</b>   | <b>11</b> |
| <b>3.1 Facteurs génétiques.....</b>   | <b>12</b> |
| <b>3.1.1 Analyse de liaisons génétiques .....</b>                             | <b>12</b> |
| <b>3.1.2 Études d'association.....</b>  | <b>13</b> |
| <b>3.1.3 Études d'association sur le génome entier (GWAS).....</b>            | <b>16</b> |
| <b>3.1.4 Gènes candidats.....</b>   | <b>16</b> |
| <b>3.2 Dysfonctionnement métabolique .....</b>                                | <b>17</b> |
| <b>3.2.1 Rôle de la mélatonine dans l'étiologie de la SIA.....</b>            | <b>17</b> |
| <b>3.2.2 Le défaut de la signalisation de la mélatonine dans la SIA.....</b>  | <b>19</b> |
| <b>3.3 Autres contributions hormonales à la SIA.....</b>                      | <b>23</b> |
| <b>3.3.1 IGF-1 .....</b>  | <b>23</b> |
| <b>3.3.2 Œstrogènes .....</b>   | <b>24</b> |
| <b>3.3.3 Rôle de la Calmoduline.....</b>                                      | <b>24</b> |
| <b>3.3.4 Ostéopontine .....</b>   | <b>25</b> |
| <b>3.4 Autres facteurs .....</b>  | <b>26</b> |

|  |            |
|--|------------|
| 3.4.1 Facteurs environnementaux .....  | 26         |
| 3.4.2 Facteurs épigénétiques .....   | 26         |
| <b>4. Co-morbidités .....</b>  | <b>29</b>  |
| 4.1 Ostéopénie.....  | 29         |
| 4.2 L'anomalie des plaquettes .....  | 30         |
| <b>5. Les endophénotypes de la SIA .....</b>   | <b>32</b>  |
| <b>6. Hypothèses et Objectifs.....</b>   | <b>35</b>  |
| <b>CHAPITRE 2 : ARTICLE I .....</b>  | <b>37</b>  |
| <b>Genome-wide expression profiling of idiopathic scoliosis endophenotypes.....</b>            | <b>38</b>  |
| <b>La contribution des auteurs : .....</b>   | <b>38</b>  |
| <b>Abstract.....</b>   | <b>41</b>  |
| <b>Introduction .....</b>  | <b>42</b>  |
| <b>Materials and methods.....</b>  | <b>44</b>  |
| <b>Results .....</b>   | <b>47</b>  |
| <b>Discussion.....</b>   | <b>51</b>  |
| <b>Conclusions .....</b>   | <b>62</b>  |
| <b>Tables and Figures .....</b>  | <b>64</b>  |
| <b>References .....</b>  | <b>102</b> |
| <b>CHAPITRE 3 : ARTICLE II.....</b>  | <b>114</b> |
| <b>Gene expression analysis of monozygotic twins discordant for idiopathic scoliosis .....</b> | <b>115</b> |
| <b>La contribution des auteurs : .....</b>   | <b>115</b> |
| <b>Abstract.....</b>   | <b>118</b> |
| <b>Introduction .....</b>  | <b>119</b> |
| <b>Materials and Methods .....</b>   | <b>121</b> |
| <b>Results .....</b>   | <b>123</b> |
| <b>Discussion.....</b>   | <b>124</b> |
| <b>Conclusions .....</b>   | <b>129</b> |
| <b>Tables and Figures .....</b>  | <b>130</b> |
| <b>References .....</b>  | <b>137</b> |

|   |            |
|---|------------|
| <b>CHAPITRE 4 : DISCUSSION.....</b>   | <b>144</b> |
| <b>4.1 Profilage moléculaire des patients SIA en fonction de leur endophénotype.....</b>  | <b>148</b> |
| 4.1.1 Contribution des résultats du profilage moléculaire dans la compréhension des aspects cliniques de la SIA .....   | 150        |
| 4.1.2 Identification des gènes impliqués dans le métabolisme osseux.....  | 151        |
| 4.1.3 Concordance des résultats de profilage moléculaire avec les études publiées par d'autres chercheurs .....   | 152        |
| 4.1.4 Discrimination des gènes dont l'expression contribue au pathomécanisme de la SIA <i>versus</i> ceux dont l'expression est modifiée en conséquence de la maladie ..... | 153        |
| 4.1.5 Contributions possibles d'autres tissus ou types de cellules dans la SIA .....  | 154        |
| <b>4.2 Études de l'expression génique chez jumelles identiques discordantes .....</b>   | <b>154</b> |
| 4.2.1 Travaux futur .....   | 158        |
| 4.2.2 Importance des méthodes d'analyses statistiques .....   | 158        |
| <b>CHAPITRE 5 : CONCLUSION .....</b>  | <b>160</b> |
| <b>CHAPITRE 6 : RÉFÉRENCES .....</b>  | <b>163</b> |

## Liste des figures

|   |            |
|---|------------|
| <b>Figure 1: Sujet sain vs Patient SIA.....</b>   | <b>3</b>   |
| <b>Figure 2: SI chez une jeune adolescente.....</b>   | <b>3</b>   |
| <b>Figure 3: Les trois plans dans la déformation complexe de la colonne vertébrale. ....</b>  | <b>5</b>   |
| <b>Figure 4: Les rayons X d'une jeune fille de 14 ans avec une scoliose de 90°.....</b>   | <b>5</b>   |
| <b>Figure 5: Les quatre principaux patrons de courbure de la SIA.....</b>   | <b>7</b>   |
| <b>Figure 6: La méthode pour mesurer l'angle de Cobb.....</b>   | <b>7</b>   |
| <b>Figure 7: Arbre généalogique d'une famille présentant des cas de scoliose.....</b>   | <b>15</b>  |
| <b>Figure 8: La transduction du signal de la mélatonine. .....</b>  | <b>21</b>  |
| <b>Figure 9 : Diagramme de Venn sur la relation entre la variation génétique, épigénétique et la maladie.....</b>                                 | <b>28</b>  |
| <b>Figure 10 : Système cible dans l'action de la mélatonine et symptômes observés chez les patients scoliotiques.....</b>                         | <b>31</b>  |
| <b>Figure 11: Évaluation de l'effet inhibiteur de la mélatonine sur l'accumulation intracellulaire d'AMPc dans les ostéoblastes humains. ....</b> | <b>33</b>  |
| <b>Figure 12: Dysfonctionnement de la signalisation démontré par la CDS. ....</b>   | <b>34</b>  |
| <b>Figure 13. Les endophénotypes, des outils pour étudier les maladies complexes.....</b>   | <b>147</b> |

## Liste des abréviations

SI : Scoliose Idiopathique

SIA : Scoliose Idiopathique de l'Adolescent

AIS : Adolescent Idiopathic Scoliosis

SRS : Scoliosis Research Society

SNPs : Single Nucleotide Polymorphism

*MTNR1B* : Melatonin 1B Receptor

*CDH7* : Chromodomain Helicase DNA-binding Protein

*TPH1* : Tryptophan Hydroxylase 1

*COL1A1* : Collagen Type I Gene

*COL1A2* : Collagen Type II Gene

*MATN1* : Matrillin-1 Gene

*ESR1(alpha)* : Estrogen Receptor α

*ESR1(beta)* : Estrogen Receptor β

*IL6* : Interleukin-6

MMPS : Matrix Metalloproteinases

*MMP1* : Matrix Metalloproteinase-1

*MMP2* : Matrix Metalloproteinase-2

*MMP3* : Matrix Metalloproteinase-3

*MMP13* : Matrix Metalloproteinase-13

*SNTG1* : γ 1-Syntrophin

GWAS : Genome Wide Association Study

*CHL1* : L1 Cell Adhesion Molecule

*DSCAM* : Down Syndrome Adhesion Molecule

*ROBO3* : Roundabout, Axon Guidance Receptor, Homolog 3

*NAT* : N-Acetyltransferase

*MT1* : Melatonin MT1 Receptor

*MT2* : Melatonin MT2 Receptor

*Gi* : Guanine Nucleotide-Binding inhibitory Protein

*AMPc* : AMP cyclique

*CDS* : Spectroscopie Cellulaire Dielectrique

*IGF1* : Insulin-Like Growth Factor-1

*IFG-R1* : Insulin-Like Growth Factor type 1 Receptor

*OPN* : Osteopontin

*SPP1* : Secreted Phosphoprotein 1

*BMD* : Bone Mineral Density

*ECM* : Extracellular Matrix

*MKX* : Mohawk homeobox

*SLPI* : Secretory Leukocyte Peptidase Inhibitor

*CLGN* : Calmegin

*LRRC17* : Leucine Rich Repeat containing 17

*RANKL* : Receptor Activator of NF-kappa-B Ligand

*CHRD1* : chordin-like 1

*ITGB3* : Integrin, Beta 3

*CHI3L1*: Chitinase 3 Like 1

*COMP* : Cartilage Oligomeric Matrix Protein

*HLA-DQBI* : Histocompatibility complex, Class II, DQ Beta 1

*MSLN* : Mesothelin

*TLR3* : Toll-Like Receptor 3

*THBS1* : Thrombospondin 1

*DDR* : Discoidin Domain Receptor Tyrosine Kinase

MHC : Major Histocompatibility Complex

FDR : False Discovery Rate

*PDE1A* : Phosphodiesterase 1A, Calmodulin-dependent

*BDKRB1* : Bradykinin Receptor B1

*BDKRB2* : Bradykinin Receptor B2

*CCRL1* : Chemokine (C-C motif) Receptor-Like 1

*CCKAR* : Cholecystokinin A Receptor

*PTGFR* : Prostaglandin F Receptor

*F2RL1* : Coagulation Factor II (thrombin) Receptor-Like 1

*Je dédie ce mémoire à ma mère et à mon père, qui m'ont encouragés tout au long de mon parcours, merci pour votre soutien moral et financier.*

*Je dédie aussi ce mémoire à mon grand-père, merci pour m'avoir guidé dans mon parcours scientifique, depuis mon enfance jusqu'à ce jour, il a été une inspiration pour moi et m'a toujours encouragée dans tout ce que j'ai entrepris.*

*Merci pour tout !*

## **Remerciements**

Un remerciement spécial à mon directeur de recherche, Dr Alain Moreau, pour m'avoir accueillie au sein de son laboratoire, pour sa patience et son encouragement continu au cours de ma maîtrise et qui est pour moi un exemple de gentillesse, d'intelligence et d'enthousiasme pour la science. Je lui souhaite le meilleur dans sa carrière et dans la vie. Un Grand Merci.

Je remercie Dr Kristen Fay Gorman pour ses conseils et sa patience, elle qui m'a donnée beaucoup de conseils précieux pour la synthèse de mon mémoire. Merci également pour nos passionnantes discussions.

Je tiens à remercier également Dr Marie-Yvonne Akoume pour son support scientifique et psychologique. C'est un grand privilège d'avoir travaillé avec Marie-Yvonne et cela m'a permis d'enrichir mon projet de recherche en utilisant la technique du CDS.

Je remercie aussi Dr Cédric Julien pour m'avoir supervisée durant une partie de mon projet de maîtrise. Merci pour vos judicieux conseils scientifiques. Je remercie également Dr Dashen Wang qui a toujours été là pour m'encourager.

Je remercie spécialement Anita Franco et Saadallah Bouhanik pour leurs supports techniques et d'avoir toujours été là pour moi notamment dans les moments plus difficiles. Je remercie aussi les autres membres du laboratoire Nancy Karam, Lauriane Bernard qui m'ont rendue le milieu de travail si agréable. Vous occuperez une place spéciale dans mes souvenirs de mon séjour au Canada.

Finalement, je tiens à remercier le Centre de recherche de CHU Saint-Justine pour le support et le magnifique environnement de recherche dont j'ai pu profiter.

## **CHAPITRE 1 : INTRODUCTION**

## **1. Introduction sur la scoliose idiopathique**

La scoliose est une déformation de la colonne vertébrale, souvent décrite comme une déformation rachidienne latérale sous ‘une forme de S’ (Figure 1). La scoliose idiopathique (SI) représente la forme la plus commune des déformations en orthopédie pédiatrique. Environ 80 % des courbes structurelles sont dites idiopathiques (IS [MIM 181800]), indiquant que la courbe anormale se développe pour des raisons inconnues. La SI est généralement observée au cours de périodes de croissance rapide au début de l’adolescence durant la période de la puberté. La SI est une maladie génétique complexe de haute prévalence et extrêmement variable qui affecte en moyenne 4% de la population mondiale en pédiatrie. Elle affecte majoritairement les filles et se manifeste comme une déformation tridimensionnelle complexe de la colonne vertébrale qui implique des changements dans le plan frontal, sagittal et transversal (Figure 2). Malgré des recherches intenses, les causes de cette maladie demeurent encore inconnues mais il est accepté de nos jours que cette maladie soit probablement multifactorielle avec une prédisposition génétique (*Blank, Raggio et al. 1999; Machida 1999; Lowe, Edgar et al. 2000; Burwell 2003; Wise, Gao et al. 2008*).

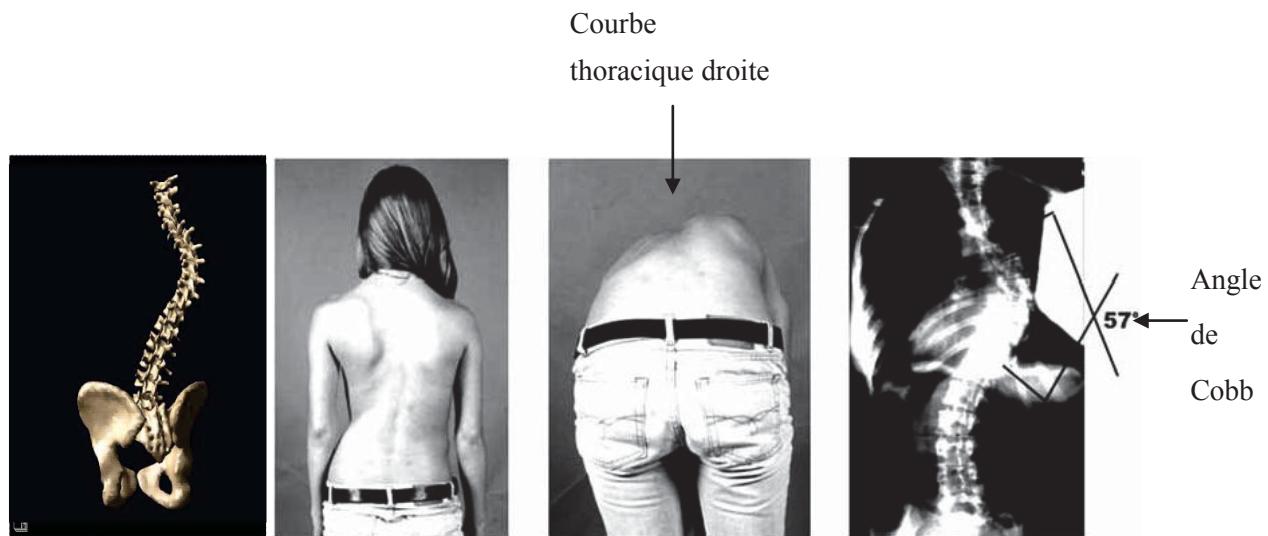
## **2. Phénotype de la Scoliose Idiopathique**

Le phénotype de la scoliose idiopathique est caractérisé par une variabilité de la morphologie, de l’amplitude/magnitude de la courbe, de l’âge d’apparition, du taux de progression de la courbure, ainsi que de la propension de la courbure à la progression, à la stabilisation ou à la régression.



**Figure 1: Sujet sain vs Patient SIA.**

Le sujet sain sur la gauche présente une vertèbre normale alors que le patient SI sur la droite présente une déformation de la colonne.



**Figure 2: SI chez une jeune adolescente.**

La radiographie aux rayons X sur la droite révèle une courbe thoracique droite de 57 degrés mais sans anomalie vertébrale ou autres.

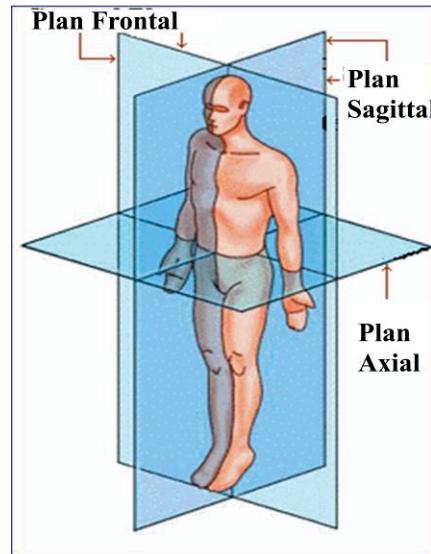
*Figure adaptée de (Wise, Gao et al. 2008)*

## 2.1 Âge

La scoliose idiopathique peut être classée en trois groupes selon l'âge d'apparition : 0-3 ans (infantile), 3-9 ans (juvénile), et 10-18 ans (adolescent) (*James 1954*). Certains chercheurs proposent simplement de distinguer deux catégories : l'apparition précoce de la scoliose (moins de huit ans) et l'apparition tardive de la scoliose (huit ans et plus). Deux catégories basées sur une prédisposition héréditaire existante (*Wynne-Davies 1968*). La scoliose idiopathique de l'adolescent (SIA) se développe évidemment durant l'adolescence, surtout chez les jeunes filles, et progresse typiquement pendant la période de croissance. La SIA représente approximativement 80% des cas de scoliose idiopathique. Concernant les scolioses infantiles et juvéniles, la probabilité que la courbure régresse avant que l'enfant n'ait atteint sa maturité est plus élevée (*Warner 2001; Dickson 2001*). Outre une étiologie inconnue, la question se pose sur la propension de certains enfants à développer une scoliose progressive tandis que chez d'autres, elle se stabilisera ou régressera spontanément ou non. Dans ce contexte, la caractérisation des profils moléculaires pourrait nous aider à identifier les facteurs de progression et à mieux comprendre leur interaction.

## 2.2 La morphologie et la magnitude des déformations rachidiennes

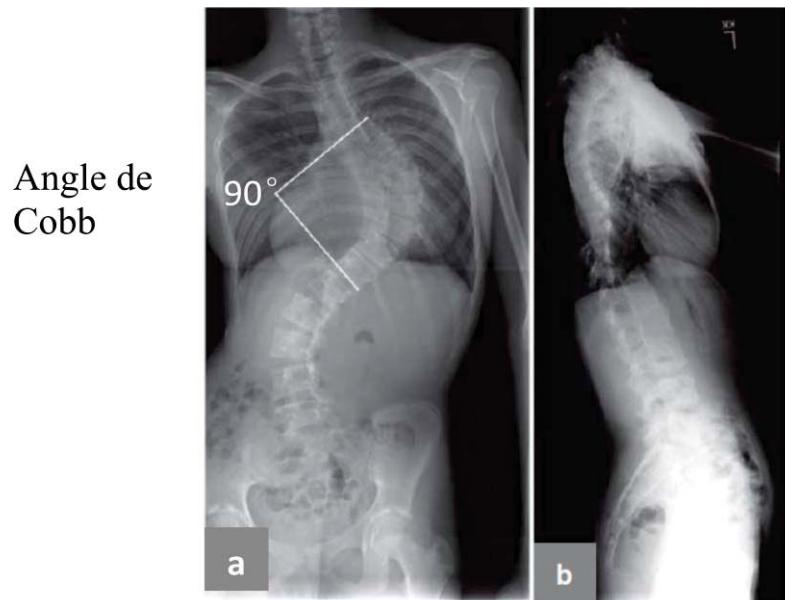
La scoliose est définie comme une déformation tridimensionnelle de la colonne vertébrale (Figure 3) (*Deacon, Flood et al. 1984; Perdriolle and Vidal 1987*).



**Figure 3:** Les trois plans dans la déformation complexe de la colonne vertébrale.

Représentation des trois plans axial, sagittal et frontal.

*Figure adaptée de Baron S. Lonner, Scoliosis: A Three-Dimensional Problem*



**Figure 4:** Les rayons X d'une jeune fille de 14 ans avec une scoliose de 90°.

(a) Vue postéro-antérieure avec un angle de Cobb indiquée. (b) Vue latérale.

*Figure adaptée de (Trobisch, Suess et al. 2010)*

Le patron de courbure de la scoliose est généralement décrit à l'emplacement de cette dernière (à l'apex de la courbure principale), tel qu'illustré à la figure 5 ci-dessous : la convexité des points de la courbe vers la droite ou vers la gauche, si la courbe présente chez deux endroits, chaque courbe doit être décrite et mesurée. On utilise la méthode de l'angle de Cobb pour mesurer le degré de la courbe (Figure 6).

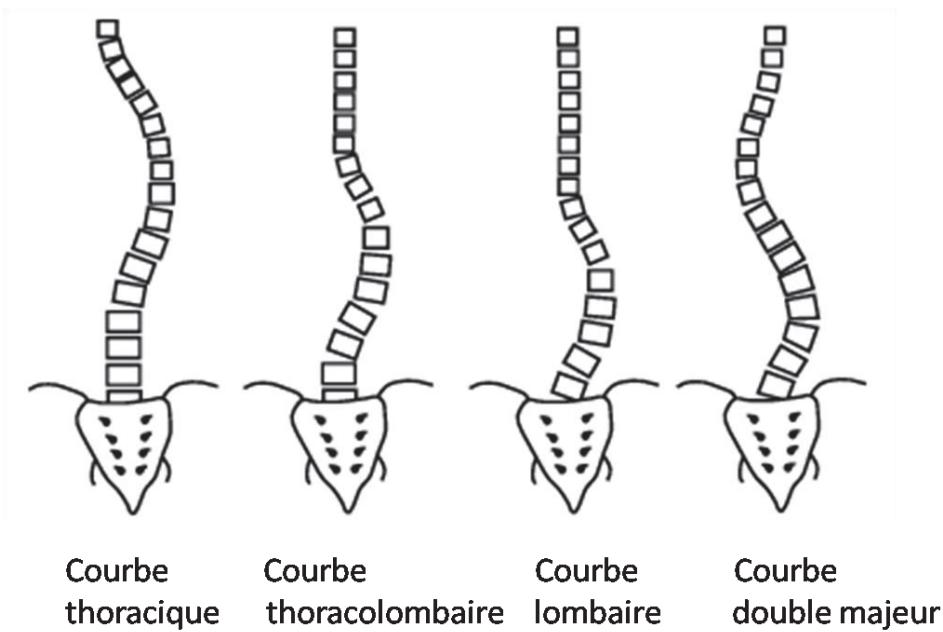


Figure 5: Les quatre principaux patrons de courbure de la SIA.

*Figure adaptée de <http://www.rad.washington.edu/>*

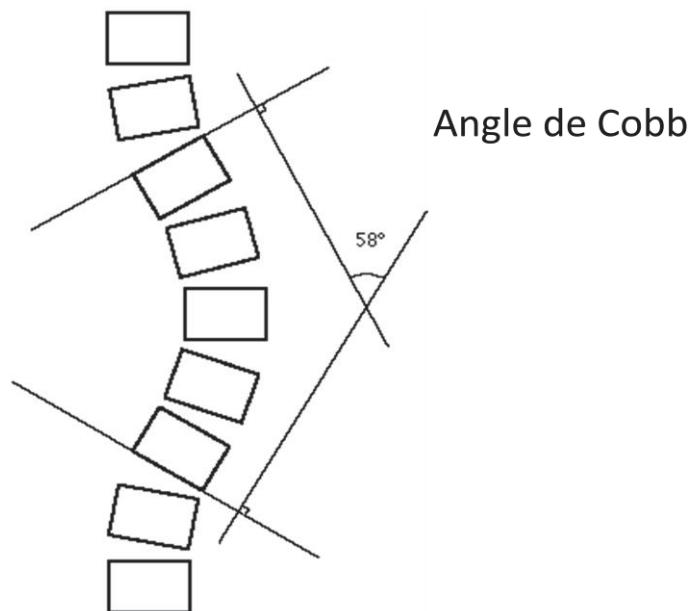


Figure 6: La méthode pour mesurer l'angle de Cobb.

*Figure adaptée de*

*<http://www.rad.washington.edu/academics/academic-sections/msk/teaching-materials/online-musculoskeletal-radiology-book/scoliosis>*

## 2.3 Prévalence

La scoliose idiopathique fait partie des maladies musculo-squelettiques qui sont rarement mortelles. De ce fait, elle ne bénéficie pas de la même visibilité que les pathologies cardiaques, respiratoires et que le cancer qui sont quand même moins répandues dans la population pédiatrique. Selon la ‘Scoliosis Research Society’ (SRS), la scoliose est une déviation latérale de plus de 10 degrés mesurés suite à une radiographie dans le plan vertical et antéropostérieur du rachis. Tout autour du globe, la population adolescente atteinte de SIA se compose d'environ 2.5 % dans la catégorie d'âge de 10-16 ans. La SIA est relativement fréquente, avec une large prévalence étudiée par de nombreuses études dans la littérature (*Shands and Eisberg 1955; Cowell 1970; Rogala, Drummond et al. 1978; Willner 1982; Burwell, James et al. 1983; Dickson 1983; Nissinen, Heliovaara et al. 1993*). D'autres études ont montré que la scoliose a une prévalence qui varie entre 2 et 10 % chez les écoliers (*Cowell 1970; Rogala, Drummond et al. 1978; Dickson 1983; Nissinen, Heliovaara et al. 1993*). En 2004, 1.26 millions de patients ont nécessité des soins de santé pour la déformation de la colonne vertébrale aux États-Unis. De ce nombre, environ 134,500 patients ont été hospitalisés, 93 % des personnes hospitalisées ont eu une scoliose diagnostiquée et plus de la moitié était âgée de moins de 18 ans (*American Academy of Orthopedic Surgeons, 2011*).

Les coûts associés aux traitements des scolioses sont très élevés pour les systèmes de santé. En 2004, le coût estimé du traitement des patients hospitalisés avec un diagnostic de scoliose idiopathique aux États-Unis seul était de plus de 3 milliards de dollars (*American Academy of Orthopedic Surgeons, 2008*), tandis que ce chiffre s'élevait à 373.3 millions d'euros en Europe (*American Academy of Orthopedic Surgeons, 2011*).

## 2.4 Le traitement et la prévention

Jusqu'à maintenant, il n'y a pas de tests spécifiques pour la détection et le pronostic de la SIA. Cependant, il y a des données fiables sur le risque estimé de progression de la scoliose, détaillées dans le tableau 1, proposées par la Scoliosis Research Society (*Scoliosis Research Society, Chicago, Illinois, USA*). Ces données ont permis de mettre en évidence que les courbes qui sont les plus susceptibles d'évoluer, nécessitent un traitement actif allant de la physiothérapie au corset et ultimement à la chirurgie pour les déformations de plus de 45°. Le corset vise à contrôler l'évolution des courbures pendant la phase de croissance via l'application de forces extérieures. Toutefois, le corset ne peut pas guérir la scoliose.

Dans le cas de courbes ayant un angle de Cobb supérieur à 45 degrés, il y a un recours à la chirurgie pour la correction de la déformation du rachis. Le principe de base consiste à appliquer un ensemble de forces par l'utilisation d'une instrumentation spécifique, afin de corriger la déformation et d'en stopper sa progression.

Actuellement, le développement de la scoliose ne peut être prévenu. Les efforts sont concentrés sur la mise en place de systèmes de diagnostique et/ou de pronostic précoce, afin de traiter les patients plus tôt et plus rapidement (*Fong, Lee et al. 2010*).

**Tableau 1 : Risque de progression de la SIA**

| Risque de progression de la SIA       |              |              |                   |
|---------------------------------------|--------------|--------------|-------------------|
| Degré de la courbe<br>(Angle de Cobb) | Âge<br>10-12 | Âge<br>13-15 | Âge<br>plus de 16 |
| <20°                                  | 25%          | 10%          | 0%                |
| 20-30°                                | 60%          | 40%          | 10%               |
| 30-60°                                | 90%          | 70%          | 30%               |
| >60°                                  | 100%         | 90%          | 70%               |

Le tableau 1 reflète la probabilité de progression de la scoliose non traitée développée par la Scoliosis Research Society. L'amplitude de la courbe initiale *vs* âge.

*Tableau adapté de (Weinstein, Dolan et al. 2008).*

### 3. Étiopathogenèse

Pratiquement toutes les structures du corps ont été incriminées dans la pathogenèse de la SI. Dans ce contexte, plusieurs modèles animaux ont été créés pour mieux comprendre cette maladie. Néanmoins, aucune cause unique n'a été identifiée à ce jour, et il est de plus en plus accepté que la SI soit une condition multifactorielle (*Lowe, Edgar et al. 2000*). Plusieurs hypothèses pour expliquer l'étiologie de cette pathologie ont donc été proposées. Différentes pistes biologiques impliquant la déficience du système endocrinien (pouvant concerner certaines hormones comme la mélatonine, les œstrogènes, l'hormone de croissance, la testostérone ou l'insuline) ont été explorées (*Misol, Ponseti et al. 1971; Skogland and Miller 1980; Fagan, Kennaway et al. 1998; Brodner, Krepler et al. 2000; Barrios, Cortes et al. 2011*). La présence de possibles anomalies affectant les tissus osseux et conjonctifs (*Zorab, Clark et al. 1971; Shapiro, Burn et al. 1989; Miller, Mims et al. 1996; Guo, Chau et al. 2005; Man, Wong et al. 2011; Wang, Yeung et al. 2011*). Des désordres au niveau du tissu musculaire ont également été investiguée (*Fidler and Jowett 1976; Reuber, Schultz et al. 1983; Gonyea, Moore-Woodard et al. 1985; Avikainen, Rezasoltani et al. 1999; Cheung, Halbertsma et al. 2005; Martinez-Llorens, Ramirez et al. 2010*). Certains chercheurs ont également proposé que cette pathologie était causée par un dysfonctionnement du système nerveux central et/ou périphérique, ainsi que par des anomalies perturbant le système vestibulaire (*Fisk and Bunch 1979; Pincott, Davies et al. 1984; Lonstein and Renshaw 1987; Manzoni and Miele 2002*). Des désordres au niveau des plaquettes (*Kahmann, Donohue et al. 1992; Lowe, Lawellin et al. 2002*) de même qu'au niveau de la signalisation moléculaire de la mélatonine (*Lombardi, Akoume et al. 2011*) ont également été observés. Aujourd'hui avec les avancées dans le domaine de la génétique/génomique, les irrégularités génétiques associées à une déformation du rachis suscitent un intérêt particulier.

### **3.1 Facteurs génétiques**

Le rôle de facteurs héréditaires ou génétiques dans le développement de la scoliose idiopathique a largement été reconnu et accepté (*Garland 1934; Cowell, Hall et al. 1972; Beals 1973; Harrington 1977; Dickson 1992; Bell and Teebi 1995*). Les premières études des cas familiaux ont montré que les proches du premier degré de patients scoliotiques ont un risque accru de développer la maladie (*Robin and Cohen 1975*). Une méta-analyse des différentes études de jumeaux a révélé un taux de concordance de 73 % entre les jumeaux monozygotes contre 36 % pour les jumeaux dizygotes (*Kesling and Reinker 1997*). Selon une étude basée sur un questionnaire (*Andersen, Thomsen et al. 2007*), le taux de concordance par paire, en Finlande, est seulement de 13 % pour les jumeaux monozygotes et nul pour les paires de jumeaux dizygotes. L'ensemble de ces données souligne la contribution de facteurs génétiques mais soulève également l'importance de facteurs environnementaux (contribution épigénétique) dans l'étiologie de la SIA.

#### **3.1.1 Analyse de liaisons génétiques**

Avec l'avancement des techniques d'analyse génétique et génomique, des études de liaisons génétiques sur plusieurs familles ont révélées plusieurs loci liés à la SIA. En effet, plusieurs études réalisées sur des cas familiaux (ceux-ci représentent 15 % environ des cas de SIA) ont démontré l'existence de plusieurs loci de susceptibilité (*Miller 2007*) et d'autres (*Raggio, Giampietro et al. 2009*) ont permis de démontrer les aspects génétiques de la SIA. En 2005, Miller *et al.* (*Miller, Justice et al. 2005*) ont démontré l'existence de loci d'intérêt présents sur les chromosomes 6, 9, 16 et 17 ainsi que sur des régions secondaires des chromosomes 1, 3, 5,

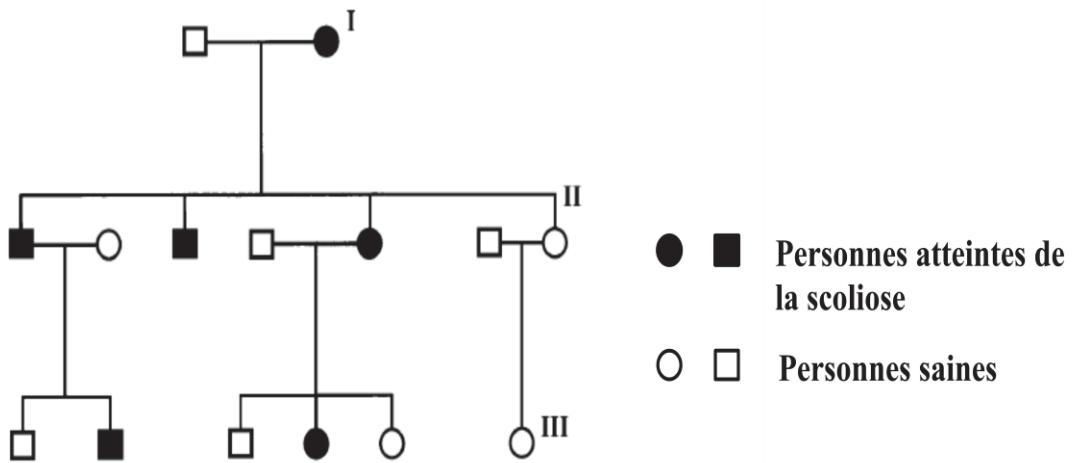
7, 8, 11, 12 et 19 (étude de 202 familles avec 1198 individus). Une étude récente a également mis en évidence un locus de susceptibilité sur le chromosome 18 (*Gurnett, Alaee et al. 2009*). Une forte hétérogénéité génétique est donc présentée dans ces études compte tenu de la multiplicité des loci et de l’implication de près de la moitié des chromosomes. La prédominance de la SIA chez les filles a conduit plusieurs groupes à examiner une possible transmission de la SIA par le chromosome X (*Justice, Miller et al. 2003*); alors que d’autres études favorisent plutôt une transmission autosomale (*Salehi, Mangino et al. 2002*); ou un modèle d’hérédité multifactorielle (Figure 7). L’ensemble de ces données génétiques obtenues dans différentes populations a mis en évidence un nombre croissant de loci de susceptibilité pour la SIA, ce qui rend difficile l’identification précise d’un ou plusieurs gènes candidats responsables de cette maladie (*Wise, Barnes et al. 2000; Miller, Schwab et al. 2001; Miller 2002; Alden, Marosy et al. 2006; Cheng, Tang et al. 2007*).

### 3.1.2 Études d’association

L’approche par “gène candidat” a été largement utilisée au cours de la dernière décennie. Des gènes candidats ont été sélectionnés sur la connaissance phénotypique observée chez les patients et l’implication de différents polymorphismes (SNPs) a été évaluée.

Parmi les gènes candidats investigués on y retrouve : MTNR1B (melatonin 1B receptor gene) (*Qiu, Tang et al. 2007*), CDH7 (chromodomain helicase DNA-binding protein 7) (*Gao, Gordon et al. 2007*), TPH1 (tryptophan hydroxylase 1) (*Wang, Wu et al. 2008*), COL1A1 (collagen type I gene), COL1A2 (collagen type II gene) (*Miller, Mims et al. 1996*), MATN1 (matrillin-1 gene) (*Montanaro, Parisini et al. 2006; Chen, Tang et al. 2009*), ESR1(α) (estrogen receptor α) (*Inoue,*

*Minami et al. 2002)*, ESR1(β) (estrogen receptor β) (*Zhang, Lu et al. 2009*), IL6 (interleukin-6), MMP3 ( matrix metalloproteinase-3) (*Aulisa, Papaleo et al. 2007*) et SNTG1 ( $\gamma$  1-syntrophin) (*Bashiardes, Veile et al. 2004*). Malheureusement, la majorité de ces études ont été réalisées sur un nombre limité de sujets et peinent dans bien des cas à atteindre des valeurs statistiquement significatives. De plus, aucune de ces études n'a été reproduite au sein des mêmes populations ou même dans d'autres ethnies ce qui ajoute à la controverse quant aux rôles proposés de ces facteurs dans l'étiopathogenèse de la SIA.



**Figure 7: Arbre généalogique d'une famille présentant des cas de scoliose.**

La figure 7 représente une famille contenant plusieurs membres atteints de la scoliose idiopathique. Les personnes touchées par la maladie, filles et garçons, sont présents à chaque génération (I, II et III), suggérant un mode de transmission autosomique dominant. Ici, les carrés représentent les hommes et les cercles, les femmes. Les individus en noir sont atteints et les individus en blanc sont sains.

*Figure adaptée de (Lowe, Edgar et al. 2000).*

### **3.1.3 Études d'association sur le génome entier (GWAS)**

Le développement des études d'association sur l'ensemble du génome (GWAS) fournit une image plus complète des gènes impliqués dans l'étiopathogénèse de la SIA. Une étude récente portant sur 419 familles a conduit au génotypage d'environ 327,000 polymorphismes (SNPs) et a démontré l'existence d'un SNP associé à la SIA sur le chromosome 3p26.3, à proximité du gène CHL1 (L1 Cell Adhesion Molecule). Un second signal a été détecté au niveau du gène DSCAM (Down Syndrome Adhesion Molecule) qui code pour une protéine de guidage axonal, appartenant à la même classe structurale que CHL1 et ROBO3 (Roundabout, Axon Guidance Receptor, Homolog 3). Ces deux gènes sont impliqués dans le guidage axonal, apportant une piste neurologique dans l'étiologie de la SIA (*Sharma, Gao et al. 2011*). D'ailleurs, le gène ROBO3 est muté dans un syndrome humain dont la scoliose est une des manifestations cliniques de ce syndrome (*MacDonald 2006; Abu-Amero 2008; Amouri, Nehdi et al. 2009; Abu-Amero, Kapoor et al. 2011*).

### **3.1.4 Gènes candidats**

Selon l'hypothèse d'une hétérogénéité génétique sous-jacente à la SIA, plusieurs gènes candidats ont été étudiés. Dans la revue publiée par Gorman *et al.*, de nombreux gènes candidats ont été classés par catégories fonctionnelles, reflétant ainsi différentes hypothèses concernant les systèmes biologiques impliqués dans le développement de la scoliose : les protéines structurales, le métabolisme osseux, la signalisation de la mélatonine et la croissance (*Gorman, Julien et al. 2012*).

La plupart de ces études ont été réalisées sur de petites cohortes et il est donc difficile de conclure si les associations observées, non reproduites dans différentes cohortes (de même nationalité ou sur des populations différentes) sont des conséquences de l'hétérogénéité génétique ou d'un biais de l'étude. La majorité des études de gènes candidats sont non concluantes, et les associations observées doivent encore être validées.

Autre que ces liens génétiques, un dysfonctionnement métabolique a aussi suscité un intérêt particulier.

## 3.2 Dysfonctionnement métabolique

### 3.2.1 Rôle de la mélatonine dans l'étiologie de la SIA

La mélatonine ou N-acétyl-5-méthoxytryptamine est une neuro-hormone de la famille des indolamines. Elle est produite en réponse à l'absence de lumière. La mélatonine est secrétée principalement par la glande pinéale, celle-ci agit comme une hormone endocrine (diffuse dans le sang). Celle produite par la rétine agit comme une hormone paracrine. La mélatonine joue un rôle dans la régulation du rythme circadien ou sommeil (*Cajochen et al., 2003*). Elle a aussi un rôle important dans le système immunitaire (*Bozenna et al., 2003*). Elle peut également agir sur le métabolisme du calcium (*Faria Poloni J et al., 2011*). Elle exerce plusieurs de ses actions physiologiques par l'interaction avec les récepteurs membranaires MT1 et MT2 qui appartiennent à la famille de récepteurs couplés aux protéines G (GPCR) (*Von Gall et al., 2002; Delagrange et al., 2008; Cardinali, D. P., et al., 2008*). L'hypothèse qu'une carence en mélatonine serait à la source de la SIA a suscité beaucoup d'intérêt et de controverses (*Bagnall, Raso et al. 1996; Hilibrand, Blakemore et al. 1996; Machida, Dubousset et al. 1996; Fagan, Kennaway et al.*

1998; Brodner, Krepler et al. 2000; Sadat-Ali, al-Habdan et al. 2000; Dubousset and Machida 2001; Fagan, Kennaway et al. 2009). Les études de Thillard *et al.* ont rapporté que la pinéalectomie induit une scoliose chez des poulets (*Thillard 1959*), produisant un phénotype similaire en plusieurs points à la maladie humaine (*Wang, Jiang et al. 1997*). La pinéalectomie de rats bipédaux induit le développement de malformations rachidiennes, non observées chez les rats quadrupèdes ayant également subit le même type d'intervention chirurgicale. Par analogie à l'Homme, ces résultats soulignent l'importance de la composante biomécanique car le fait de se déplacer debout (bipédalisme) semble être un facteur essentiel pour le développement de la scoliose (*Machida 1999*). Depuis, l'importance de la mélatonine a été renforcée par une autre étude réalisée dans les souris C57BL/6J qui sont naturellement déficientes en mélatonine (mutation naturelle dans le gène NAT (N-Acetyltransferase) impliqué dans la voie de synthèse de la mélatonine et qui cause une réduction substantielle dans sa synthèse). Un grand nombre de souris bipèdes C57BL/6 (50 à 100% selon les études) développent une scoliose suite à l'amputation des membres supérieurs. De plus, le développement de la scoliose semble être évité par la réimplantation intramusculaire de la glande pinéale ou par injection de mélatonine (*Machida, Dubousset et al. 1993; Machida, Dubousset et al. 1995*). Ceci suggère une cause hormonale dans le développement de la scoliose. Par contre, la pertinence biologique de la mélatonine dans la SIA reste cependant controversée puisque la plupart des études cliniques n'ont pu mettre en évidence une différence significative du taux de mélatonine circulante chez les patients SIA par comparaison avec des sujets sains (*Dubousset 1983; Machida, Dubousset et al. 1993; Bagnall, Raso et al. 1999; Bagnall, Beuerlein et al. 2001; Machida, Dubousset et al. 2006*). D'autre part, la scoliose ne peut pas être induite chez les animaux supérieurs comme le singe après ablation de la glande pinéale malgré une carence complète en mélatonine (*Cheung, Wang et al. 2005*). De même, la destruction de la glande pinéale par radiothérapie chez des

enfants atteints d'un cancer rare de la glande pinéale ne conduit pas au développement d'une scoliose (*Lowe, Edgar et al. 2000*). Dans l'ensemble, ces faits suggèrent que la carence en mélatonine pourrait possiblement être un facteur étiologique chez les animaux inférieurs mais pas chez les primates et par conséquent ces résultats ne peuvent pas nécessairement être extrapolés aux êtres humains (*Cheung, Wang et al. 2005*).

### **3.2.2 Le défaut de la signalisation de la mélatonine dans la SIA**

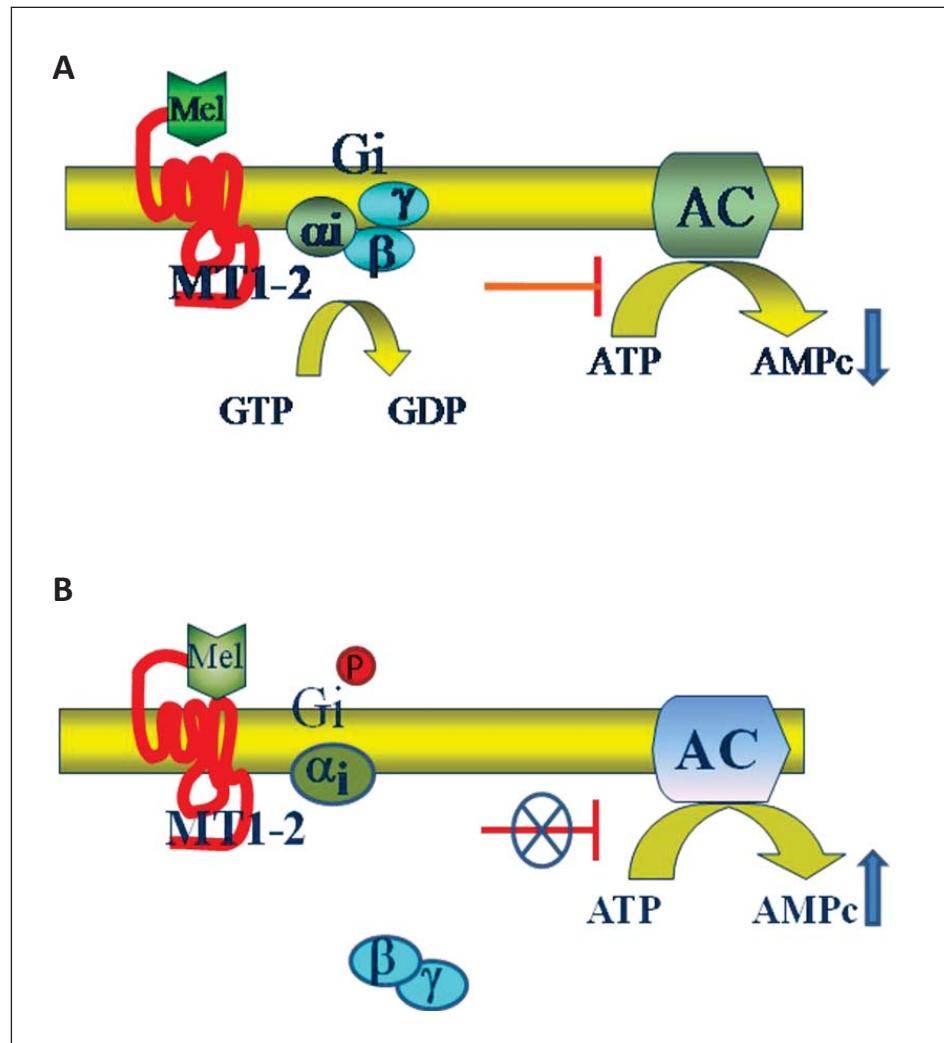
Les travaux du Dr. Moreau ont permis d'apporter un éclairage nouveau dans la compréhension de l'étiopathogénèse de la SIA et de mieux comprendre le rôle de la mélatonine dans cette maladie. En effet, ces travaux ont mis en évidence un défaut dans la transmission du signal de la mélatonine chez des patients gravement atteints de SIA montrant pour la première fois que le problème n'était pas au niveau du signal lui-même (la mélatonine) mais plutôt dans une ou plusieurs composantes de la machinerie protéique impliquée dans sa transmission.

La mélatonine peut se lier avec une haute affinité à ses récepteurs membranaires MT1 et MT2. Ces récepteurs transduisent un signal en réponse au ligand par l'intermédiaire de protéines G inhibitrices ou Gi (guanine nucleotide-binding protein) (*Levoye, Jockers et al. 2006; Azeddine, Letellier et al. 2007; Farce, Chugunov et al. 2008; Jockers, Maurice et al. 2008*). Ces protéines Gi dans les ostéoblastes exhibent différentes affinités et efficacité selon le récepteur sollicité (*MacKenzie, Melan et al. 2002; Witt-Enderby, Bennett et al. 2003; Radio, Doctor et al. 2006; Peng, Bencsik et al. 2008; Millard, Louie et al. 2011*). Les protéines G sont constituées d'une sous-unité  $\alpha$  qui se lie et hydrolyse du GTP ainsi que d'un sous-unité  $\beta$  et un sous-unité  $\gamma$ . Il existe trois sous-unités  $\alpha$  pour les protéines Gi donc Gi1, Gi2 et Gi3. Ces sous-unités  $\alpha$  forment

des trimères avec des sous-unités  $\beta$  et des sous-unités  $\gamma$ . Ces protéines Gi sont activées par des récepteurs membranaires, qui catalysent l'échange du GDP pour GTP dans le site de liaison de guanine nucléotide de la sous-unité  $\alpha$ . Les trimères forment ensemble une unité fonctionnelle capable de réguler de nombreuses protéines effectrices comme les adénylates cyclases (*Federman et al. 1992*), la phospholipase C (*Katz et al. 1992*), les canaux ioniques potassiques ou calciques (*Kim et al. 1989 ; Ikeda et al. 1996 ; Medina et al. 2000*). La sous-unité  $\beta$  interagit avec la sous-unité  $\alpha$  tandis que la sous-unité  $\gamma$  détermine la spécificité de l'effecteur et aurait un rôle dans l'interaction avec le récepteur M2 de l'acide muscarinique (muscarinic receptor) (*Hou et al. 2000; Smrcka et al. 2008*).

Normalement, lorsque la mélatonine est en contact avec ses récepteurs, leur activation conduit à une inhibition de l'activité des adénylates cyclases, enzymes responsables de la synthèse d'AMP cyclique (AMPc) qui est un puissant second messager intracellulaire. En présence de mélatonine, les niveaux en AMPc sont réduits (Figure 8 A).

Chez les patients atteints de SIA, les sous-unités  $G\alpha$  des protéines Gi couplées aux récepteurs membranaires sont hypofonctionnelles suite à une phosphorylation différentielle de certaines séries conduisant ainsi à un dysfonctionnement de la signalisation de la mélatonine. Cette phosphorylation des sous-unités alpha des protéines Gi altère la formation de trimères avec les sous-unités beta et gamma. L'activité des protéines Gi est alors inhibée et l'activité des adénylates cyclases augmente, entraînant ainsi une augmentation du niveau d'AMPc intracellulaire (*Moreau, Wang et al. 2004*). La phosphorylation des protéines Gi a été démontrée chez les patients SIA (*Azeddine, Letellier et al. 2007*) (Figure 8 B).



**Figure 8: La transduction du signal de la mélatonine.**

A. La mélatonine se fixe sur ses récepteurs MT1 et MT2 (GPCR) couplés aux protéines Gi, et inhibe l'accumulation de l'AMPe chez les sujets sains. B. Les récepteurs de la mélatonine sont couplés à des protéines Gi phosphorylées et hypofonctionnelles. Il y a donc une levée d'inhibition de l'activité des adénylates cyclases et une augmentation des niveaux intracellulaires en AMPc.

Le test d'AMPc présente quelques limitations au niveau de la sensibilité et de la reproductibilité, car l'AMPc se dégrade rapidement. Pour cela, notre équipe a développé un test diagnostique fonctionnel basé sur la détection du défaut des protéines Gi dans la voie de signalisation déclenchée par la mélatonine, plutôt que sur le second messager (l'AMPc), en utilisant la technologie de spectroscopie diélectrique cellulaire (cellular dielectric spectroscopy ou CDS).

Le nouveau test fonctionnel utilise la technologie CDS avec un appareil appelé CellKey™ (MDS Scieux, South San Francisco, CA, USA). Cet appareil permet de mesurer l'activation des récepteurs cellulaires de surface endogènes dans des cellules vivantes et produit des données qui permettent l'identification des mécanismes de couplage entre des ligands et des récepteurs inconnus. Cette technologie est potentiellement utile pour l'identification et la découverte de récepteurs endogènes fonctionnels, l'identification de voies de signalisation et enfin pour des analyses pharmacologiques (*Verdonk, Johnson et al. 2006*). Le principe de fonctionnement de la technologie CDS est le suivant : quand les cellules sont exposées à un stimulus, par exemple un ligand, l'événement de transduction du signal active des mécanismes cellulaires tels que la modulation du cytosquelette d'actine, qui provoque des changements dans la forme et le volume de la cellule, au niveau de son adhérence et dans les interactions cellule-cellule. Ces changements, individuellement ou collectivement, affectent le flux de courant extracellulaire ou transcellulaire, ce qui se traduit par des variations de magnitudes et d'impédance. Le principal avantage de cette technologie est qu'elle permet d'étudier la fonction des récepteurs endogènes, ce qui nous donne une information biologiquement plus pertinente et une simplification des procédures. En outre, cette technologie permet d'effectuer un test diagnostique fonctionnel de la SIA, directement à travers l'observation de la réponse des GPCR à la mélatonine sans avoir à mesurer l'AMPc.

Dans le cadre du développement d'un test minimalement invasif, il est malvenu d'utiliser des ostéoblastes car ces cellules sont responsables de la formation des os et sont obtenues seulement par voie chirurgicale ou par biopsie. C'est pourquoi, les lymphocytes ont été adoptés pour la réalisation de ce test diagnostique, car ils sont obtenus par une simple prise de sang et ils possèdent aussi des récepteurs à la mélatonine (*Garcia-Perganeda, Pozo et al. 1997*). Des essais soutenus par la technologie CDS ont démontré que les résultats obtenus avec les lymphocytes sont analogues à ceux obtenus avec les ostéoblastes (*Moreau, Akoume Ndong et al. 2009; Akoume, Azeddine et al. 2010*) (Figure 12).

### **3.3 Autres contributions hormonales à la SIA**

#### **3.3.1 IGF-1**

De nombreuses études essayant d'expliquer les anomalies de croissance chez les patients atteints de SIA ont basé leur hypothèse sur un défaut de la vitesse de croissance chez ces patients. En effet, l'expression irrégulière de l'hormone de croissance et de l'insulin-like growth factor-1 (IGF-1) a été désignée comme pouvant perturber le métabolisme hormonal et causer une asymétrie de croissance. Cette hormone est connue pour jouer un rôle important dans la croissance osseuse et pourrait être impliquée dans la progression des déformations rachidiennes dans la SIA (*Wang, Zhou et al. 1999; Yang, Wu et al. 2009*). Sanders *et al.* rapportent un niveau élevé d'IGF-1 chez les filles atteintes de SIA, et le niveau d'IGF-1 est corrélé avec la vitesse de croissance chez les filles pré-pubères (*Sanders, Browne et al. 2006*). Le profil et la régulation de la sécrétion de l'IGF-1 chez les patientes en particulier, dès la maturation squelettique, devrait

faire l'objet d'études plus poussées (*Nordwall and Willner 1975; Hagglund, Karlberg et al. 1992; Goldberg, Dowling et al. 1993; Burwell and Dangerfield 2002*).

### 3.3.2 Oestrogènes

La SIA se développe et progresse durant la puberté. Cette période critique coïncide avec de nombreux changements biologiques liés aux oestrogènes. Ces hormones sont suspectées d'interférer avec l'activité de la mélatonine. Letellier *et al.* ont montré que l'estriadiol améliore le défaut de signalisation de la mélatonine chez une minorité de patients SIA mais contribue à l'aggraver chez la majorité d'entre eux. C'est le premier modèle qui fait le lien entre l'incidence de la maladie à la puberté, et la prévalence chez un plus grand nombre de filles (*Letellier, Azeddine et al. 2008*).

### 3.3.3 Rôle de la Calmoduline

La calmoduline, une protéine liant les récepteurs du calcium, a également été incriminée dans le développement de la SIA. La calmoduline régule les propriétés contractiles des muscles et est impliquée dans l'activation des plaquettes à travers son interaction avec l'actine et la myosine, et sa régulation des flux de calcium (relargué par le réticulum sarcoplasmique). Des niveaux accrus de calmoduline dans les plaquettes ont été avancés comme étant associés à la progression de la SIA (*Kindsfater, Lowe et al. 1994*). De plus, la mélatonine peut indirectement moduler l'activité de la calmoduline, par régulation des flux de calcium.

### **3.3.4 Ostéopontine**

L'origine biologique de la scoliose, impliquant notamment l'ostéopontine et la mélatonine, suscite un intérêt particulier dans notre laboratoire.

L'ostéopontine (OPN), codée par le gène SPP1, est une glycoprotéine phosphorylée avec un poids moléculaire compris entre 44 et 75 kDa. L'OPN est une cytokine ubiquitaire, présente dans de nombreux tissus et secrétée dans plusieurs liquides biologiques (sang, urine, lait et bile). Cette cytokine joue un rôle de mécano-régulateur au niveau tissulaire, aidant les cellules à intégrer les informations suite à des changements biomécaniques. L'OPN est une cytokine multifonctionnelle régulée par plusieurs facteurs dont la mélatonine. Elle est essentielle pour déclencher l'apparition de la scoliose et stimuler la progression des déformations rachidiennes. Dans notre laboratoire, on a démontré qu'une augmentation anormale de l'expression de l'OPN survenait uniquement chez les poulets pinéalectomisés développant une scoliose et chez les souris bipèdes C57BL/6. Ce dernier modèle a d'ailleurs été utilisé pour démontrer le rôle essentiel de l'OPN dans l'initiation de la scoliose. En effet, l'inactivation génique de l'OPN chez les souris C57BL/6-OPN-KO prévient la formation de scoliose chez les souris bipèdes (*Moreau et al. 2013 manuscrit en préparation*). Au niveau clinique, l'équipe du Dr Moreau a également montré une élévation des niveaux circulants en OPN chez les patients scoliotiques et que cette élévation corrèle avec la sévérité de la maladie (*Moreau et al. 2013 manuscrit en préparation*).

### **3.4 Autres facteurs**

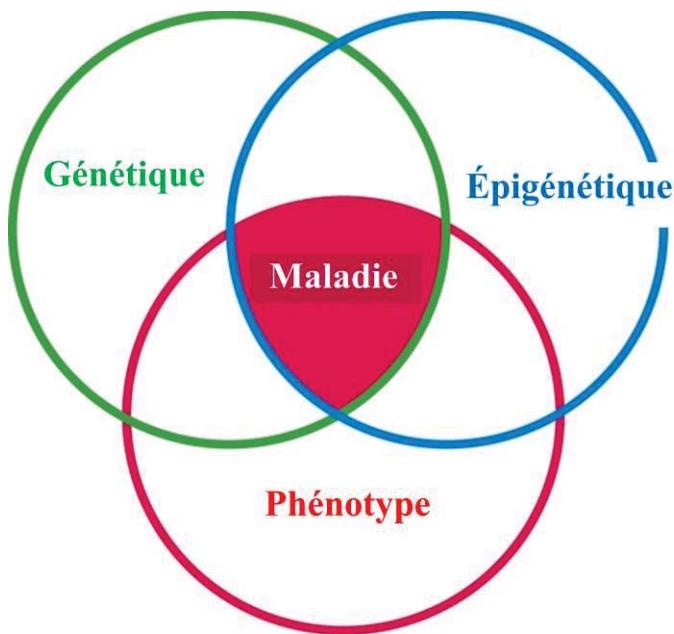
#### **3.4.1 Facteurs environnementaux**

Il s'agit ici d'aborder les liens possibles entre les facteurs environnementaux (tels que le style de vie, l'alimentation, les apports en calcium, en vitamine D, et le niveau d'exercice) et la prévalence de la SIA (*Lee, Leung et al. 1997; Lee, Cheung et al. 2005; Cheung, Lee et al. 2006; Burwell, Dangerfield et al. 2011*). La consommation en calcium est un facteur alimentaire important pour déterminer l'accrétion minérale osseuse durant l'adolescence (*Lee, Leung et al. 1994; Lee, Leung et al. 1995*). Généralement, une faible masse osseuse a été observée chez les patients SIA ainsi qu'une consommation de calcium insuffisante pour la minéralisation osseuse normale (*Cheung, Lee et al. 2006*). McMaster *et al.* ont rapporté que les enfants atteints de SIA font rarement de l'exercice physique (*McMaster M 2006*). Des chercheurs ont donc suggéré que la nutrition étant un facteur environnemental dans l'étiologie de SIA, en associant une mauvaise nutrition avec le développement de la pathologie de la SI (*Worthington and Shambaugh 1993*).

#### **3.4.2 Facteurs épigénétiques**

Les facteurs épigénétiques sont à prendre sérieusement en considération dans l'étude de la SIA. Effectivement, lors de la division cellulaire, des modifications épigénétiques vont se produire que ce soit chez les personnes saines ou malades au niveau de l'ADN et des protéines associées comme les histones qui transportent l'information (*Bjornsson, Fallin et al. 2004; Feinberg 2008; Nelissen, van Montfoort et al. 2011*). Les modifications épigénétiques peuvent alors affecter le système biologique d'une personne avec l'environnement, de façon directe ou

indirecte. Les mécanismes épigénétiques comme la méthylation ou l'acétylation, la modification des histones, les ARN non codant sont des processus dynamiques qui sont influencés par le stade de développement, le type tissulaire, les facteurs environnementaux, comportementaux ou par des événements stochastiques (*Petronis, Gottesman et al. 2003*). Parce que les modifications épigénétiques sont dynamiques et donc partiellement stables, il est possible qu'il puisse y avoir une variation épigénétique potentielle chez les patients SIA. Ces changements épigénétiques pourraient contribuer à l'hétérogénéité de la maladie en modulant l'expression de certains gènes responsables de la progression des déformations rachidiennes (Figure 9).



**Figure 9 : Diagramme de Venn sur la relation entre la variation génétique, épigénétique et la maladie.**

Ce diagramme montrant la relation entre la variation génétique, la variation épigénétique et la maladie. La zone de chevauchement entre la génétique et le phénotype souligne que la plupart des résultats génétiques doivent être considérés dans un contexte épigénétique.

*Figure adaptée de (Feinberg 2010).*

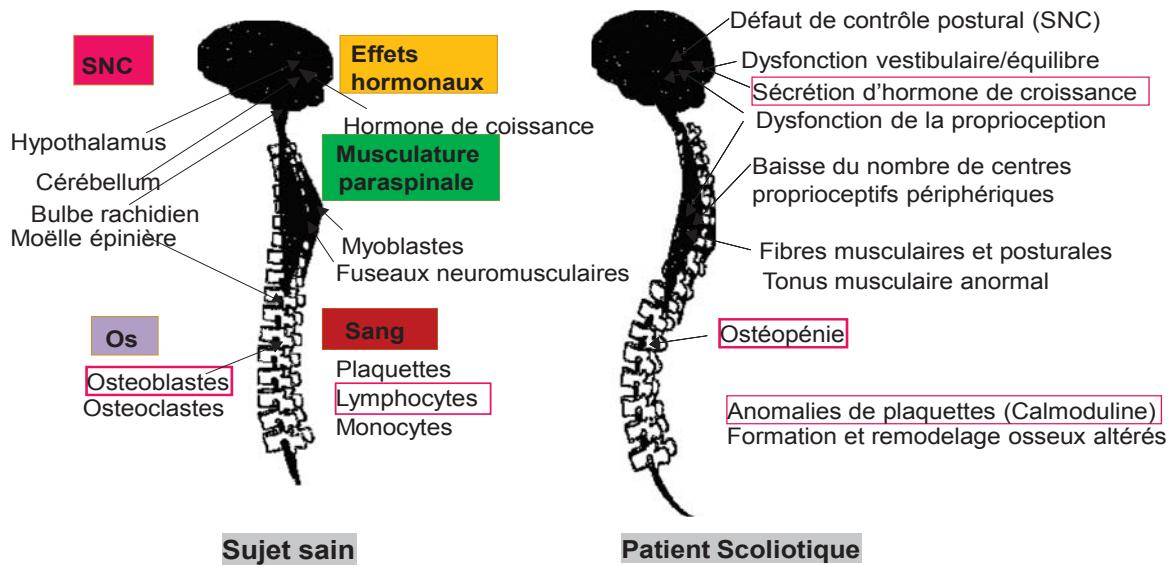
## 4. Co-morbidités

### 4.1 Ostéopénie

Chez des patients atteints de SIA, les travaux de l'équipe du Dr Jack Cheng ont mis en évidence que 30 % d'entre eux développent une ostéopénie, c'est-à-dire une perte de densité osseuse, qui se traduit au début de la vie adulte chez ces patients SIA par le développement d'une ostéoporose accélérée (*Cheng and Guo 1997; Guo, Chau et al. 2005; Lee, Cheung et al. 2005*). Une diminution de la densité osseuse (BMD: bone mineral density) chez les patients de SIA a également été rapportée par d'autres équipes (*Cook, Harding et al. 1987; Thomas, Cook et al. 1992*), suggérant un désordre du métabolisme osseux systématique chez les filles ayant la SIA. L'ostéopénie est également considérée par certains comme un facteur de risque associé à la progression de la courbure chez les patients SIA (*Hung, Qin et al. 2005*). La diminution de la densité osseuse chez les patients atteints de SIA pourrait être la conséquence d'une minéralisation anormale couplée à une accélération de la croissance osseuse pendant la puberté. De récents travaux de l'équipe du Dr Cheng ont mis en évidence une dysfonction de la signalisation intracellulaire de l'IGF-1 (Insulin-like Growth Factor 1) au niveau de cellules osseuses provenant de patients SIA. Ces travaux préliminaires ont également démontré que la mélatonine réprime normalement l'expression du récepteur membranaire de l'IGF-1 (IGF-R1) dans les cellules normales mais pas des cellules SIA, suggérant que cet effet est médié par la signalisation de la mélatonine via l'un de ses deux récepteurs membranaires MT1 et MT2 (*Yeung, Tang et al. 2006*). La perturbation de la signalisation de l'IGF-1 chez les patients SIA pourrait donc avoir un lien avec cette perte de masse osseuse compte tenu de son rôle dans la croissance osseuse.

## 4.2 L'anomalie des plaquettes

Des anomalies plaquettaires sont aussi observées chez les patients atteints de la SIA. Une augmentation du calcium et du phosphore intracellulaire dans les plaquettes de patients SIA, ainsi qu'une augmentation du taux de calmoduline dans les plaquettes de patients SIA seraient associées à la SIA et notamment à la progression de la courbure de la scoliose (*Kindsfater, Lowe et al. 1994*). De plus, les travaux des Docteurs Enouf et Moreau ont mis en évidence un défaut dans la différenciation et dans la maturation des plaquettes sanguines de sujets atteints de SIA (*Bredoux, Corvazier et al. 2006*). Cependant, les défauts observés dans les plaquettes de ces individus pourraient être la conséquence de la scoliose plutôt qu'un évènement primaire à l'origine de la SIA.



**Figure 10 : Système cible dans l'action de la mélatonine et symptômes observés chez les patients scoliotiques.**

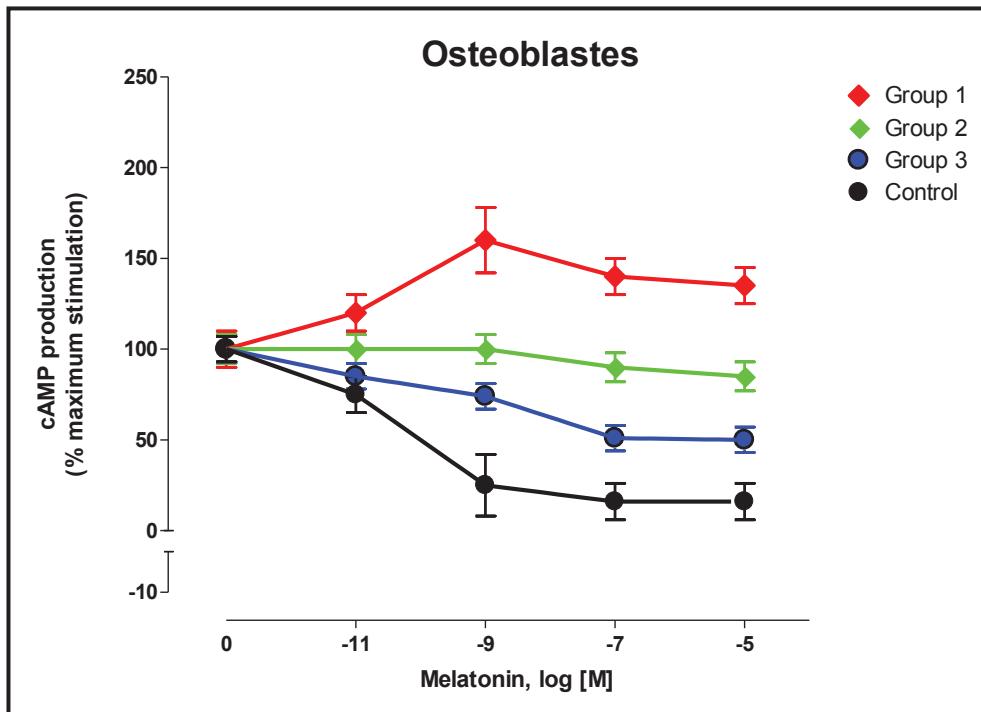
Représentation schématique des effets physiologiques de la mélatonine chez les sujets normaux et leurs liens avec les pathologies observées chez les patients SIA. Tissus ciblés par l'action de la mélatonine. Une défaillance de la signalisation de cette hormone dans ces systèmes pourrait expliquer la majorité des anomalies et des symptômes observés chez les patients SIA.

*Figure adaptée de (Letellier, Azeddine et al. 2007).*

## 5. Les endophénotypes de la SIA

La complexité phénotypique et l'hétérogénéité génétique de la SIA ont compliqué les efforts mis en place pour identifier le pathomécanisme causant la SIA en utilisant des approches de génétique traditionnelle. En effet, les études génétiques de la scoliose idiopathique reposent sur le phénotype principal soit la courbure de la colonne vertébrale. Cependant, ce phénotype est très variable (différents patrons de courbure et de magnitude des déformations rachidiennes) au sein des familles présentant plusieurs individus scoliotiques. Ceci nous a donc poussé à considérer d'autres paramètres cliniques, biochimiques ou métaboliques (endophénotypes) afin de mieux définir l'étiopathogenèse de la SIA. Des évidences indiquent que la génétique est le facteur le plus important dans l'étiologie de la SIA, et le concept actuel considère la SIA comme une condition associée à une variété de gènes dominants en interaction avec des facteurs environnementaux encore méconnus. C'est d'ailleurs cette signalisation croisée avec l'environnement qui pourrait expliquer cette grande variabilité phénotypique. Dans ce contexte, la découverte d'un défaut dans la transmission du signal de la mélatonine chez des patients gravement atteints de SIA a conduit à une classification fonctionnelle des patients en trois différents groupes fonctionnels (Figure 11) (Figure 12) (*Moreau, Wang et al. 2004; Akoume, Azeddine et al. 2010*). Ces groupes fonctionnels représentent des endophénotypes spécifiques, un concept couramment utilisé comme outil en génétique dans d'autres maladies complexes, notamment en psychiatrie, pour simplifier l'étude de ces maladies complexes. Nous avons donc utilisé ces nouveaux endophénotypes afin de mieux identifier les marqueurs prédictifs de la SIA, ce qui représente un objectif majeur dans notre recherche.

L'utilisation d'endophénotypes pour les études de SIA dans de grandes cohortes de patients et de témoins peut objectivement réduire l'hétérogénéité génétique et augmenter les chances de détecter une vraie association.



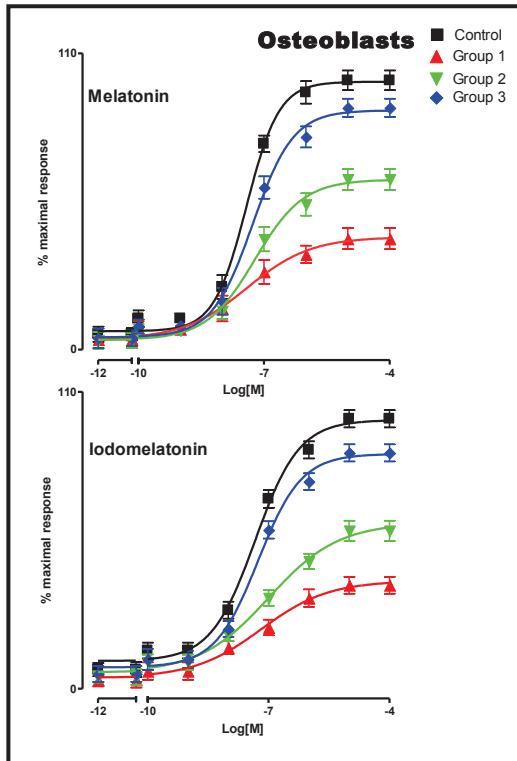
**Figure 11:** Évaluation de l'effet inhibiteur de la mélatonine sur l'accumulation intracellulaire d'AMPc dans les ostéoblastes humains.

*Figure adaptée d' (Azeddine, Letellier et al. 2007)*

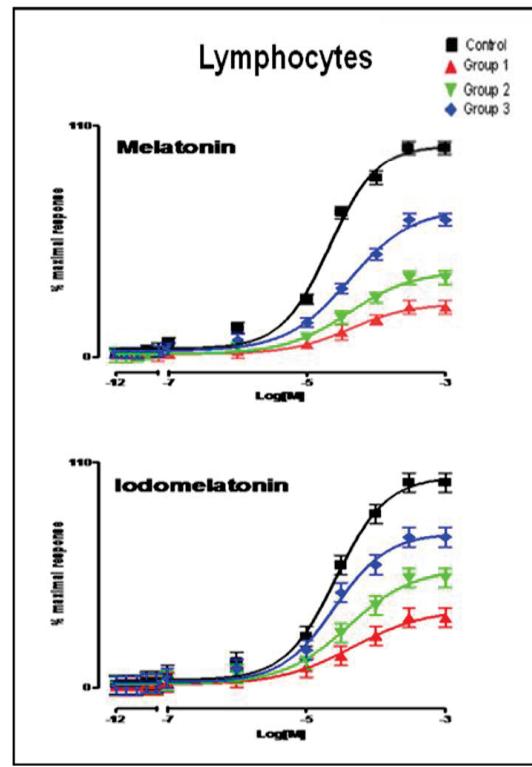
Effet de doses croissantes de mélatonine ( $10^{-11}$  à  $10^{-5}$ M) sur les niveaux d'AMPc intracellulaire induit par la stimulation de l'activité des adénylates cyclases par la forskoline ( $10^{-5}$ M) dans les ostéoblastes dérivés de patients scoliotiques. Les courbes d'inhibition ont permis de classifier les patients SIA en trois groupes fonctionnels.

## Cellular Dielectric Spectroscopy

A



B



**Figure 12: Dysfonctionnement de la signalisation démontré par la CDS.**

A. Les courbes dose-réponse de la mélatonine et de l'iodomélatonine obtenues par la CDS avec les ostéoblastes de patients SIA et de sujets témoins.

B. Les courbes dose-réponse de la mélatonine et de l'iodomélatonine obtenues par la CDS avec les lymphocytes de patients SIA et de sujets témoins.

## **6. Hypothèses et Objectifs**

### **Hypothèses :**

**H1.** Les trois groupes fonctionnels (endophénotypes) observés chez les patients SIA sont la résultante de changements génétiques distincts qui devraient se traduire par des profiles d'expression particuliers nous permettant d'identifier de gènes potentiellement impliqués dans chaque groupe fonctionnel.

**H2.** Outre ces gènes activateurs ou principaux de la SIA, la comparaison entre deux jumelles monozygotique, une diagnostiquée avec la scoliose et l'autre non. Patients SIA partageant un même patron de courbure devrait aussi permettre d'identifier certains ensembles de gènes responsables de la formation d'un patron de courbure particulier (gènes dits secondaires). Ainsi des gènes dont le changement d'expression génique pourrait être influencé par des modifications d'épigénétiques liées à l'environnement.

**H3.** Nous proposons que l'ostéopénie et certains troubles de l'humeur (anxiété et dépression) puissent être associés plus spécifiquement à un groupe fonctionnel donné et résultent de changements moléculaires au niveau de gènes spécifiques.

### **Objectifs :**

**Objectif principal:** Examiner les profiles d'expression moléculaires des patients SIA après les avoir stratifiés selon la classification fonctionnelle développée par Moreau et al. afin d'identifier les marqueurs moléculaires responsables du défaut de signalisation observé dans chaque endophénotype (groupe fonctionnel).

### **Objectifs spécifiques:**

**S1:** Identifier les gènes qui sont plutôt responsable de l'initiation de la SIA en comparant les profils moléculaires des patients SIA d'un même groupe fonctionnel et du défaut différentiel observé dans la signalisation des récepteurs couplés aux protéines Gi.

**S2:** Vérifier les fonctionnalités et les mécanismes des gènes spécifiques à certains groupes, pouvant contribuer au développement de certaines co-morbidités et/ou au risque d'aggravation des déformations rachidiennes, dans l'affirmative d'identifier les gènes jouant un rôle dans le métabolisme osseux (croissance et minéralisation) ou dans la signalisation de l'ostéopontine.

## **CHAPITRE 2 : ARTICLE I**

## **Introduction sur le premier article : Genome-wide expression profiling of idiopathic scoliosis endophenotypes**

Nous avons déjà démontré un dysfonctionnement de la signalisation différentielle agissant par les récepteurs couplés aux protéines G inhibitrices chez les patients SIA, conduisant à leur stratification en trois groupes fonctionnels distincts (endophénotypes). Les endophénotypes servent à simplifier la complexité phénotypique de cette maladie. Ces défauts sont la résultante de changements génétiques distincts qui devraient se traduire par des profils d'expression particuliers, nous permettant ainsi d'identifier des gènes que nous considérons comme les activateurs principaux de la SIA.

L'objectif principal est d'examiner les profils moléculaires de patients SIA après les avoir stratifiés selon la classification fonctionnelle développée par Moreau *et al.* afin d'identifier les marqueurs moléculaires responsables du défaut de signalisation observé dans chaque groupe fonctionnel et aussi responsables des différents profils d'expression génétique des trois endophénotypes. Nous avons effectué cette étude en utilisant la technologie de la société Affymetrix sur des biopuce d'ADN afin d'examiner l'expression génique du génome entier et d'identifier les gènes associés avec chaque groupe fonctionnel. Par la suite, nous avons étudié les fonctions de ces gènes d'intérêt et valider leur contribution dans des fonctions biologiques importantes.

### **La contribution des auteurs :**

Qing Yuan a réalisé l'étude expérimentale et a contribué à la préparation du manuscrit et des analyses fonctionnelles des gènes. Kristen Fay Gorman a réalisé des analyses quantitatives par Anova, évalué et interprété les données et a aidé à la préparation du manuscrit. Cédric Julien a effectué une partie des analyses par qPCR et a participé à la validation des gènes candidats.

Dashen Wang a aidé dans les expériences et a isolé les ostéoblastes des patients scoliotiques et des sujets sains.

Alain Moreau a conçu l'étude, a supervisé le projet dans son ensemble, a évalué et interprété les données et a aidé lors de la préparation du manuscrit.

Tous les auteurs ont lu et approuvé le manuscrit final.

# Genome-wide expression profiling of idiopathic scoliosis endophenotypes

Qing Yuan<sup>1</sup> and Kristen Fay Gorman<sup>1</sup>, Cédric Julien<sup>1</sup>, Saadallah Bouhnik<sup>1</sup>, Anita Franco<sup>1</sup>, Karen Choquet<sup>1</sup>, Da Shen Wang<sup>1</sup>, Marie-Yvonne Akoume<sup>1</sup>, Alain Moreau<sup>1,2</sup> 

<sup>1</sup>*Viscogliosi Laboratory in Molecular Genetics of Musculoskeletal Diseases, Sainte-Justine University Hospital Research Center; Department of Biochemistry, Faculty of Medicine, Université de Montréal, Québec, Canada.*

<sup>2</sup>*Department of Stomatology, Faculty of Dentistry, Université de Montréal, Québec, Canada.*



Phone: 514-345-4931 ext: 3476

Fax: 514-345-4801

## **Abstract**

Idiopathic scoliosis (IS) is the most common deformity arising during childhood, having phenotypic complexity and an unknown aetiology. We have previously demonstrated a differential signalling dysfunction acting through G inhibitor protein-coupled receptors in IS patients, leading to their stratification into three distinct groups (endophenotypes). Furthermore, we have demonstrated a pivotal role for osteopontin in scoliosis onset and progression, and that the biochemical endophenotypes have distinct osteopontin (OPN) activity profiles. However, the molecular pathways of osteopontin and its molecular partners related IS are unknown. Herein we examine the gene expression profiles of the three IS endophenotypes. We performed this study using genome-wide expression microarray technology in order to assess gene expression patterns that are associated with the different groups (endophenotypes). We found that the gene expression profile for group 1 patients is distinct from that of group 2 and group 3, and that the gene expression from group 2 and group 3 are similar. Importantly, this difference in gene expression profiles reflects established biochemical and signaling activities such that group 1 is resistant to osteopontin and group 2 and group 3 are sensitive to osteopontin. We further analyzed the function of expression profiles for each group. We hypothesize that functional or genetic differences between groups relate to important biological roles for the regulation of osteopontin, and Gi protein signaling pathways related to IS. We further validate candidate genes for single-gene analysis via quantitative real time PCR. Candidates were chosen partly based on the ANOVA and partly from hypotheses regarding genes detected in our exploratory survey. Our data further support our method of stratification of IS patients using endophenotypes and identifies candidate genes associated with IS to a better understanding of the etiology of this complex disease.

## **Introduction**

Idiopathic scoliosis (IS) is characterized as an abnormal three-dimensional curvature of the spine and is the most prevalent type of spinal deformity, affecting 3-4% of the global pediatric population, on average. The onset of IS can occur at any age between birth and sexual maturity and so the syndrome has been classified as infantile, juvenile, or adolescent IS, based on the age when a curve first presents itself [1, 2]. Adolescent Idiopathic Scoliosis (AIS), the most common form of IS, occurs between the ages of 10 and 18, with girls being affected more severely than boys [3].

The etiology of IS remains unclear, however, it appears to be multifactorial [4]. That the primary etiology and the pathogenesis of curvature are influenced by genetic factors is widely accepted [5-21]. Support for a genetic basis comes from a number of family surveys that showed a higher incidence of scoliosis between family members than the remainder population [5, 7, 8, 22, 23], and from a number of twin studies [24-30]. The IS phenotype is highly variable and so it has been suggested that multiple genes influence the deformity [4]. Linkage analysis and genome-wide association studies suggest that IS is linked/associated with chromosomes 6p, 10q, 17p, Xq, 9q, 18q, 3q, 5q, 19p, 17q, 12p, 8q [21, 31-40]. However, these regions are often large, and so can encompass hundreds or even thousands of genes across many megabases of DNA [41], making it difficult to identify candidate genes using a non-hypothesis based approach. Moreover, candidate gene studies conducted thus far have failed to confirm hypotheses [42]. This is not surprising considering that IS is a complex deformity that develops during growth and so many biological systems are impacted. Examples of biological systems hypothesized to be involved in the etiopathogenesis of IS are: the nervous and vestibular systems [43-46] (e.g.

balance and gait studies) [47-51]; bone and connective tissues [52-57]; muscle [58-63]; platelets [64, 65]; the endocrine system (*i.e.* melatonin, estrogen, testosterone, growth hormone, insulin growth factor) [66-70]; and molecular signaling [71]. As a consequence of this, many different tissues have been considered as affected. In addition, IS occurs during a dynamic period with changing hormone levels and gene expression levels. Studies among monozygotic twins demonstrates that environmental and/or epigenetic factors are also important considerations for IS study.

Phenotypic complexity and possible genetic heterogeneity has complicated efforts to identify the biology of IS using traditional genetic approaches. With complex diseases such as IS, the use of endophenotypes is highly advantageous because genetic variation can be partitioned, thus giving a study greater power to detect a genetic effect [72-74]. We previously demonstrated three distinct biochemical endophenotypes for IS based on a differential signaling dysfunction of Gi-coupled receptors in osteoblasts and other cell types isolated from AIS patients [71, 75]. The goal of this study was to investigate the molecular differences that distinguish the three endophenotypes. To accomplish this we performed a genome-wide expression study using microarray technology. We demonstrate that the osteoblasts of AIS patients show distinct expression profiles associated with resistance or sensitivity to osteopontin, thus closely mirroring the biochemical activity of the endophenotypes previously defined [75]. Genes that are strongly differentially expressed between subgroups defined by endophenotypes are novel candidates and will help to elucidate important molecular pathways involved in the regulation of osteopontin, as it relates to IS.

## **Materials and methods**

### **Study populations**

The Institutional Review Board of CHU Sainte-Justine approved this study. Parents or legal guardians of all participants gave their informed written consent, and minors their assent. The microarray analysis was performed on osteoblasts obtained from intraoperative biopsies of 15 female patients with severe AIS who were between the ages of 12.7 and 18.5 years, residents of Quebec, and of European descent. All were diagnosed with AIS by one of the orthopedic surgeons at St. Justine University Hospital, Montreal, Quebec, Canada. The magnitude of curvature ranged from 34 to 78 degrees (Cobb angle). We restricted our microarrays to patients with either right thoracic or right/left double curve types (Table 1). For controls, we performed microarray analysis on osteoblasts obtained from traumatic cases, three females non-IS patients who were between the ages of 11.6 and 15.5 years (Table 2).

Real time quantitative PCR (RT-QPCR) analysis was used to validate candidate genes from the microarray analysis. RT-QPCR was performed on osteoblasts from 24 non-IS controls and 35 randomly selected AIS patients with severe curvature: 14 from group 1, 11 from group 2, and 10 from group 3 (Tables 3 and 4). All clinical variables indicated were obtained at the moment of the surgery.

### **Isolation and characterization of Osteoblasts**

For all AIS cases, osteoblasts were derived from bone specimens surgically removed from affected vertebrae (the vertebral number sampled varied from T3 to L4) as a part of correctional

surgery for severe scoliosis. For the non-IS trauma cases that were used as controls, bone specimens were obtained from other anatomic sites (tibia or femur).

Bony fragments were manually reduced into smaller pieces with a bone cutter under sterile conditions and then incubated in 1x Dulbecco's Modification Eagle's Medium ( $\alpha$ DMEM) (Wisent Inc, Montreal QC, Canada) supplemented with 10% HyClone Fetal Bovine Serum (FBS) (Thermo Fischer Scientific, Logan UT, USA) and 1% antibiotic (Invitrogen) at 37°C with 5% CO<sub>2</sub> until confluent (10-14 days). At confluence, osteoblasts emerging from the bone pieces were isolated by trypsinization. Functional classification of AIS derived osteoblast cultures was conducted as described in Akoume, *et al.*, 2010.

### **Extraction of RNA**

Total RNA from osteoblasts of patients and control subjects was extracted using Trizol (Invitrogen, CA, USA) according to manufacturer's instructions. Briefly, cells were lysed in 1ml of Trizol and homogenized with pipette and transferred to clean RNA-free tube. Samples were layered with 100  $\mu$ l of BCP (1-bromo-3-chloropropane) (MRC, OH, USA), inverted 15 times, and incubated 3 min at room temperature. Lysates were centrifuged at 4°C for 15 min at 12,000 g and the aqueous phase transferred to a clean tube. RNA was precipitated with equal volumes of isopropanol for 10 min at room temperature. Precipitates were spun for 10 min at 4°C at 12,000 g. Pellets were washed with 70% Ethanol, and dried at room temperature. RNA was resuspended in 50 $\mu$ l RNase free water (Qiagen, Canada). An aliquot of total RNA diluted in RNase free water was set aside to estimate RNA quantity and integrity, and the remaining sample was stored at -80°C until gene expression analysis. Total RNA quantity and integrity was assessed using the spectrophotometry and all samples maintained a 28S/18S rRNA ratio of 1.5 or greater.

For RNA samples sent to process the MicroArray analysis at Centre d’Innovation de Génome Québec à Montréal, RNA integrity was confirmed by RNA Quality Control provided by Genome Québec, according to instructions.

### **Microarray and Analysis of Data**

Experiments were performed in triplicate using the Affymetrix, GeneChip® Human Exon 1.0 ST array, at the Centre d’Innovation de Génome Québec, Montréal. All protocols were conducted as described in the Affymetrix GeneChip Expression Analysis technical manuals. We used Wright and Simons (EB) method [76] in FlexArray program [77] to identify genes whose expression differed between groups. Raw data were normalized by the RMA method (Robust Multi-array Analysis). To explore genes differentially expressed between groups, probe sets were filtered by expression change being at least  $\pm 3$  fold with group 1 and group 2 versus group 3 comparisons were employed by two class unpaired analysis. As a quantitative analysis, the log<sub>2</sub> ratio for each probe set was calculated and *p* values from one-way ANOVA followed by false recovery rate (FDR)-correction were calculated according to Benjamini and Hochberg [78]. Probes were considered significant if their FDR-corrected *p* value was  $\leq 0.05$ . We categorized the functionalities of the identified genes using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) software [79]. Candidate genes selected for single-gene analysis via quantitative real time PCR were partly based on the ANOVA and partly from hypotheses regarding genes detected in our exploratory survey.

## **Quantitative Real-Time PCR**

Candidate genes having differential expression patterns among the patient groups were validated by quantitative real time PCR (RT- QPCR) using PerfeCTa SYBR Green SuperMix (Quanta Biosciences) on a 7900HT Real-Time PCR system (Applied Biosystems) according to the manufacturer's protocol. For each subject, 1 $\mu$ g of mRNA was reverse transcribed into cDNA using the Tetro cDNA Synthesis kit (Bioline, MA, USA) according to the manufacturer's protocol. The fold-change in gene expression was determined by the program RQ manager [80] ([www.appliedbiosystems.com](http://www.appliedbiosystems.com), *Applied Biosystems 7900HT Fast Real-Time PCR System Relative Quantitation Using Comparative Ct Getting Started Guide*). To determine the correlation between the microarray and RT-QPCR results, we calculated the median fold change in expression (for a given gene) for AIS vs. non-IS controls, and compared it to the fold change seen by microarray to determine whether the gene was still differentially expressed in the same direction. We used an ANOVA or Mann-Whitney tests to examine whether expression levels were different among groups of AIS patients (GraphPad Prism v4).

## **Results**

The Affymetrix, GeneChip® Human Exon 1.0 ST array has 54,000 probe sets that analyzed the expression of over 47,000 transcripts and variants from over 30,000 well-characterized human genes. Results of the genome-wide expression analysis of AIS patients are broken into 3 sections: Data exploration, Quantitative analysis, and Candidate gene testing.

### *Data exploration*

To assess gene expression patterns that are associated with the different functional groups (endophenotypes) we used FlexArray to compare the expression profile of group 1 (G1) to the expression profile of group 2 (G2), and then to the expression profile of group 3 (G3). Similarly we compared the expression profiles of group 2 to that for group 3. We used a 3-fold expression change relative to each group as a cut-off and identified a total of 347 genes. Considering that a 3-fold expression change may be too stringent and thus mask genes that are important for generating hypotheses for differences between groups, we also surveyed the data using a 2-fold threshold and the same comparisons. The 2-fold analysis revealed a total of 800 additional genes. How genes are distributed for each of the between-group comparisons, under each threshold model, is summarized in Table 5 and Figure 1. We discovered that in general, the gene expression profile for functional G1 is dramatically different from that of G2 and G3, and that there is not much of a difference in gene expression between G2 and G3.

We used the **D**atabase for **A**nnotation, **V**isualization and **I**ntegrated **D**iscovery (**DAVID**) v6.7 to classify the differentially expressed genes. To keep our functional analysis manageable we explored genes with a 3-fold expression change between groups. However, because there are only four genes differentially expressed between groups 2 and 3 under such a threshold, we used a 2-fold cut-off for our exploration of differentially expressed genes in this category, thus increasing the number of genes to 59. There are 116 genes shared in the differentially expressed gene profiles of G1 vs. G2 and G1 vs. G3, suggesting that these genes are uniquely over-expressed and under-expressed in G2 and G3 compared to G1. Because the number of genes is too long to list, we classify them according to functional clusters in Table 6, and then we

further explored the biological activity of these functional clusters in Table 7. Likewise, there are 21 genes shared in the gene expression profiles of G1 vs. G3 and G2 vs. G3, suggesting that these genes are uniquely expressed in G3. These genes are listed in Table 8 and their functional qualities are explored in Table 9. There are no genes that appear to be shared in the G1 vs. G2 and G2 vs. G3 profiles. Among the comparisons, there are 12 genes exclusive to the G1 vs. G2 profile, these genes are listed in Table 10. There are 79 genes exclusive to the G1 vs. G3 profile, and because the number of genes is too long to list, we classify them into functional clusters in Table 11, and then the biological activity of these functional clusters in Table 12. Lastly, there are 28 genes exclusive to the G2 vs. G3 profile, and these genes are listed in Table 13 and their functional terms are listed in Table 14. Consideration of the known functional roles or the hypothesized functional roles (based on our own unpublished studies) for genes, and their uncorrected p-values, in light of biological systems impacted by IS, guided our selection of candidate genes that were tested via RT-QPCR.

### *Quantitative analysis*

To survey differentially expressed genes using a non-hypothesis based approach, we conducted an analysis of variance (ANOVA) on RMA (robust means analysis) normalized Log 2 transcribed values for all 54,000 probe sets, using Flexarray. We employed the Benjamin and Hohberg algorithm for a False Discovery Rate correction for multiple testing and considered genes with a corrected p-value  $\leq 0.05$  as significantly differentially expressed. We conducted the analysis with and without control subjects.

We identified 25 genes and one micro RNA with a corrected  $p \leq 0.05$  that are differentially expressed among the biochemical endophenotypes (groups) (Table 15). From our analysis of differentially expressed genes among the groups that included controls from trauma cases, we identified 6 genes with a corrected  $p \leq 0.05$ . All the genes identified in the analysis with controls are among the genes identified in the analysis without controls (Table 15 and Figure 2).

### *Candidate genes*

We used RT-QPCR on an expanded cohort to validate 22 interesting genes from our data exploration, or quantitative analysis (Table 16). We tested candidate genes on 14 additional G1 patients, 11 additional G2, and 10 additional G3 patients and compared the relative expression levels to either controls (24 additional controls) or among the groups. Interesting candidate genes are those that relate to systems hypothesized to be involved in the IS etiopathogenesis, either from publications or based on unpublished research in our lab. Five candidates are associated with the extracellular matrix (*MMP1*, *CHI3L1*, *SPPI*, *EFEMP1*, *COMP*), one of which (*COMP*) is located in a putative locus for IS [38]. *PDE1A* is of interest because it is functionally related to calmodulin. And integrin beta 3 (*ITGB3*) is known to enhance IGF-I-stimulated signaling, a known growth factor [81]. Also IGF1 is reported to have a strong relationship with estrogen at different ages in women in a breast cancer study [82]. Furthermore, *ITGB3* is a receptor for matrix metalloproteinase-2, osteopontin [83]. *ACTG2* and *ITGA8* are associated with function in cytoskeleton, *LRRC17* is functional related to osteogenesis [84], the other genes are the most over-expressed or under-expressed genes in the group comparisons.

Of the 22 Genes tested by RT-QPCR, only *SPP1* (the osteopontin gene) was significantly differently expressed among the three groups (Figure 3a). Three genes that were significantly differentially expressed in our quantitative analysis (*MKX*, *SLPI*, *CLGN*) did not show significant differences among groups in the expanded cohort. However, we noticed that among the genes tested there was an emerging pattern of expression where G1 was different from G2 and G3. As discussed below, this expression pattern seems to follow a functional attribute that separates G1 from G2 and G3. Therefore, we used a t-test to examine whether G1 had a significantly different expression pattern from G2 &G3 combined. From this test *MKX*, *SLPI*, and *CLGN* showed significantly different expression patterns (Fig. 3 b, c, d).

## Discussion

The goal of this study was to explore gene expression differences that are associated with biochemically defined idiopathic scoliosis (IS) endophenotypes [75, 85, 86]. IS is a complex deformity that has been defined by clinical features, although a genetic basis is recognized. The biology for the syndrome has remained enigmatic since the time of Hippocrates and there has been little progress with regard to treatment or screening. We recently demonstrated that IS is associated with a defect in *Gi* protein signaling and that this defect can be stratified into three distinct groups [75]. These groups define the endophenotypes for IS, which is an important start for understanding the immense clinical variability on a biological level. Further characterization of these endophenotypes has elucidated a pivotal role for osteopontin in scoliosis onset and progression (unpublished data). Considering the ubiquity of osteopontin, its precise relationship with IS must be articulated by examining its partner molecules, thus defining critical molecular pathways for the IS etiopathogenesis. By using microarray technology we were able to screen IS

patients for genome-wide that we previously defined. We found that the gene expression profile for group 1 patients is distinct from that of group 2 and group 3. Importantly, this difference in gene expression patterns reflects observed biochemical and signaling activities such that group 1 is resistant to osteopontin and group 2 and group 3 are sensitive to osteopontin. We expect that the 25 genes and 1 miRNA significantly differentially expressed between osteopontin resistant (G1) and sensitive (G2 and G3) groups relate to important biological roles for the regulation of osteopontin pathways related to the primary etiology in IS.

Phenotypic variability has complicated genetic studies for IS [42]. The magnitude of curvature (Cobb angle), the morphology of curvature, the age of curve onset, and the propensity for progression are all highly variable aspects of IS. Complex inheritance and the large number of loci identified as linked or associated with IS have encouraged the idea that in addition to being polymorphic, it is a genetically heterogeneous syndrome. To control for possible heterogeneity in gene expression and thus increase the power to detect important gene expression changes that are related to the endophenotypes, we only selected patients with adolescent IS (AIS), with severe curves, and two types of curve morphology. Of 15 patients, 6 have a double major curve pattern, and 9 have a right thoracic curve (we will consider gene expression as it relates to curve pattern in future analysis). Because we controlled for the curve type and the age and the severity of curvature, we expect that the gene expression differences observed here relate primarily to the biochemically unique endophenotypes.

As a first step in the analysis of microarray data, we explored similarities and differences between gene expression profiles of patients according to their endophenotypes (groups). We

chose not to include controls in these comparisons because considering that these tissues were derived from anatomical locations other than the spine, we could not rule out the possibility that differential gene expression in the control profile was linked to location rather than the IS relevant signaling dysfunction. Considering that the goal of this study is to distinguish gene expression patterns specific to each group, we feel that exclusion of controls in this step of analysis is justified. Among the 402 genes analyzed (Table 5), there are 116 genes shared in G2 vs. G1 and G3 vs. G1 comparisons and thus co-expressed (over-expressed or under-expressed) in G2 and G3 compared to G1 (Tables 6 & 7). Similarly, there are 21 genes co-expressed in G1 and G2 compared to G3 (Tables 8 & 9). There are no shared genes in the G1 and G3 expression profiles. This is interesting because biochemically G1 and G3 have the most different signaling profiles [75]. Furthermore, there are 79 genes exclusive to the G3 vs. G1 profile (Tables 11 &12), 12 genes exclusive to the G2 vs. G1 profile (Table 10), and 28 genes exclusive to the G2 vs. G3 profile (Tables 13 & 14). The genes exclusive to each profile are candidates for molecular differences between these groups. To explore functional similarities and differences between endophenotype (group) expression profiles, we used DAVID to identify functional clusters.

Functions shared in G1 and G2 profiles relate to shared aspects of these groups' biochemistry, and likewise for G2 and G3 profiles. Within the shared molecular profile for G2 and G3 we identified 4 significant functional clusters (Table 6), which were further examined by analyzing key biological modules shared among the molecules (Table 7). Importantly, one gene cluster is related to G-protein coupled receptor signaling, with 6 genes in this cluster: *BDKRB1* (bradykinin receptor B1), *BDKRB2* (bradykinin receptor B2), *CCRL1* (chemokine (C-C motif) receptor-like 1), *CCKAR* (cholecystokinin A receptor), *PTGFR* (prostaglandin F receptor),

*F2RL1* (coagulation factor II (thrombin) receptor-like 1). Considering that we have previously demonstrated a signaling dysfunction specific to Gi-coupled receptors in the IS patients [75, 87], these genes may play a role in this Gi-coupled receptor signaling dysfunction. However, Bradykinin receptors (BDKRB1 and BDKRB2) are normally coupled to Gq-coupled receptors [88-90], and Gq-coupled receptors are not affected in AIS patients (unpublished data), so we can exclude these two. Nine genes belong to the functional cluster that relates to signal transduction, containing six genes from the previous cluster, and *FMOD* (fibromodulin), *SFRP4* (Secreted frizzled-related protein 4). G protein-coupled receptor (GPCRs) signaling is described as a series of molecular signals that proceeds with an activated receptor promoting the exchange of GDP for GTP on the alpha-subunit of an associated heterotrimeric G-protein complex. The GTP-bound activated alpha-G-protein then dissociates from the beta- and gamma-subunits to further transmit the signal within the cell. The pathway begins with receptor-ligand interaction and Gi-mediated signal transduction, which might explain why these 2 functional clusters share genes [91-94]. GPCR signaling is regulated not only by small molecule binding to allosteric or orthosteric sites, but also by protein/protein interactions, most with intracellular proteins such as the canonical G proteins, GPCR kinases and arrestins [95]. G protein-coupled receptors (GPCRs) can engage multiple pathways to activate ERK1/2 via both G proteins and/or β-arrestin. *CCRL1* belongs to the G-protein coupled receptor 1 family, chemokines mediate their diverse activities through G protein-coupled receptors [96], researches have shown that C-C chemokine receptor subfamily exist in teleost fish [97], as Gorman *et al.* demonstrated spinal curvature also exist in teleosts [98], we suspect there may be un chemokine receptor dysfunction exist in teleost fish as well, suggest a possible role of C-C chemokine receptor in scoliosis development. Abera *et al.* demonstrated that *PTGER2* (prostaglandin E receptor 2) and *PTGFR* (prostaglandin F receptor) are often co-expressed in the same cells and use different intracellular signaling pathways. EP2 receptors

couple to  $G\alpha_s$ , resulting in increased formation of cAMP, while FP receptors couple to  $G\alpha_q$  which in turn results in release of inositol-1,4,5-triphosphate (IP) and dialcylglycerol (DAG). Moreover, they used siRNA molecules targeted against the adenylyl cyclase 3 (AC3) isoform, showing that AC3 is responsible for the cross-talk between the FP and EP2 receptors [99]. We need to further investigate their functions to scoliosis. *FMOD* (fibromodulin) is extracellular matrix protein may participate in the assembly of the extracellular matrix as it interacts with type I and type II collagen fibrils, it also participates in tooth formation, the absence of Fmod impaired dentin mineralization, increased the diameter of the collagen fibrils [100]. Embree *et al.* have showed that Fmod are critical for TMJ OA (Temporomandibular joint osteoarthritis) subchondral bone integrity, Fmod knock out mice have a type I collagen deficient in subchondral bone [101]. *SFRP4* (Secreted frizzled-related protein 4) has a very closed relationship with osteoblasts, Nakanishi *et al.* have reported a single nucleotide polymorphisms in the *Sfrp4* gene are responsible for low peak BMD in senescence-accelerated mouse (SAM) P6 [102], they showed an overexpression of *Sfrp4* in osteoblasts suppressed osteoblast proliferation, resulting in a decrease in bone formation in vivo, such result suggest the role of *SFRP4* in osteoblast [103]. The same result have demonstrated by Cho HY *et al.* *Sfrp4* negatively regulates bone formation [104]. However, it shows an underexpression of *SFRP4* in G2 and G3 of IS patients which may increase bone formation; it needs to be further investigated.

Other genes that are not showing in the functional clusters: *COMP* (Cartilage oligomeric matrix protein) is identified as a potential marker for AIS [105], it is essential for the normal development of cartilage and for its conversion to bone during growth. We showed a low expression of *COMP* in G2 and G3, the abnormal expression may prevent normal bone to grow, and other research showed that Comp promotes attachment via two mechanisms involving CD47 and alphaVbeta 3 integrin [106]. Also, both COMP and OPN are considered as of ECM

components [107]. *TGFBR3* (transforming growth factor, beta receptor III) is bone mass candidate genes, genes that are associated with bone mineral density, which is a prominent osteoporosis risk factor [108]. Abnormal expression of *TGFBR3* in G2 and G3 indicates a high risk of osteoporosis in these IS patients.

*SEPP1* (selenoprotein P) underexpressed in G2 and G3 patients, while a selenium deficiency have showed in both hairs and serum of patients with IS [109]. *CLU* (Clusterin) also has a low expression in G2 and G3 patients compare to patients in G1, Zhu Y *et al.*, *Clu* was down-regulated in the serum of degenerative scoliosis patients [110]. *CHI3L1* (Cartilage glycoprotein-39) also known as *YKL-40* is considered a new biomarker in the spine disease by Tsuji *et al.*, they showed it is expressed in cerebrospinal fluid as a result of damage or stress to the neural structure [111]. Moreover, *CHI3L1* present in human lumbar disc demonstrated a potential role in the pathophysiology of discal degeneration and inflammation [112]. In our expression level comparison, it is the most underexpressed gene in G2 and G3 patients compare to G1, we should better clarify the role of this protein in idiopathic scoliosis or other related pathology.

Finally, it is worth noting that of these 116 genes, the expression level difference between G3 vs. G1 is generally greater than between G2 vs. G1, suggesting that there may be a bigger difference between G3 and G1 than G2 and G1.

There are only 21 genes shared in the molecular profile for G1 and G2 (Table 8), classified these according to their individual functions (Table 9). Three extracellular matrix genes are under-expressed in group 3 and this is interesting because the extracellular matrix is hypothesized to play a role in the IS pathology. The MMPs (Matrix Metalloproteases), *MMP3*

and *MMP1*, Matrix metalloproteinases regulate degradation and remodeling of the extracellular matrix (ECM), and are important candidate structure proteins genes in IS [113-116]. Interestingly, Grivas *et al.* reported an overexpression of *MMP1* in experimentally induced scoliotic deformity on a rat tail intervertebral disc resulting in an asymmetrical distribution of *MMP1* between the convex and concave side, which was dependent on the degree of the deformity [117]. Another group, identified in the human *MMP3* promoter a gene polymorphism (5A/6A or rs3025058) associated with IS. The frequency of the 5A/5A genotype of *MMP3* gene polymorphism in patients with scoliosis was almost 3 times higher than in controls (30.2% vs. 11.2%, p 0.001) in the Italian population [115]. However, Liu *et al.* did not find any significant association of this SNP indicating that the *MMP3* promoter polymorphism is not associated with IS in the Chinese population [114]. It cannot be ruled-out that *MMP3* may modulate curve susceptibility when interacting with *BMP4* [113]. Also, *ECM2* (Extracellular matrix protein 2) plays a role in extracellular matrix organization, and an extracellular matrix structure mal-function may occur in scoliosis [118]. Considering that many of the common genes shared in G1 and G2 and likewise in G2 and G3 are involved in the G-protein coupled receptor protein signaling pathway, and that there is a differential dysfunction of Gi protein coupled receptor signaling pathway in all three groups of AIS patients, those functional profiles further support the existence of the Gi-protein signaling pathway dysfunction inside our AIS patients.

We identified genes that are exclusive to each comparison; we classified 12 genes differentially expressed between G1 and G2 (Table 10), 79 genes between G1 and G3 (Table 11 and 12), and 28 between G2 and G3 (Table 13 and 14). The number of exclusive genes in group comparison G1 and G3 is much more than in G1 and G2 comparison, further suggesting that there are more

molecular differences between G1 and G3 than between G1 and G2. Our analysis of the functional classifications for genes associated each of the above comparisons, may be important for the biochemical differences between them. There are 11 functional terms associated in the G1 and G3 comparison, suggesting a large biochemical difference between G1 and G3 (Table 12). The functional differences are evident in extracellular matrix binding, ECM-receptor interactions, regulation of actin-cytoskeleton, integrin-mediated pathways. Of the genes identified in our array, integrin beta 3 (*ITGB3*) is a receptor for matrix metalloproteinase-2, osteopontin, osteomodulin, and thrombospondin [83]. Matrix metalloproteinases regulate degradation and remodeling of the extracellular matrix (ECM), and are important candidate structure proteins genes in IS. Thus, it is possible that *ITGB3* is a relevant growth gene related to IS progression. With 3.1 fold over-expression of *ITGB3* in G3 compare to G1, this may contribute to the resistance vs. sensitive of OPN to G1 to G3. Furthermore, OPN has been shown to regulate the actin cytoskeleton reorganization and bone resorption [119-121].

There are molecular similarities between G2 and G3 in that they both show a sensitivity of OPN, but their gene expression profiles have 28 genes differentially expressed. Previous work in our lab analysis with IS patients classified in the surgical cohort (late stage disease cohort) showed a more prevalent distribution of G2 (46%) compared to individuals classified either in G1 (20%) or G3 (34%). Logistic regression analysis allowed us to determine an odds ratio of 5-fold representing the risk of developing a severe scoliosis for IS patients classified in G2, while patients classified in the G1 showed a reduced risk of developing a severe scoliosis when compared to the other subgroups. An important functional difference between G2 and G3 is the classification term “regulation of growth”. This may correlate with the severity of the disease,

because the genes that are responsible for the regulation of growth are expressed differently between G2 and G3, thus explaining why patients in G2 represent a higher risk of developing a severe scoliosis.

Although we screened hundreds of genes functions simultaneously, we noticed that many genes in our comparisons did not get assigned to any of the functional gene clusters. The possible reasons for why a gene is excluded from our functional clustering are: it does not relate with other genes which have a similar threshold; it is related with only a few other genes, but together they do not have enough members to form a gene cluster based on minimum requirements defined in DAVID; or the gene was omitted in error, as the current algorithm in DAVID could have up to a 2 % false negative rate.

To analyze the microarray data as a whole we used an ANOVA [122]. Each assayed gene is tested independently statistically for a difference in expression between groups. The output of the analysis is a probability (*p* value) that a difference in expression could have been observed by chance. As our previous approach replied on non-statistically motivated criteria such as "3 fold change" to select genes, it was limited in its sensitivity and reliability, except for the largest and most robust changes in expression [123]. The ANOVA test can reveal genes that show small but highly significant changes in expression [124]. With both methods, we are ensured to reveal important candidate genes for IS. The method of ANOVA allows one to examine whether the variability due to a particular experimental factor, or combination of factors, is statistically significant compared to the measured variability due to random sources. Thus, we used ANOVA to examine differences between groups. In addition to family-wise error rate (*p* value) at 0.05, we

used false discovery rate (FDR) corrected *p* value to increase the stringency of the test [125, 126]. We analyzed the data with and without controls. Genes revealed in the analysis with controls are considered to be highly significantly associated with either the spine or IS, however, as previously discussed, the controls are imperfect and so we considered the results without controls to be informative for the purpose of this study.

Analysis of the data without controls gave us 25 significantly differentially expressed genes among the endophenotypes, after correction for multiple testing (Table 15). These genes are important candidates for understanding differences between the groups that must be confirmed in an expanded cohort via RT-QPCR. Fifteen of these genes are included in the DAVID analysis. Furthermore, *CLGN*, *MKX*, *SLPI*, *LRRK2*, *GALNTL2*, *SLC2A12*, *KCNJ2*, *CHRD1*, *ANGPTL5* belong to the genes shared in G2 and G3, but none of them falls into the biological modules in the gene cluster. Therefore, we searched main function of each of these genes using DAVID, in an attempt to relate functionality to the G-protein signaling or OPN sensitivity, or other hypotheses for IS. We found that chordin-like 1 (*CHRD1*) plays a role in the skeletal system development, bone development, ossification, and BMP signaling pathway [127, 128]. We have demonstrated before that MMP3 may modulate curve susceptibility when interacting with BMP4 [113]. We hypothesized that chordin-like 1 may interact in the BMP signaling pathway and play a role in the curve development in IS. *ISMI*, *TBC1D8*, *HTR1F* belong to the genes exclusive to G3 vs. G1, 5-hydroxytryptamine (serotonin) receptor 1F (*HTR1F*) involves in G-protein coupled receptor protein signaling pathway, also, a number of molecular genetic studies have investigated if serotonin (5-HT) receptor subtypes are involved in the pathogenesis of depression [129-131]. Depression is considered as one of the co-morbidities for IS patients (unpublished data), and

serotonin is known as the major hormone regulating mood, its receptors thus have a strong link to this co-morbidity [132]. We are currently validating this theory in our lab. The dysfunction of Gi-protein coupled receptor signaling pathway is systemic, this might be the reason for this co-morbidity in IS patients. Finally, *PAX9*, *MOXD1* and *SERPINII* belong to the exclusive genes to G2 vs. G1. Downregulation of *PAX9* in G2 is particularly of interest because this transcription factor regulates vertebral column during development [133].

Five genes showed significance at FDR corrected  $p \leq 0.05$  inclusive the controls, from the heat map in Figure 2, *SLC39A8* showed the most overexpression compared to control, *SLC39A8* encodes a key Zinc transporter, while zinc deficiency has been shown in AIS patients. Author showed decreased zinc content in back muscle of patients with IS to be a secondary disturbance associated with the deformation of the spine [109, 134].

We started confirmation of the significantly differentially expressed genes by testing 3 (*MKX*, *SLPI*, *CLGN*) by RT-QPCR. Confirmation of the other 22 genes is ongoing in the lab and beyond the scope of this paper. These confirmed genes show that there is a distinct expression difference between the OPN resistant versus sensitive groups, and the expression patterns are similar to the biochemical functional profiles in IS patients, G1 separated from G2 and G3 (Fig 3 b,c,d).

Among the other genes confirmed using RT-QPCR, the gene that encodes the osteopontin (OPN) protein (*SPP1*), was significantly differently expressed among the three groups. *SPP1* has very low expression in G1, biochemically characterized as OPN resistant, and it has higher expression level in G2 and G3, the 2 groups are characterized as OPN sensitive. From the expression level

differences between Group2 and Group3, we can tell the G3 has more OPN sensitivity than G2, as is also evident from the fact that there are much more gene expression level differences among the 116 genes in G3 than in G2 when compare to G1. Furthermore, we demonstrated that the groups defined by the biochemical endophenotypes have distinct osteopontin (OPN) activity profiles, demonstrating that this is a key molecule in the initiation of IS. We hypothesized that osteopontin is essential to trigger scoliosis onset and stimulates deformity progression in AIS patients. Articulation of the biochemistry for each of these groups (unpublished data) has facilitated our choice for candidate genes. The expression profile of OPN in the three groups reflect our IS endophenotypes.

## Conclusions

Phenotypic variability is the major difficulty in decoding the real etiopathogenesis of IS. Although large a number of loci have been identified as linked or associated with IS, and many genes have been examined as candidates for IS, none of them has been confirmed. In order to reveal novel candidates for this genetically heterogeneous syndrome, we examined gene expression for the three unique IS associated biochemical endophenotypes that are defined by a dysfunction in Gi protein signaling. Microarray gene expression technology provides us with information on the activities of all genes simultaneously therefore allowing us to apply the field of functional genomics to each endophenotype. Such an approach to a complex phenotype such as IS has allowed for the discovery of novel molecular pathways not revealed by sequence analysis alone. We have identified several genes involved in Gi protein signaling, that are associated with three groups, and OPN related genes associated with G1 alone and G2, G3 respectively, suggesting that they play a role in the distinct osteopontin (OPN) activity profiles.

The genes indicated in this study are a first step in understanding the biological pathways that are associated with the pathogenesis of IS. It is anticipated that identification of specific molecular markers to each group will be used to identify patients' prognosis, help the clinician to better predict their clinical outcomes and response to personal therapies in the future.

## Tables and Figures

**Table 1. Clinical Data of AIS Patients used for Microarray (Osteoblasts)**

| Case No. | Gender | Age at Surgery<br>(yrs) | Cobb Angle<br>(degrees) | Curve          | Functional |
|----------|--------|-------------------------|-------------------------|----------------|------------|
|          |        |                         |                         | pattern        | Group      |
| 1075     | Female | 13.13                   | 60-77                   | R/L double     | 1          |
| 1060     | Female | 14.41                   | 44-38                   | R/L double     | 1          |
| 1059     | Female | 14                      | 61-54                   | R/L double     | 2          |
| 1012     | Female | 12                      | 52-40                   | R/L double     | 2          |
| 1020     | Female | 14                      | 59-55                   | R/L double     | 3          |
| 1064     | Female | 15                      | 70-70                   | R/L double     | 3          |
| 1322     | Female | 13.11                   | 51                      | Right thoracic | 1          |
| 1311     | Female | 14.46                   | 78                      | Right thoracic | 1          |
| 1377     | Female | 16.18                   | 58                      | Right thoracic | 2          |
| 1069     | Female | 12.72                   | 63                      | Right thoracic | 2          |
| 1014     | Female | 14.97                   | 30                      | Right thoracic | 2          |
| 1144     | Female | 5.42                    | 53                      | Right thoracic | 2          |
| 1056     | Female | 18.49                   | 34                      | Right thoracic | 3          |
| 1319     | Female | 13.73                   | 54                      | Right thoracic | 3          |

|      |        |       |    |                |   |
|------|--------|-------|----|----------------|---|
| 1040 | Female | 16.29 | 60 | Right thoracic | 3 |
|------|--------|-------|----|----------------|---|

Functional groups are based on endophenotypes previously published in Akoume *et al.*, 2010. R/L double curve patterns are manifest as right/left thoracic and lumbar curves. The Cobb angle is the standard measure of curve magnitude for IS, where a higher angle denotes a more severe curve. For double curve patterns the first number represents the thoracic curve and the second number represents the lumbar curve.

**Table 2. Clinical Data of Control Subjects Studied for Microarray (Osteoblast)**

| Case No. | Gender | Age  |
|----------|--------|------|
| T25      | Female | 13.8 |
| T14      | Female | 11.6 |
| T19      | Female | 15.5 |

Three female non-IS subjects for Microarray studies, age of 11.6 to 15.5

**Table 3. Clinical Data of AIS Patients Studied For RT-QPCR (Osteoblasts)**

| Case No. | Gender | Age (yrs) | Cobb Angle | Curve pattern       | Functional Group |
|----------|--------|-----------|------------|---------------------|------------------|
| 1090     | Female | 17.97     | 45-57      | R/L double          | 1                |
| 1423     | Female | 14.57     | 73-45      | R/L double          | 1                |
| 1012     | Female | 11.81     | 74-56      | R/L double          | 1                |
| 1032     | Female | 11.44     | 62-47      | R/L double          | 1                |
| 1335     | Female | 17.60     | 47-50      | R/L double          | 1                |
| 1038     | Female | 17.16     | 47         | Left thoracolumbar  | 1                |
| 1081     | Female | 13.64     | 60         | Left thoracolumbar  | 1                |
| 1174     | Female | 18        | 37         | Left thoracolumbar  | 1                |
| 1228     | Female | 13.18     | 57         | Right thoracic      | 1                |
| 1311     | Female | 14.46     | 78         | Right thoracic      | 1                |
| 1322     | Female | 13.11     | 51         | Right thoracic      | 1                |
| 1264     | Female | 15.15     | 40         | Right thoracolumbar | 1                |
| 1017     | male   | 16.41     | 45         | Left lumbar         | 1                |
| 1083     | Female | 14.14     | 76         | Left lumbar         | 1                |
| 1042     | Female | 14.24     | 70-48      | R/L double          | 2                |

---

|      |        |       |       |                     |   |
|------|--------|-------|-------|---------------------|---|
| 1060 | Female | 12.76 | 53-55 | R/L double          | 2 |
| 1029 | Female | 14.40 | 56-53 | R/L double          | 2 |
| 1387 | Female | 15.01 | 55-46 | R/L double          | 2 |
| 1025 | Female | 15.32 | 53-42 | R/L double          | 2 |
| 1033 | Female | 16.19 | 35    | Left thoracolumbar  | 2 |
| 1066 | Female | 17.33 | 53    | Right thoracic      | 2 |
| 1063 | Female | 14.81 | 67    | Right thoracic      | 2 |
| 1377 | Female | 16.18 | 58    | Right thoracic      | 2 |
| 1013 | Female | 14.96 | 54    | Right thoracic      | 2 |
| 1674 | Female | 16.49 | 54    | Right thoracolumbar | 2 |
| 1058 | Female | 14.40 | 90-66 | R/L double          | 3 |
| 1373 | Female | 14.61 | 41-48 | R/L double          | 3 |
| 1361 | Female | 16.60 | 65-95 | R/L double          | 3 |
| 1665 | Female | 12.81 | 45-30 | R/L double          | 3 |
| 1064 | Female | 15.16 | 90-90 | R/L double          | 3 |
| 1036 | Female | 13.02 | 48    | Right thoracic      | 3 |
| 1439 | Female | 17.72 | 69    | Right thoracic      | 3 |

---

---

|      |        |       |    |                     |   |
|------|--------|-------|----|---------------------|---|
| 1653 | Female | 11.20 | 68 | Right thoracic      | 3 |
| 1658 | Female | 19.66 | 46 | Right thoracolumbar | 3 |
| 1043 | Female | 14.12 | 50 | Right thoracolumbar | 3 |

---

Curve patterns are defined as follows: R/L double curve patterns are manifest as right/left thoracic and lumbar curves; Right thoracic curve patterns are concentrated in the thoracic (mid-back) region and curves to the right; Right thoracolumbar curve patterns indicate the major curve is to the right in the thoracic and lumbar region; Left thoracolumbar curve patterns indicate the major curve is to the left in the thoracic and lumbar region; Left lumbar curve pattern indicate the major curve is to the left and is concentrated in the lumbar region.

**Table 4. Clinical Data of Control Subjects Studied For RT-QPCR (Osteoblast)**

| Case No. | Gender | Age   |
|----------|--------|-------|
| T32      | Female | 8.8   |
| T14      | Female | 11.6  |
| T20      | Female | 13    |
| T22      | Female | 14    |
| T2       | Female | 14.1  |
| T12      | Female | 14.2  |
| T51      | Female | 14.98 |
| T8       | Female | 15.1  |
| T45      | Female | 15.2  |
| T19      | Female | 15.5  |
| T11      | Female | 15.7  |
| T1       | Female | 15.9  |
| T13      | Female | 18.7  |
| T18      | Female | 19    |
| T39      | Female | 20.4  |

---

|     |      |      |
|-----|------|------|
| T10 | Male | 12   |
| T42 | Male | 12.1 |
| T15 | Male | 13   |
| T25 | Male | 13.8 |
| T30 | Male | 14.2 |
| T5  | Male | 15.4 |
| T43 | Male | 16.2 |
| T7  | Male | 16.6 |
| T41 | Male | 17.2 |

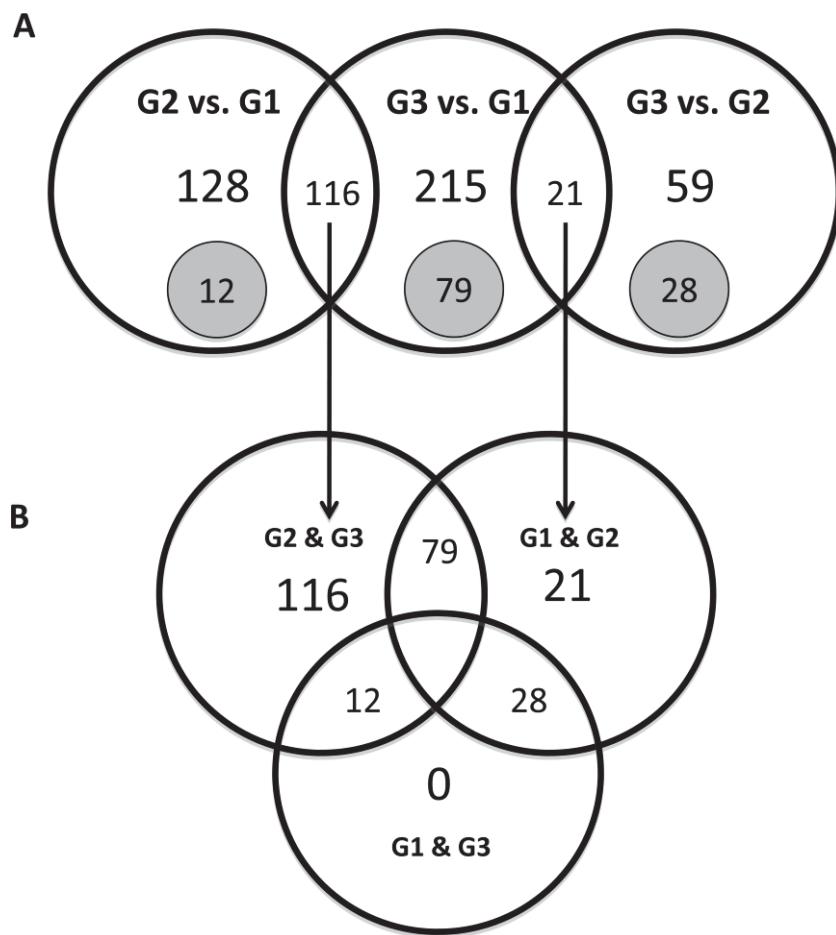
---

**Table 5. The number of genes differentially expressed between groups with fold-change**

| Fold change | G1 vs. G2 | G1 vs. G3 | G2 vs. G3 | Total |
|-------------|-----------|-----------|-----------|-------|
| 2           | 409       | 679       | 59        | 1147  |
| 3           | 128       | 215       | 4         | 347   |

Using a 2-fold change in levels of expression as our threshold for our between group comparisons identified 800 more genes than a 3-fold threshold. In order investigate a manageable number of genes we considered those with 3-fold expression differences for group 1 (G1) vs. group 2 (G2) and G1 vs. group 3 (G3). However, considering that so few genes are differentially expressed between G2 and G3 with a 3-fold threshold, we considered the 59 genes revealed using a 2-fold threshold. Therefore, in sum we considered 402 genes in our functional analysis.

**Figure 1. The distribution of genes differentially expressed in between-group comparisons**



- A.** Using a 3-fold threshold, there are 128 genes differentially expressed between groups 1 (G1) and 2 (G2), and 215 differentially expressed between groups 1 (G1) and 3 (G3). Using a 2-fold threshold, there are 59 genes differentially expressed between G2 and G3. Among these comparisons, there are 79 genes exclusive to the G1 vs. G3 profile (Tables 12 & 13), 12 genes exclusive to the G1 vs. G2 profile (Table 10), and 28 genes exclusive to the G2 vs. G3 (Tables 14 & 15). **B.** There are 116 genes that are shared in the differentially expressed gene profiles of G1 vs. G2 and G1 vs. G3 (Table 6 & 7). There are 21 genes shared in the differentially expressed gene profiles of G1 vs. G3 and G2 vs. G3 (Tables 8 & 9). There are no genes that appear to be shared in the G1 vs. G2 and G2 vs. G3 profiles.

**Table 6. Functional classification of genes shared in groups 2 and 3 expression profiles**

| Gene                  | Location | Gene           | Title   | Fold Change |           | Enrichment Score |
|-----------------------|----------|----------------|---|-------------|-----------|------------------|
|                       |          |                |   | G3 vs.G1    | G2 vs. G1 |                  |
|                       |          |                |   |             |           |                  |
| <b>Gene cluster 1</b> |          |                |   |             |           | <b>15.54</b>     |
| 1q32.1                |          | <i>CHI3L1</i>  | chitinase 3-like 1 (cartilage glycoprotein-39)  | -38.5       | -18.6     |                  |
| 7q11.23               |          | <i>FGL2</i>    | fibrinogen-like 2   | -19.2       | -13.8     |                  |
| 9q22                  |          | <i>OGN</i>     | osteoglycin   | -12.9       | -10       |                  |
| 7p14.1                |          | <i>SFRP4</i>   | secreted frizzled-related protein 4   | -12.1       | -4.2      |                  |
| 11p15.4               |          | <i>OLFML1</i>  | olfactomedin-like 1   | -8.8        | -10.7     |                  |
| 8p22-p21.3            |          | <i>PDGFRL</i>  | platelet-derived growth factor receptor-like  | -8.1        | -5.3      |                  |
| 11q22.1               |          | <i>ANGPTL5</i> | angiopoietin-like 5   | -7.5        | -6.3      |                  |
| 4q21                  |          | <i>IGJ</i>     | immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides | -4.5        | -3.9      |                  |
| 9q34                  |          | <i>ANGPTL2</i> | angiopoietin-like 2   | -4.4        | -3.1      |                  |
| 1p13.3                |          | <i>CHI3L2</i>  | chitinase 3 like 2  | -3.7        | -3.7      |                  |
| <b>Gene cluster 2</b> |          |                |   |             |           | <b>7.41</b>      |
| 9q22                  |          | <i>OGN</i>     | osteoglycin   | -12.9       | -10       |                  |
| 1q32                  |          | <i>FMOD</i>    | fibromodulin  | -7.4        | -4.8      |                  |
| 1q32                  |          | <i>PRELPL</i>  | proline/arginine-rich end leucine-rich repeat protein                                     | -6.3        | -4.5      |                  |
| 9q22                  |          | <i>ASPN</i>    | asporin   | -6.0        | -5.7      |                  |

|                       |               |  |       |             |
|-----------------------|---------------|--|-------|-------------|
| Xp22.33               | <i>MXRA5</i>  | matrix-remodelling associated 5                  | -5.4  | -3.4        |
| 9q22.31               | <i>OMD</i>    | osteomodulin                                     | -5.3  | -4.9        |
| 7q22.1                | <i>LRRC17</i> | leucine rich repeat containing 17                | 6.4   | 4.0         |
| 12q13.11              | <i>AMIGO2</i> | adhesion molecule with Ig-like domain 2          | 11.3  | 7.8         |
| <b>Gene cluster 3</b> |               |  |       | <b>5.78</b> |
| 8p22-p21.3            | <i>PDGFRL</i> | platelet-derived growth factor receptor-like     | -8.1  | -5.3        |
| Xp22.33               | <i>MXRA5</i>  | matrix-remodelling associated 5                  | -5.4  | -3.4        |
| 11q23-q24             | <i>CDON</i>   | Cdon homolog (mouse)                             | -3.5  | -3.1        |
| 3q12-q13              | <i>CD200</i>  | CD200 molecule                                   | 3.5   | 3.1         |
| 12q13.11              | <i>AMIGO2</i> | adhesion molecule with Ig-like domain 2          | 11.3  | 7.8         |
| <b>Gene cluster 4</b> |               |  |       | <b>3.17</b> |
| 3q22                  | <i>CCRL1</i>  | chemokine (C-C motif) receptor-like 1            | -22.1 | -14.6       |
| 14q32.1-q32.2         | <i>BDKRB1</i> | bradykinin receptor B1                           | -7.0  | -3.6        |
| 14q32.1-q32.2         | <i>BDKRB2</i> | bradykinin receptor B2                           | -5.3  | -3.6        |
| 1p31.1                | <i>PTGFR</i>  | prostaglandin F receptor (FP)                    | -4.6  | -4.2        |
| 14q22                 | <i>PTGER2</i> | prostaglandin E receptor 2 (subtype EP2)         | -3.6  | -4.9        |
| 4p15.1-p15.2          | <i>CCKAR</i>  | cholecystokinin A receptor                       | -3.2  | -4.0        |
| 3q12-q13              | <i>CD200</i>  | CD200 molecule                                   | 3.5   | 3.1         |
| 5q13                  | <i>F2RL1</i>  | coagulation factor II (thrombin) receptor-like 1 | 5.1   | 3.2         |

Using the 3-fold threshold to compare genes differentially expressed between G1 and G2 and then between G1 and G3, we identified 116 genes. These genes were then analyzed using The

Gene Functional Classification Tool with default settings in DAVID. There are 4 main functional clusters with significant enrichment for annotation terms. The group enrichment score ranks the biological significance of gene groups based on overall EASE scores (modified Fischer Exact p-values) of all enriched annotation terms. Higher enrichment scores denote the strongest associations.

**Table 7. Key biological modules for each gene cluster**

| Functional Clustering                           | Gene          | Title  | P value |
|---|---------------|--|---------|
| 1. Peptide receptor activity, G-protein coupled |               |  | 1.8E-6  |
|   | <i>CCRL1</i>  | chemokine (C-C motif) receptor-like 1            |         |
|   | <i>BDKRB1</i> | bradykinin receptor B1                           |         |
|   | <i>BDKRB2</i> | bradykinin receptor B2                           |         |
|   | <i>CCKAR</i>  | cholecystokinin A receptor                       |         |
|   | <i>F2RL1</i>  | coagulation factor II (thrombin) receptor-like 1 |         |
| 2. Signal transduction                          |               |  | 1.7E-4  |
|   | <i>CCRL1</i>  | chemokine (C-C motif) receptor-like 1            |         |
|   | <i>BDKRB1</i> | bradykinin receptor B1                           |         |
|   | <i>BDKRB2</i> | bradykinin receptor B2                           |         |
|   | <i>CCKAR</i>  | cholecystokinin A receptor                       |         |
|   | <i>CDON</i>   | Cdon homolg (mouse)                              |         |
|   | <i>F2RL1</i>  | coagulation factor II (thrombin) receptor-like 1 |         |
|   | <i>FMOD</i>   | fibromodulin                                     |         |
|   | <i>PTGFR</i>  | prostaglandin F receptor (FP)                    |         |
|   | <i>SFRP4</i>  | secreted frizzled-related protein 4              |         |
| 3. Calcium signaling pathway                    |               |  | 3.9E-4  |
|   | <i>BDKRB1</i> | bradykinin receptor B1                           |         |
|   | <i>BDKRB2</i> | bradykinin receptor B2                           |         |

---

|   |  |        |
|---|--|--------|
| <i>CCKAR</i>  | cholecystokinin A receptor                       |        |
| <i>PTGFR</i>  | prostaglandin F receptor (FP)                    |        |
| 4. chitinase activity                                   |  | 2.5E-3 |
| <i>CHI3L1</i>   | chitinase 3-like 1 (cartilage glycoprotein-39)   |        |
| <i>CHI3L2</i>   | chitinase 3 like 2                               |        |
| 5. G-protein coupled receptor protein signaling pathway |  | 4.1E-3 |
| <i>CCRL1</i>  | chemokine (C-C motif) receptor-like 1            |        |
| <i>BDKRB1</i>   | bradykinin receptor B1                           |        |
| <i>BDKRB2</i>   | bradykinin receptor B2                           |        |
| <i>CCKAR</i>  | cholecystokinin A receptor                       |        |
| <i>F2RL1</i>  | coagulation factor II (thrombin) receptor-like 1 |        |
| <i>PTGFR</i>  | prostaglandin F receptor (FP)                    |        |
| 6. Cellular ion homeostasis                             |  | 6.1E-3 |
| <i>BDKRB1</i>   | bradykinin receptor B1                           |        |
| <i>BDKRB2</i>   | bradykinin receptor B2                           |        |
| <i>CCKAR</i>  | cholecystokinin A receptor                       |        |
| <i>F2RL1</i>  | coagulation factor II (thrombin) receptor-like 1 |        |
| 7. Regulation of leukocyte migration                    |  | 8.8E-3 |
| <i>BDKRB1</i>   | bradykinin receptor B1                           |        |
| <i>F2RL1</i>  | coagulation factor II (thrombin) receptor-like 1 |        |

---

---

|  |               |  |
|--|---------------|--|
| 8. Response to external stimulus                                   |               | 2.1E-2   |
|  | <i>CCRL1</i>  | chemokine (C-C motif) receptor-like 1            |
|  | <i>BDKRB1</i> | bradykinin receptor B1                           |
|  | <i>BDKRB2</i> | bradykinin receptor B2                           |
|  | <i>CCKAR</i>  | cholecystokinin A receptor                       |
|  | <i>F2RL1</i>  | coagulation factor II (thrombin) receptor-like 1 |
| 9. Aminoglycan catabolic process, Polysaccharide catabolic process |               | 2.5E-2   |
|  | <i>CHI3L1</i> | chitinase 3-like 1 (cartilage glycoprotein-39)   |
|  | <i>CHI3L2</i> | chitinase 3 like 2                               |
| 10. Negative regulation of protein amino acid phosphorylation      |               | 3.8E-2   |
|  | <i>BDKRB1</i> | bradykinin receptor B1                           |
|  | <i>BDKRB2</i> | bradykinin receptor B2                           |

---

To consider the biological roles for those genes with co-expression (over-expression and under-expression) in G2 and G3, we used DAVID to classify the genes of each cluster identified in Table 6 according to shared biological modules or functional terms.

\* The count signifies number of genes for the corresponding term. Significance was defined as p-value < 0.05 and was determined using Fisher Exact test

**Table 8. Comparision of genes deregulated in IS functional group 3**

| Location                     | Gene<br>Symbol  | Title   | Fold change |             |  |
|------------------------------|-----------------|---|-------------|-------------|--|
|                              |                 |   | G3 vs<br>G1 | G3 vs<br>G2 |  |
| <b>OVER-EXPRESSED GENES</b>  |                 |   |             |             |  |
| 8p12                         | <i>PLAT</i>     | Plasminogen activator, tissue                             | 7.2         | 3.2         |  |
| 12q21.2                      | <i>E2F7</i>     | E2F transcription factor 7                                | 4.1         | 2.1         |  |
| 5p15.1-p14                   | <i>CDH6</i>     | Cadherin 6, type 2, K-cadherin (fetal kidney)             | 4.1         | 2.9         |  |
| 18q21.3                      | <i>SERPINB2</i> | Serpin peptidase inhibitor, clade B (ovalbumin), member 2 | 4           | 2.1         |  |
| 4q28.3                       | <i>PCDH10</i>   | Protocadherin 10  | 3.9         | 2.8         |  |
| 7p21.1                       | <i>TSPAN13</i>  | Tetraspanin 13  | 3.9         | 2.8         |  |
| 1q23.1                       | <i>RGS5</i>     | Regulator of G-protein signaling 5                        | 3.7         | 3.0         |  |
| 5q23-q31                     | <i>ITGA2</i>    | Integrin, alpha 2   | 3.5         | 3.0         |  |
| 8q21.13                      | <i>FABP5</i>    | Fatty acid binding protein 5 (psoriasis-associated)       | 3.1         | 2.1         |  |
| <b>UNDER-EXPRESSED GENES</b> |                 |   |             |             |  |
| 12q24.1                      | <i>OASI</i>     | 2',5'-oligoadenylate synthetase 1, 40                     | -3.0        | -2.1        |  |
| 3q25                         | <i>PTX3</i>     | Pentraxin-related gene, rapidly induced by IL-1 beta      | -3.3        | -2.1        |  |
| 6q25                         | <i>FNDC1</i>    | Fibronectin type III domain containing 1                  | -3.3        | -2.2        |  |
| 4q13-q21                     | <i>IL8</i>      | Interleukin 8   | -4.1        | -2.1        |  |
| 4q31-q32                     | <i>SCRG1</i>    | Stimulator of chondrogenesis 1                            | -4.1        | -2.3        |  |
| 9q22.3                       | <i>ECM2</i>     | Extracellular matrix protein 2, female organ and          | -4.1        | -2.1        |  |

---

|         |               |  |      |      |  |
|---------|---------------|--|------|------|--|
|         |               | adipocyte specific   |      |      |  |
| 4q21    | <i>CXCL1</i>  | Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha) | -4.2 | -2.3 |  |
| 8q24.1  | <i>ENPP2</i>  | Ectonucleotide pyrophosphatase/phosphodiesterase 2                             | -5.4 | -2.6 |  |
| 2q12    | <i>IL1R1</i>  | Interleukin 1 receptor, type I   | -5.4 | -2.2 |  |
| 16p12.2 | <i>NPIPL3</i> | Nuclear pore complex interacting protein-like 3                                | -6.1 | -2.7 |  |
| 11q22.3 | <i>MMP3</i>   | Matrix metallopeptidase 3 (stromelysin 1, progelatinase)                       | -6.7 | -3.7 |  |
| 11q22.3 | <i>MMP1</i>   | Matrix metallopeptidase 1 (interstitial collagenase)                           | -9.9 | -6.2 |  |

---

Using the 3-fold threshold to compare genes differentially expressed between G1 and G3 and then between G2 and G3, we identified 21 genes. Fold change reflects gene expression in the Functional group 3 patients with IS relative to other 2 functional groups

**Table 9. Functional Clustering of genes differentially expressed in Functional group 3 patients with IS**

| Functional Clustering     | Gene            | Title  | P value |
|---------------------------|-----------------|--|---------|
| 1. Response to wounding   |                 |  | 2.8E-4  |
|                           | <i>PLAT</i>     | Plasminogen activator, tissue  |         |
|                           | <i>SERPINB2</i> | Serpin peptidase inhibitor, clade B (ovalbumin), member 2                      |         |
|                           | <i>ITGA2</i>    | Integrin, alpha 2  |         |
|                           | <i>PTX3</i>     | Pentraxin-related gene, rapidly induced by IL-1 beta                           |         |
|                           | <i>IL8</i>      | Interleukin 8  |         |
|                           | <i>CXCL1</i>    | Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha) |         |
| 2. Immune response        |                 |  | 9.3E-4  |
|                           | <i>OASI</i>     | 2',5'-oligoadenylate synthetase 1, 40  |         |
|                           | <i>PTX3</i>     | Pentraxin-related gene, rapidly induced by IL-1 beta                           |         |
|                           | <i>IL8</i>      | Interleukin 8  |         |
|                           | <i>CXCL1</i>    | Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha) |         |
|                           | <i>ENPP2</i>    | Ectonucleotide pyrophosphatase/phosphodiesterase 2                             |         |
|                           | <i>IL1R1</i>    | Interleukin 1 receptor, type I   |         |
| 3. Endopeptidase activity |                 |  | 1.0E-2  |
|                           | <i>PLAT</i>     | Plasminogen activator, tissue  |         |

---

|                           |   |        |
|---------------------------|---|--------|
| <i>SERPINB2</i>           | Serpin peptidase inhibitor, clade B (ovalbumin), member 2           |        |
| <i>MMP3</i>               | Matrix metallopeptidase 3 (stromelysin 1, progelatinase)            |        |
| <i>MMP1</i>               | Matrix metallopeptidase 1 (interstitial collagenase)                |        |
| 4. Polysaccharide binding |   | 1.5E-2 |
| <i>PTX3</i>               | Pentraxin-related gene, rapidly induced by IL-1 beta                |        |
| <i>ECM2</i>               | Extracellular matrix protein 2, female organ and adipocyte specific |        |
| <i>ENPP2</i>              | Ectonucleotide pyrophosphatase/phosphodiesterase 2                  |        |
| 5. Cell motion            |   | 1.7E-2 |
| <i>PLAT</i>               | Plasminogen activator, tissue                                       |        |
| <i>ITGA2</i>              | Integrin, alpha 2   |        |
| <i>IL8</i>                | Interleukin 8   |        |
| <i>ENPP2</i>              | Ectonucleotide pyrophosphatase/phosphodiesterase 2                  |        |
| 6. Calcium ion binding    |   | 2.3E-2 |
| <i>CDH6</i>               | Cadherin 6, type 2, K-cadherin (fetal kidney)                       |        |
| <i>PCDH10</i>             | Protocadherin 10  |        |
| <i>ITGA2</i>              | Integrin, alpha 2   |        |
| <i>MMP3</i>               | Matrix metallopeptidase 3 (stromelysin 1, progelatinase)            |        |
| <i>MMP1</i>               | Matrix metallopeptidase 1 (interstitial collagenase)                |        |
| 7. Defense response       |   | 3.4E-2 |
| <i>PTX3</i>               | Pentraxin-related gene, rapidly induced by IL-1 beta                |        |

---

---

|                  |  |
|------------------|--|
| <i>IL8</i>       | Interleukin 8  |
| <i>CXCL1</i>     | Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha) |
| <i>IL1R1</i>     | Interleukin 1 receptor, type I   |
| 8. Cell adhesion | 4.7E-2   |
| <i>CDH6</i>      | Cadherin 6, type 2, K-cadherin (fetal kidney)                                  |
| <i>PCDH10</i>    | Protocadherin 10   |
| <i>ITGA2</i>     | Integrin, alpha 2  |
| <i>ECM2</i>      | Extracellular matrix protein 2, female organ and adipocyte specific            |

---

Functional grouping of genes shared in groups 1 and 2 has identified here are eight significant functional clusters ( $p$ -value  $< 0.05$ ). Significance is determined by Fisher Exact test in DAVID.

**Table 10. Genes exclusively expressed in G1 vs. G2 patients**

| <b>Location</b>              | <b>Gene</b>     | <b>Title</b>  | <b>Fold<br/>change</b> |
|------------------------------|-----------------|---|------------------------|
| <b>OVER-EXPRESSED GENES</b>  |                 |   | <b>G2 vs. G1</b>       |
| 5q14.3                       | <i>HAPLN1</i>   | hyaluronan and proteoglycan link protein 1                  | 3.5                    |
| <b>UNDER-EXPRESSED GENES</b> |                 |   |                        |
| 18pter-p11.3                 | <i>COLEC12*</i> | collectin sub-family member 12                              | -3                     |
| 14q12-q13                    | <i>PAX9*</i>    | paired box 9  | -3.2                   |
| 3q11.2                       | <i>PROS1*</i>   | protein S (alpha)   | -3.2                   |
| 18q11.2                      | <i>ANKRD29</i>  | ankyrin repeat domain 29                                    | -3.3                   |
| 1q21                         | <i>CTSK*</i>    | cathepsin K   | -3.3                   |
| 6q23.1-q23.3                 | <i>MOXD1</i>    | monooxygenase, DBH-like 1                                   | -3.5                   |
| 12p13.3                      | <i>CD9*</i>     | CD9 molecule  | -3.5                   |
| 6p21.3                       | <i>TAPBP</i>    | TAP binding protein (tapasin)                               | -3.5                   |
| 8q21.13                      | <i>STMN2*</i>   | stathmin-like 2   | -3.8                   |
| 1p22.1-p21.3                 | <i>ARHGAP29</i> | Rho GTPase activating protein 29                            | -3.9                   |
| 3q26.1                       | <i>SERPNI1*</i> | serpin peptidase inhibitor, clade I (neuroserpin), member 1 | -4.0                   |

Using a 3-fold threshold, we identified 12 genes that are expressed exclusively in the comparison between G1 and G2 patients. \*Denotes genes that were functionally clustered into “multicellular organismal processes” ( $p=5.0E-2$ ) using DAVID.

**Table 11. Genes exclusively expressed in G1 vs. G3 patients**

| Location              | Gene           | Title  | Fold Change  | Enrichment Score |
|-----------------------|----------------|--|--------------|------------------|
|                       |                |  | G3 vs. G1    |                  |
| <b>Gene cluster 1</b> |                |  | <b>11.46</b> |                  |
| 17q25                 | <i>SECTM1</i>  | secreted and transmembrane 1                     | -4.8         |                  |
| 1q25.2                | <i>ANGPTL1</i> | angiopoietin-like 1                              | -3.5         |                  |
| 1q41                  | <i>HHIPL2</i>  | HHIP-like 2                                      | -3.0         |                  |
| 20p12.1               | <i>ISM1</i>    | isthmin 1 homolog (zebrafish)                    | -3.0         |                  |
| <b>Gene cluster 2</b> |                |  | <b>6.86</b>  |                  |
| 17q25                 | <i>SECTM1</i>  | secreted and transmembrane 1                     | -4.8         |                  |
| 14q24-q32             | <i>FLRT2</i>   | fibronectin leucine rich transmembrane protein 2 | -4.5         |                  |
| 11q12.3               | <i>LRRN4CL</i> | LRRN4 C-terminal like                            | -3.6         |                  |
| 2q37.3                | <i>CXCR7</i>   | chemokine (C-X-C motif) receptor 7               | -3.5         |                  |
| 4q31                  | <i>PCDH18</i>  | protocadherin 18                                 | -3.4         |                  |
| Xq13.1-q28            | <i>IL13RA2</i> | interleukin 13 receptor, alpha 2                 | -3.3         |                  |
| 17q11.2               | <i>EVI2A</i>   | ecotropic viral integration site 2A              | -3.1         |                  |
| 3q12.2                | <i>GPR128</i>  | G protein-coupled receptor 128                   | -3.1         |                  |
| 3p12                  | <i>HTR1F</i>   | 5-hydroxytryptamine (serotonin) receptor 1F      | 3.0          |                  |
| 3q12.1 3              | <i>DCBLD2</i>  | discoidin, CUB and LCCL domain containing 2      | 3.2          |                  |

---

| Gene cluster 3 |               |  | 5.23 |
|----------------|---------------|--|------|
| 14q24-q32      | <i>FLRT2</i>  | fibronectin leucine rich transmembrane protein 2           | -4.5 |
| 1p32.3         | <i>PODN</i>   | podocan  | -4.1 |
| 12q21.33       | <i>DCN</i>    | Decorin  | -3.2 |
| 12q14.1        | <i>LRIG3</i>  | leucine-rich repeats and immunoglobulin-like domains 3     | -3.1 |
| 15q23-q24      | <i>ISLR</i>   | immunoglobulin superfamily containing leucine-rich repeat  | -3.1 |
| Gene cluster 4 |               |  | 4.68 |
| 14q24.2        | <i>SMOC1</i>  | SPARC related modular calcium binding 1                    | -3.5 |
| 2p16           | <i>EFEMP1</i> | EGF-containing fibulin-like extracellular matrix protein 1 | -3.5 |
| 4p16.1         | <i>CPZ</i>    | carboxypeptidase Z   | -3.2 |
| 3p25.1         | <i>FBLN2</i>  | fibulin 2  | -3.0 |

---

Using the 3-fold threshold to compare genes differentially expressed between G1 and G3, we identified 79 genes. These genes were then analyzed using The Gene Functional Classification Tool with default settings in DAVID. There are 4 main functional clusters with significant enrichment for annotation terms. The group enrichment score ranks the biological significance of gene groups based on overall EASE scores (modified Fischer Exact p-values) of all enriched annotation terms. Higher enrichment scores denote the strongest associations.

**Table 12. The key biology modules for each gene cluster**

| Functional Clustering           | Gene            | Title   | Fold change | P value |
|---------------------------------|-----------------|---|-------------|---------|
| 1. Extracellular matrix binding |                 |   |             | 1.5E-4  |
|                                 | <i>SMOC1</i>    | SPARC related modular calcium binding 1                     | -3.5        |         |
|                                 | <i>DCN</i>      | decorin   | -3.2        |         |
|                                 | <i>FBLN2</i>    | fibulin 2   | -3.0        |         |
| 2.Blood circulation             |                 |   |             | 2.4E-4  |
|                                 | <i>ADM</i>      | adrenomedullin  | -4.3        |         |
|                                 | <i>NPR3</i>     | natriuretic peptide receptor C                              | -4.3        |         |
|                                 | <i>MEOX2</i>    | mesenchyme homeobox 2                                       | -3.8        |         |
|                                 | <i>TBC1D8</i>   | TBC1 domain family, member 8 (with GRAM domain)             | -3.5        |         |
|                                 | <i>SERPING1</i> | serpin peptidase inhibitor, clade G (C1 inhibitor), member1 | -3.4        |         |
|                                 | <i>OXTR</i>     | oxytocin receptor   | 3.4         |         |
|                                 | <i>ACTG2</i>    | actin, gamma 2, smooth muscle, enteric                      | 4.2         |         |
| 3.Calcium ion binding           |                 |   |             | 1.9E-3  |
|                                 | <i>C1S</i>      | complement component 1, s subcomponent                      | -5.4        |         |
|                                 | <i>EFEMP1</i>   | EGF-containing fibulin-like extracellular matrix protein 1  | -3.5        |         |
|                                 | <i>SMOC1</i>    | SPARC related modular calcium binding 1                     | -3.5        |         |
|                                 | <i>PCDH18</i>   | protocadherin 18  | -3.4        |         |
|                                 | <i>SULF2</i>    | sulfatase 2   | -3.2        |         |

---

|                         |  |        |
|-------------------------|--|--------|
| <i>AMY1A</i>            | amylase, alpha 1A (salivary)   | -3.1   |
| <i>ITGA8</i>            | integrin, alpha 8  | -3.1   |
| <i>FBLN2</i>            | fibulin 2  | -3.0   |
| <i>NRXN3</i>            | neurexin 3   | 3.2    |
| <i>TRPV2</i>            | transient receptor potential cation channel, subfamily V, member 2           | 3.8    |
| <i>TRPC4</i>            | transient receptor potential cation channel, subfamily C, member 4           | 4.2    |
| <i>ITGA3</i>            | integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor)         | 5.1    |
| 4.Inflammatory response |  | 4.3E-3 |
| <i>C1S</i>              | complement component 1, s subcomponent                                       | -5.4   |
| <i>CXCL6</i>            | chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2)         | -4.6   |
| <i>CD302</i>            | CD302 molecule   | -4.2   |
| <i>TNFAIP6</i>          | tumor necrosis factor, alpha-induced protein 6                               | -4.2   |
| <i>PDPN</i>             | podoplanin   | -3.6   |
| <i>CD55</i>             | CD55 molecule, decay accelerating factor for complement (Cromer blood group) | -3.5   |
| <i>SERPING1</i>         | serpin peptidase inhibitor, clade G (C1 inhibitor), member 1                 | -3.4   |

---

---

|   |  |        |
|---|--|--------|
| 5.Cell-cell signaling                               |  | 7.1E-3 |
| <i>CXCL6</i>  | chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2) | -4.6   |
| <i>ADM</i>  | adrenomedullin   | -4.3   |
| <i>TNFAIP6</i>                                      | tumor necrosis factor, alpha-induced protein 6                       | -4.2   |
| <i>MME</i>  | membrane metallo-endopeptidase                                       | -3.9   |
| <i>PDE7B</i>  | phosphodiesterase 7B   | -3.3   |
| <i>EFNA5</i>  | ephrin-A5  | -3.2   |
| <i>HTR1F</i>  | 5-hydroxytryptamine (serotonin) receptor 1F                          | 3      |
| <i>NRXN3</i>  | neurexin 3   | 3.2    |
| <i>OXTR</i>   | oxytocin receptor  | 3.4    |
| 6.ECM-receptor interaction                          |  | 1.1E-2 |
| <i>LAMA2</i>  | laminin, alpha 2   | -4.0   |
| <i>ITGA8</i>  | integrin, alpha 8  | -3.1   |
| <i>ITGB3</i>  | integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)          | 3.1    |
| <i>ITGA3</i>  | integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor) | 5.1    |
| 7.Glycosaminoglycan binding, Polysaccharide binding |  | 2.4E-2 |
| <i>CXCL6</i>  | chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2) | -4.6   |
| <i>TNFAIP6</i>                                      | tumor necrosis factor, alpha-induced protein 6                       | -4.2   |
| <i>FGFR2</i>  | fibroblast growth factor receptor 2                                  | -3.4   |

---

---

|   |  |        |
|---|--|--------|
| <i>DCN</i>                              | decorin  | -3.2   |
| 8.Regulation of actin cytoskeleton      |  | 2.9E-2 |
| <i>PDGFD</i>                            | platelet derived growth factor D                                     | -4.2   |
| <i>FGFR2</i>                            | fibroblast growth factor receptor 2                                  | -3.4   |
| <i>ITGA8</i>                            | integrin, alpha 8  | -3.1   |
| <i>ITGB3</i>                            | integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)          | 3.1    |
| <i>ITGA3</i>                            | integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor) | 5.1    |
| 9. Response to oxygen levels            |  | 2.9E-2 |
| <i>DPP4</i>                             | dipeptidyl-peptidase 4   | -5.2   |
| <i>ADM</i>                              | adrenomedullin   | -4.3   |
| <i>PDPN</i>                             | podoplanin   | -3.6   |
| <i>OXTR</i>                             | oxytocin receptor  | 3.4    |
| 10. Cell adhesion                       |  | 3.5E-2 |
| <i>FLRT2</i>                            | fibronectin leucine rich transmembrane protein 2                     | -4.5   |
| <i>PCDH18</i>                           | protocadherin 18   | -3.4   |
| <i>DCBLD2</i>                           | discoidin, CUB and LCCL domain containing 2                          | 3.2    |
| 11. Integrin-mediated signaling pathway |  | 4.3E-2 |
| <i>ITGA8</i>                            | integrin, alpha 8  | -3.1   |
| <i>ITGB3</i>                            | integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)          | 3.1    |
| <i>ITGA3</i>                            | integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor) | 5.1    |

---

---

VLA-3 receptor)

---

To consider the biological roles for those genes with over-expression and under-expression in G1 and G3, we used DAVID to classify the genes of each cluster identified in Table 11 according to shared biological modules or functional terms. \* The count signifies number of genes for the corresponding term. Significance was defined as p-value < 0.05 and was determined using Fisher Exact test

**Table 13. Genes exclusively expressed in G3 vs. G2 patients**

| <b>Location</b>              | <b>Gene</b>    | <b>Title</b>   | <b>Fold<br/>change</b> |
|------------------------------|----------------|--|------------------------|
|                              |                | <b>Symbol</b>  |                        |
| <b>OVER-EXPRESSED GENES</b>  |                |  |                        |
| 5q15                         | <i>ERAP2</i>   | endoplasmic reticulum aminopeptidase 2   | 2.9                    |
| 7p12.1                       | <i>SEMA3A</i>  | sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A | 2.5                    |
| 4q13.3                       | <i>EPGN</i>    | epithelial mitogen homolog (mouse)   | 2.3                    |
| 7q11.2                       | <i>CD36</i>    | CD36 molecule (thrombospondin receptor)  | 2.3                    |
| 7p11                         | <i>IGF2BP3</i> | insulin-like growth factor 2 mRNA binding protein 3                                    | 2.3                    |
| 4q28.1                       | <i>PRSS12</i>  | protease, serine, 12 (neurotrypsin, motopsin)  | 2.1                    |
| 5q11.2                       | <i>ESM1</i>    | endothelial cell-specific molecule 1   | 2.1                    |
| 14q22.3                      | <i>DLGAP5</i>  | discs, large (Drosophila) homolog-associated protein 5                                 | 2.1                    |
| 16q11.2                      | <i>SHCBP1</i>  | SHC SH2-domain binding protein 1   | 2.0                    |
| 6p21.3                       | <i>SCUBE3</i>  | signal peptide, CUB domain, EGF-like 3   | 2.0                    |
| 7q22                         | <i>TFPI2</i>   | tissue factor pathway inhibitor 2  | 2.0                    |
| <b>UNDER-EXPRESSED GENES</b> |                |  |                        |
| 13q13.3                      | <i>EPST11</i>  | epithelial stromal interaction 1 (breast)  | -2.0                   |
| 21q21.2                      | <i>JAM2</i>    | junctional adhesion molecule 2   | -2.0                   |
| 10q23-q25                    | <i>IFIT2</i>   | interferon-induced protein with tetratricopeptide repeats 2                            | -2.1                   |
| 8q24.1-q24.3                 | <i>WISPI1</i>  | WNT1 inducible signaling pathway protein 1   | -2.1                   |

---

|           |                |  |      |
|-----------|----------------|--|------|
| 17p13.1   | <i>XAF1</i>    | XIAP associated factor 1   | -2.1 |
| 1q32-q41  | <i>KCNH1</i>   | potassium voltage-gated channel, subfamily H (eag-related), member 1               | -2.1 |
| 11p15.5   | <i>IGF2</i>    | insulin-like growth factor 2 (somatomedin A)                                       | -2.1 |
| 2q24      | <i>IFIH1</i>   | interferon induced with helicase C domain 1  | -2.2 |
| 12q23.3   | <i>TMEM119</i> | transmembrane protein 119  | -2.2 |
| 12q15     | <i>IL26</i>    | interleukin 26   | -2.2 |
| 21q22.3   | <i>MX1</i>     | myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse) | -2.3 |
| 17q12-q21 | <i>RNU2-1</i>  | RNA, U2 small nuclear 1  | -2.3 |
| 10q25-q26 | <i>IFIT1</i>   | interferon-induced protein with tetratricopeptide repeats 1                        | -2.4 |
| 6p21.3    | <i>UBD</i>     | ubiquitin D  | -2.4 |
| 11q22.3   | <i>MMP13</i>   | matrix metallopeptidase 13 (collagenase 3)   | -2.5 |
| 11q13.4   | <i>SPCS2</i>   | signal peptidase complex subunit 2 homolog (S. cerevisiae)                         | -2.5 |
| 2p25.2    | <i>RSAD2</i>   | radical S-adenosyl methionine domain containing 2                                  | -2.6 |

---

Using a 2-fold threshold, we identified 28 genes that are expressed differently only between G2 and G3 patients.

**Table 14. Functional Clustering of genes exclusively expressed in G2 vs. G3 patients**

| Functional<br>Clustering                             | Gene          | Title  | Fold change | P value |
|--|---------------|--|-------------|---------|
| 1. Regulation of Growth                              |               |  |             | 5.3E-4  |
|  | <i>CD36</i>   | CD36 molecule (thrombospondin receptor)  | 2.3         |         |
|  | <i>ESM1</i>   | endothelial cell-specific molecule 1   | 2.1         |         |
|  | <i>WISP1</i>  | WNT1 inducible signaling pathway protein 1   | -2.1        |         |
|  | <i>IGF2</i>   | insulin-like growth factor 2 (somatomedin A)                                       | -2.1        |         |
| 2. Positive regulation of mitosis (nuclear division) |               |  |             | 6.7E-4  |
|  | <i>EPGN</i>   | epithelial mitogen homolog (mouse)   | 2.3         |         |
|  | <i>DLGAP5</i> | discs, large (Drosophila) homolog-associated protein 5                             | 2.1         |         |
|  | <i>IGF2</i>   | insulin-like growth factor 2 (somatomedin A)                                       | -2.1        |         |
| 3. Response to virus (defense)                       |               |  |             | 1.2E-2  |
|  | <i>IFIH1</i>  | interferon induced with helicase C domain 1  | -2.2        |         |
|  | <i>MX1</i>    | myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse) | -2.3        |         |
|  | <i>RSAD2</i>  | radical S-adenosyl methionine domain containing 2                                  | -2.6        |         |
| 4. Protease activity                                 |               |  |             | 2.7E-2  |
|  | <i>ERAP2</i>  | endoplasmic reticulum aminopeptidase 2   | 2.9         |         |
|  | <i>PRSS12</i> | protease, serine, 12 (neurotrypsin, motopsin)                                      | 2.1         |         |

---

|              |  |      |
|--------------|--|------|
| <i>MMP13</i> | matrix metallopeptidase 13 (collagenase 3)                 | -2.5 |
| <i>SPCS2</i> | signal peptidase complex subunit 2 homolog (S. cerevisiae) | -2.5 |

---

\* Significance was defined as p-value < 0.05

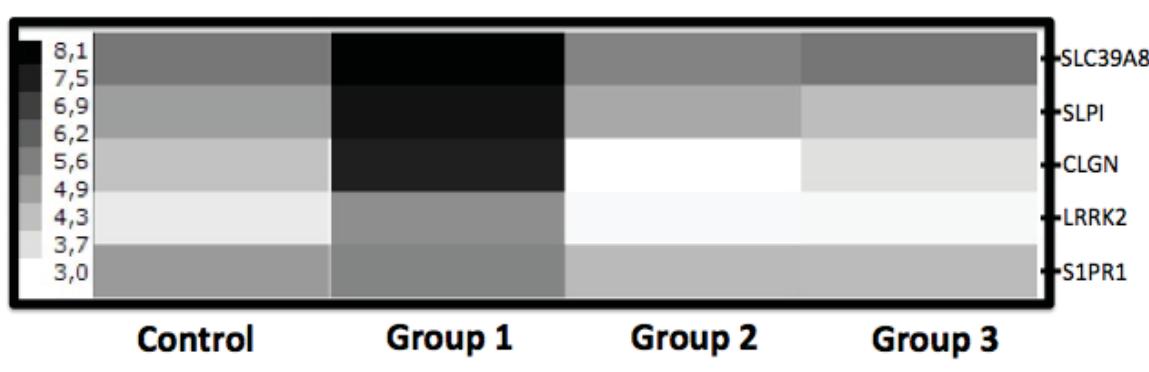
The table describes four functional clusters with significance: regulation of growth, positive regulation of mitosis, response to virus, and protease activity. Here we compare the fold change difference between Functional group 2 and Functional group 3, we manage to use the 2-fold threshold instead of 3-fold in other 2 comparison in order to increase the number of genes in this group by group comparison.

**Table 15. The ANOVA Table**

| Gene               | Location      | FDR           |          | G1<br>mean | G1<br>variance | G2<br>mean | G2<br>variance | G3<br>mean | G3<br>variance |
|--------------------|---------------|---------------|----------|------------|----------------|------------|----------------|------------|----------------|
|                    |               | Correcte<br>P | Raw P    |            |                |            |                |            |                |
| <i>CLGN*</i>       | 4q28.3-q31.1  | 0.00882       | 2.65E-07 | 7.481      | 0.1538         | 3.014      | 0.1228         | 3.014      | 0.1228         |
| <i>ISM1</i>        | 20p12.1       | 0.01568       | 9.42E-07 | 5.448      | 0.0560         | 3.954      | 0.0594         | 3.954      | 0.0594         |
| <i>MKX</i>         | 10p12.1       | 0.01582       | 1.43E-06 | 7.760      | 0.0841         | 5.448      | 0.1633         | 5.448      | 0.1634         |
| <i>MOXD1</i>       | 6q23.1-q23.3  | 0.02152       | 2.59E-06 | 8.881      | 0.0018         | 7.030      | 0.1014         | 7.030      | 0.1014         |
| <i>SLPI*</i>       | 20q12         | 0.02520       | 4.54E-06 | 7.782      | 0.4590         | 4.751      | 0.3653         | 4.751      | 0.3653         |
| <i>ABCC1</i>       | 16p13.1       | 0.02520       | 3.92E-06 | 8.670      | 0.0727         | 7.297      | 0.0688         | 7.209      | 0.0148         |
| <i>C13orf18</i>    | 13q14.13      | 0.02551       | 5.36E-06 | 4.719      | 0.2041         | 3.322      | 0.0183         | 3.322      | 0.0183         |
| <i>LRRK2*</i>      | 12q12         | 0.03608       | 1.04E-05 | 5.278      | 0.1236         | 3.114      | 0.1464         | 3.114      | 0.1464         |
| <i>GALNTL2*</i>    | 3p25.1        | 0.03608       | 1.04E-05 | 8.392      | 0.0126         | 5.813      | 0.1995         | 5.812      | 0.1995         |
| <i>SIPRI*</i>      | 1p21          | 0.03608       | 1.08E-05 | 5.468      | 0.0106         | 4.378      | 0.0560         | 4.378      | 0.0560         |
| <i>TBC1D8</i>      | 2q11.2        | 0.04046       | 1.55E-05 | 6.142      | 0.2340         | 4.657      | 0.1042         | 4.657      | 0.1042         |
| <i>RNFT2</i>       | 12q24.22      | 0.04046       | 2.19E-05 | 4.147      | 0.0029         | 4.623      | 0.0088         | 4.623      | 0.0088         |
| <i>SLC2A12</i>     | 6q23.2        | 0.04046       | 2.17E-05 | 7.709      | 0.2470         | 4.052      | 0.8294         | 4.052      | 0.8294         |
| <i>Has-mir-29c</i> | chr1          | 0.04046       | 1.89E-05 | 4.651      | 0.0331         | 2.978      | 0.1991         | 2.978      | 0.1991         |
| <i>SLC39AB*</i>    | 4q22-q24      | 0.04046       | 1.66E-05 | 8.131      | 0.6473         | 5.491      | 0.1372         | 5.491      | 0.1372         |
| <i>KCNJ2</i>       | 17q23.1-q24.2 | 0.04046       | 1.91E-05 | 6.831      | 0.9678         | 4.323      | 0.0517         | 4.323      | 0.0517         |
| <i>CHRDL1</i>      | Xq23          | 0.04046       | 1.54E-05 | 6.874      | 0.0361         | 4.549      | 0.1965         | 4.549      | 0.1965         |

|                 |            |         |          |       |        |       |        |       |        |
|-----------------|------------|---------|----------|-------|--------|-------|--------|-------|--------|
| <b>MUSK</b>     | 9q31.3-q32 | 0.04046 | 2.05E-05 | 3.023 | 0.0056 | 2.584 | 0.0048 | 2.584 | 0.0048 |
| <b>ANKRD28</b>  | 3p25.1     | 0.04560 | 2.60E-05 | 8.338 | 0.0991 | 7.441 | 0.0234 | 7.441 | 0.0234 |
| <b>ANGPTL5</b>  | 11q22.1    | 0.04746 | 2.85E-05 | 5.674 | 1.0328 | 3.070 | 0.1822 | 3.067 | 0.1822 |
| <b>FAM49A</b>   | 2p24.2     | 0.04963 | 3.13E-05 | 6.157 | 0.0331 | 4.711 | 0.0810 | 4.711 | 0.0810 |
| <b>SCL38A6</b>  | 14q23.1    | 0.05053 | 3.93E-05 | 7.685 | 0.0041 | 6.624 | 0.0575 | 6.623 | 0.0575 |
| <b>IRS2</b>     | 13q34      | 0.05053 | 3.79E-05 | 6.068 | 0.0128 | 5.619 | 0.0225 | 5.619 | 0.0225 |
| <b>HTR1F</b>    | 3p12       | 0.05053 | 3.95E-05 | 3.068 | 0.0499 | 3.794 | 0.1521 | 3.794 | 0.1521 |
| <b>SERPINI1</b> | 3q26.1     | 0.05053 | 3.56E-05 | 5.665 | 0.0442 | 3.614 | 0.1675 | 3.614 | 0.1675 |
| <b>PAX9</b>     | 14q12-q13  | 0.05053 | 3.70E-05 | 6.468 | 0.4096 | 4.763 | 0.0628 | 4.763 | 0.0627 |

The sample size for G1 was 3; for G2 sample size was 6; for G3 sample size was 5. The ANOVA was performed with 2 DF. All mean and variance values are shown as Log2 transformed values. \*Indicates genes that are significant at FDR corrected  $p \leq 0.05$ .



**Figure 2. Significantly differentially expressed genes inclusive of controls**

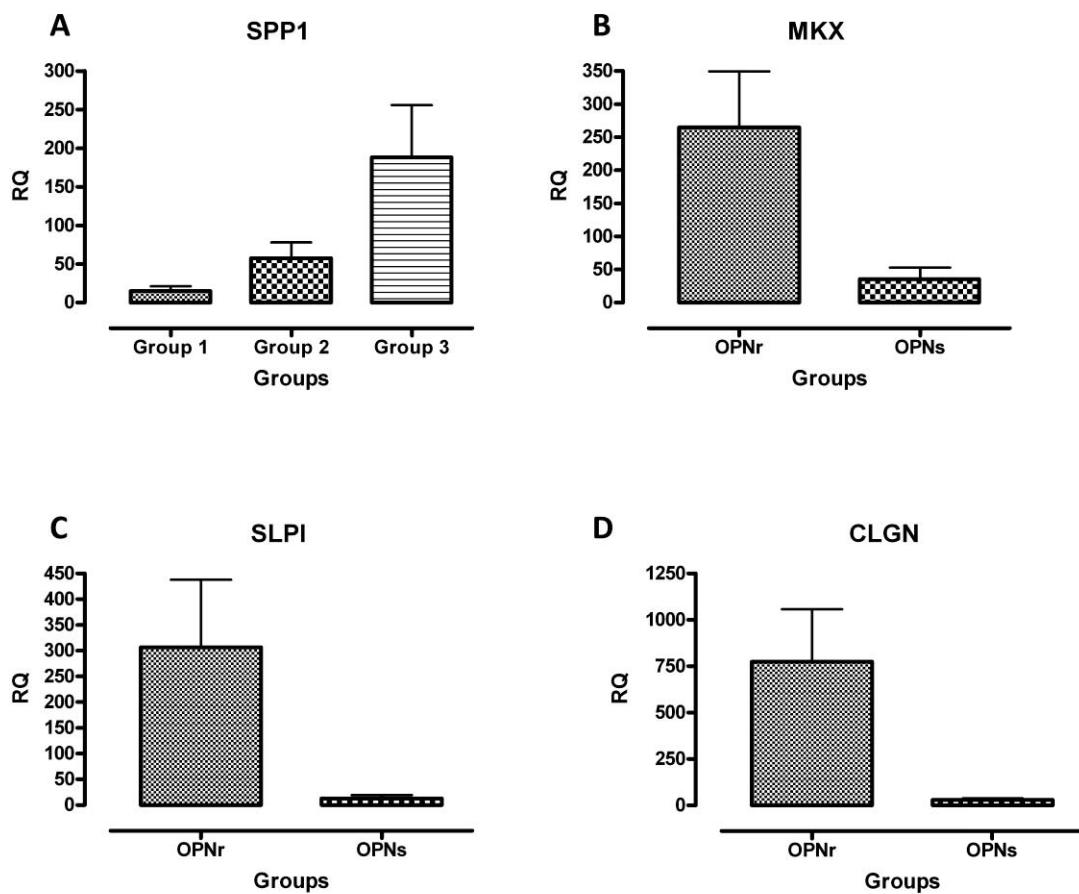
The heatmap above shows the expression profiles for each of the 5 genes that are significantly differentially expressed after correction for multiple testing (FDR corrected p-value  $\leq 0.05$ ). The scale to the left correlates log2 transformed mean expression values with the intensity of greyscale. A darker shade is representative of relative overexpression.

**Table 16. 22 candidate genes were chosen to be tested on a larger cohort using RT-QPCR.**

| GENE          | GENE   | LOCATION      | RESULTS                      |
|---------------|--|---------------|------------------------------|
| <i>MMP1</i>   | matrix metallopeptidase 1                                  | 11q22.3       | Not associated               |
| <i>CHI3L1</i> | chitinase 3-like 1 (cartilage glycoprotein-39)             | 1q32.1        | Not associated               |
| <i>EFEMP1</i> | EGF-containing fibulin-like extracellular matrix protein 1 | 2p16          | Not associated               |
| <i>SPP1</i>   | secreted phosphoprotein 1                                  | 4q21-q25      | Associated with Groups       |
| <i>COMP</i>   | cartilage oligomeric matrix protein                        | 19p13.1       | Not associated               |
| <i>PDE1A</i>  | phosphodiesterase 1A,<br>calmodulin-dependent              | 2q32.1        | Not associated               |
| <i>CLU</i>    | clusterin  | 8p21-p12      | Not associated               |
| <i>LRRC17</i> | leucine rich repeat containing 17                          | 7q22.1        | Not associated               |
| <i>RGS4</i>   | regulator of G-protein signaling 4                         | 1q23.3        | Not associated               |
| <i>PTGER2</i> | Homo sapiens prostaglandin E receptor 2 (subtype EP2)      | 14q22         | Not associated               |
| <i>BDKRB1</i> | bradykinin receptor B1                                     | 14q32.1-q32.2 | Not associated               |
| <i>FGL2</i>   | fibrinogen-like 2  | 7q11.23       | Not associated               |
| <i>CDH2</i>   | cadherin 2, type 1   | 18q11.2       | Not associated               |
| <i>CLGN</i>   | calmegin   | 4q28.3-q31.1  | Resistant vs. Sensitive only |

|                     |  |          |                              |
|---------------------|--|----------|------------------------------|
| <b><i>SLPI</i></b>  | secretory leukocyte peptidase inhibitor                                    | 20q12    | Resistant vs. Sensitive only |
| <b><i>MKX</i></b>   | mohawk homeobox  | 10p12.1  | Resistant vs. Sensitive only |
| <b><i>ACTG</i></b>  | actin, gamma 2, smooth muscle, enteric                                     | 2p13.1   | Not associated               |
| <b><i>ITGB3</i></b> | integrin, beta 3   | 17q21.32 | Not associated               |
| <b><i>SEPP1</i></b> | selenoprotein P, plasma, 1   | 5q31     | Not associated               |
| <b><i>OXTR</i></b>  | oxytocin receptor  | 3p25     | Not associated               |
| <b><i>SVEP1</i></b> | sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1 | 9q32     | Not associated               |
| <b><i>ITGA8</i></b> | integrin, alpha 8  | 10p13    | Not associated               |

**Figure 3. Candidate genes significantly differentially expressed via RT-QPCR**



Confirmation of candidate genes from qualitative exploration (*SPP1*) or quantitative analysis (*MKX*, *SLPI*, *CLGN*) show expression patterns similar to the biochemical profiles. **A.** *SPP1* has very low expression in G1, biochemically characterized as OPN resistant.

## References

1. Dickson, R., *Early-onset idiopathic scoliosis*. In The Pediatric Spine, 2nd edition, Stewart L. Weinstein, ed. Lippincott Williams & Wilkins, Philadelphia, PA, USA, 2001.
2. Warner, W., *Juvenile idiopathic scoliosis*. In The Pediatric Spine, 2 edition, Stewart L. Weinstein, ed. Lippincott Williams & Wilkins, Philadelphia, PA, USA, 2001.
3. Miller, N., *Adolescent idiopathic scoliosis: etiology*. In The Pediatric Spine, 2nd edition, Stewart L. Weinstein, ed. Lippincott Williams & Wilkins, Philadelphia, PA, USA., 2001.
4. Lowe, T.G., et al., *Etiology of idiopathic scoliosis: current trends in research*. J Bone Joint Surg Am, 2000. **82-A**(8): p. 1157-68.
5. Garland, H.G., *Hereditary Scoliosis*. Br Med J, 1934. **1**(3816): p. 328.
6. Cowell, H.R., *Genetic aspects of orthopedic diseases*. Am J Nurs, 1970. **70**(4): p. 763-7.
7. Filho NA, T.M., *Genetic studies in scoliosis*. J Bone Joint Surg [Am], 1971. **53**: p. 199.
8. Riseborough, E.J. and R. Wynne-Davies, *A genetic survey of idiopathic scoliosis in Boston, Massachusetts*. J Bone Joint Surg Am, 1973. **55**(5): p. 974-82.
9. Robin, G.C. and T. Cohen, *Familial scoliosis. A clinical report*. J Bone Joint Surg Br, 1975. **57**(2): p. 146-8.
10. Byrd, J.A., 3rd, *Current theories on the etiology of idiopathic scoliosis*. Clin Orthop Relat Res, 1988(229): p. 114-9.
11. Bell, M. and A.S. Teebi, *Autosomal dominant idiopathic scoliosis?* Am J Med Genet, 1995. **55**(1): p. 112.
12. Miller, N.H., *Cause and natural history of adolescent idiopathic scoliosis*. Orthop Clin North Am, 1999. **30**(3): p. 343-52, vii.
13. Hadley Miller, N., *Spine update: genetics of familial idiopathic scoliosis*. Spine (Phila Pa 1976), 2000. **25**(18): p. 2416-8.

14. Kouwenhoven, J.W. and R.M. Castelein, *The pathogenesis of adolescent idiopathic scoliosis: review of the literature*. Spine (Phila Pa 1976), 2008. **33**(26): p. 2898-908.
15. Burwell, R.G., *Aetiology of idiopathic scoliosis: current concepts*. Pediatr Rehabil, 2003. **6**(3-4): p. 137-70.
16. Giampietro, P.F., et al., *Congenital and idiopathic scoliosis: clinical and genetic aspects*. Clin Med Res, 2003. **1**(2): p. 125-36.
17. Cheng, J.C., et al., *Genetic association of complex traits: using idiopathic scoliosis as an example*. Clin Orthop Relat Res, 2007. **462**: p. 38-44.
18. Miller, N.H., *Genetics of familial idiopathic scoliosis*. Clin Orthop Relat Res, 2007. **462**: p. 6-10.
19. Cheung, K.M., et al., *Recent advances in the aetiology of adolescent idiopathic scoliosis*. Int Orthop, 2008. **32**(6): p. 729-34.
20. Heary, R.F. and K. Madhavan, *Genetics of scoliosis*. Neurosurgery, 2008. **63**(3 Suppl): p. 222-7.
21. Wise, C.A., et al., *Understanding genetic factors in idiopathic scoliosis, a complex disease of childhood*. Curr Genomics, 2008. **9**(1): p. 51-9.
22. Wynne-Davies, R., *Familial (idiopathic) scoliosis. A family survey*. J Bone Joint Surg Br, 1968. **50**(1): p. 24-30.
23. Cowell, H.R., J.N. Hall, and G.D. MacEwen, *Genetic aspects of idiopathic scoliosis. A Nicholas Andry Award essay, 1970*. Clin Orthop Relat Res, 1972. **86**: p. 121-31.
24. Esteve, R., *Idiopathic scoliosis in identical twins*. J Bone Joint Surg Br, 1958. **40-B**(1): p. 97-9.
25. Gaertner, R.L., *Idiopathic scoliosis in identical (monozygotic) twins*. South Med J, 1979. **72**(2): p. 231-4.
26. Carr, A.J., *Adolescent idiopathic scoliosis in identical twins*. J Bone Joint Surg Br, 1990. **72**(6): p. 1077.

27. Kesling, K.L. and K.A. Reinker, *Scoliosis in twins. A meta-analysis of the literature and report of six cases*. Spine (Phila Pa 1976), 1997. **22**(17): p. 2009-14; discussion 2015.
28. Inoue, M., et al., *Idiopathic scoliosis in twins studied by DNA fingerprinting: the incidence and type of scoliosis*. J Bone Joint Surg Br, 1998. **80**(2): p. 212-7.
29. Andersen, M.O., K. Thomsen, and K.O. Kyvik, *Adolescent idiopathic scoliosis in twins: a population-based survey*. Spine (Phila Pa 1976), 2007. **32**(8): p. 927-30.
30. The, N.S., L.S. Adair, and P. Gordon-Larsen, *A study of the birth weight-obesity relation using a longitudinal cohort and sibling and twin pairs*. Am J Epidemiol, 2010. **172**(5): p. 549-57.
31. Wise, C.A., et al., *Localization of susceptibility to familial idiopathic scoliosis*. Spine (Phila Pa 1976), 2000. **25**(18): p. 2372-80.
32. Salehi, L.B., et al., *Assignment of a locus for autosomal dominant idiopathic scoliosis (IS) to human chromosome 17p11*. Hum Genet, 2002. **111**(4-5): p. 401-4.
33. Justice, C.M., et al., *Familial idiopathic scoliosis: evidence of an X-linked susceptibility locus*. Spine (Phila Pa 1976), 2003. **28**(6): p. 589-94.
34. Ocaka, L., et al., *Assignment of two loci for autosomal dominant adolescent idiopathic scoliosis to chromosomes 9q31.2-q34.2 and 17q25.3-qtel*. J Med Genet, 2008. **45**(2): p. 87-92.
35. Gurnett, C.A., et al., *Genetic linkage localizes an adolescent idiopathic scoliosis and pectus excavatum gene to chromosome 18 q*. Spine (Phila Pa 1976), 2009. **34**(2): p. E94-100.
36. Edery, P., et al., *New disease gene location and high genetic heterogeneity in idiopathic scoliosis*. Eur J Hum Genet, 2011. **19**(8): p. 865-9.
37. Chan, V., et al., *A genetic locus for adolescent idiopathic scoliosis linked to chromosome 19p13.3*. Am J Hum Genet, 2002. **71**(2): p. 401-6.
38. Alden, K.J., et al., *Idiopathic scoliosis: identification of candidate regions on chromosome 19p13*. Spine (Phila Pa 1976), 2006. **31**(16): p. 1815-9.

39. Raggio, C.L., et al., *A novel locus for adolescent idiopathic scoliosis on chromosome 12p*. J Orthop Res, 2009. **27**(10): p. 1366-72.
40. Gao, X., et al., *CHD7 gene polymorphisms are associated with susceptibility to idiopathic scoliosis*. Am J Hum Genet, 2007. **80**(5): p. 957-65.
41. Cardon, L.R. and J.I. Bell, *Association study designs for complex diseases*. Nat Rev Genet, 2001. **2**(2): p. 91-9.
42. Gorman, K.F., C. Julien, and A. Moreau, *The genetic epidemiology of idiopathic scoliosis*. Eur Spine J, 2012.
43. Pincott, J.R., J.S. Davies, and L.F. Taffs, *Scoliosis caused by section of dorsal spinal nerve roots*. J Bone Joint Surg Br, 1984. **66**(1): p. 27-9.
44. Fisk, J.R. and W.H. Bunch, *Scoliosis in neuromuscular disease*. Orthop Clin North Am, 1979. **10**(4): p. 863-75.
45. Lonstein, J.E. and T.S. Renshaw, *Neuromuscular spine deformities*. Instr Course Lect, 1987. **36**: p. 285-304.
46. Manzoni, D. and F. Miele, *Vestibular mechanisms involved in idiopathic scoliosis*. Arch Ital Biol, 2002. **140**(1): p. 67-80.
47. Herman, R., et al., *Idiopathic scoliosis and the central nervous system: a motor control problem. The Harrington lecture, 1983*. Scoliosis Research Society. Spine (Phila Pa 1976), 1985. **10**(1): p. 1-14.
48. Guo, X., et al., *Balance control in adolescents with idiopathic scoliosis and disturbed somatosensory function*. Spine (Phila Pa 1976), 2006. **31**(14): p. E437-40.
49. Lao, M.L., et al., *Impaired dynamic balance control in adolescents with idiopathic scoliosis and abnormal somatosensory evoked potentials*. J Pediatr Orthop, 2008. **28**(8): p. 846-9.

50. Mahaudens, P., et al., *Gait in adolescent idiopathic scoliosis: energy cost analysis*. Eur Spine J, 2009. **18**(8): p. 1160-8.
51. Domenech, J., et al., *Motor cortical hyperexcitability in idiopathic scoliosis: could focal dystonia be a subclinical etiological factor?* Eur Spine J, 2010. **19**(2): p. 223-30.
52. Zorab, P.A., et al., *Bone collagen turnover in idiopathic scoliosis estimated from total hydroxyproline excretion*. Arch Dis Child, 1971. **46**(250): p. 828-32.
53. Shapiro, J.R., et al., *Osteoporosis and familial idiopathic scoliosis: association with an abnormal alpha 2(I) collagen*. Connect Tissue Res, 1989. **21**(1-4): p. 117-23; discussion 124.
54. Miller, N.H., et al., *Genetic analysis of structural elastic fiber and collagen genes in familial adolescent idiopathic scoliosis*. J Orthop Res, 1996. **14**(6): p. 994-9.
55. Guo, X., et al., *Relative anterior spinal overgrowth in adolescent idiopathic scoliosis--result of disproportionate endochondral-membranous bone growth? Summary of an electronic focus group debate of the IBSE*. Eur Spine J, 2005. **14**(9): p. 862-73.
56. Man, G.C., et al., *Abnormal melatonin receptor 1B expression in osteoblasts from girls with adolescent idiopathic scoliosis*. J Pineal Res, 2011. **50**(4): p. 395-402.
57. Wang, W.J., et al., *Top theories for the etiopathogenesis of adolescent idiopathic scoliosis*. J Pediatr Orthop, 2011. **31**(1 Suppl): p. S14-27.
58. Fidler, M.W. and R.L. Jowett, *Muscle imbalance in the aetiology of scoliosis*. J Bone Joint Surg Br, 1976. **58**(2): p. 200-1.
59. Reuber, M., et al., *Trunk muscle myoelectric activities in idiopathic scoliosis*. Spine (Phila Pa 1976), 1983. **8**(5): p. 447-56.
60. Gonyea, W.J., et al., *An evaluation of muscle pathology in idiopathic scoliosis*. J Pediatr Orthop, 1985. **5**(3): p. 323-9.
61. Avikainen, V.J., A. Rezasoltani, and H.A. Kauhanen, *Asymmetry of paraspinal EMG-time characteristics in idiopathic scoliosis*. J Spinal Disord, 1999. **12**(1): p. 61-7.

62. Cheung, J., et al., *A preliminary study on electromyographic analysis of the paraspinal musculature in idiopathic scoliosis*. Eur Spine J, 2005. **14**(2): p. 130-7.
63. Martinez-Llorens, J., et al., *Muscle dysfunction and exercise limitation in adolescent idiopathic scoliosis*. Eur Respir J, 2010. **36**(2): p. 393-400.
64. Kahmann, R.D., et al., *Platelet function in adolescent idiopathic scoliosis*. Spine (Phila Pa 1976), 1992. **17**(2): p. 145-8.
65. Lowe, T., et al., *Platelet calmodulin levels in adolescent idiopathic scoliosis: do the levels correlate with curve progression and severity?* Spine (Phila Pa 1976), 2002. **27**(7): p. 768-75.
66. Misol, S., et al., *Growth hormone blood levels in patients with idiopathic scoliosis*. Clin Orthop Relat Res, 1971. **81**: p. 122-5.
67. Skogland, L.B. and J.A. Miller, *Growth related hormones in idiopathic scoliosis. An endocrine basis for accelerated growth*. Acta Orthop Scand, 1980. **51**(5): p. 779-80.
68. Fagan, A.B., D.J. Kennaway, and A.D. Sutherland, *Total 24-hour melatonin secretion in adolescent idiopathic scoliosis. A case-control study*. Spine (Phila Pa 1976), 1998. **23**(1): p. 41-6.
69. Brodner, W., et al., *Melatonin and adolescent idiopathic scoliosis*. J Bone Joint Surg Br, 2000. **82**(3): p. 399-403.
70. Barrios, C., et al., *Anthropometry and body composition profile of girls with nonsurgically treated adolescent idiopathic scoliosis*. Spine (Phila Pa 1976), 2011. **36**(18): p. 1470-7.
71. Lombardi, G., et al., *Biochemistry of adolescent idiopathic scoliosis*. Adv Clin Chem, 2011. **54**: p. 165-82.
72. O'Tuathaigh, C.M., et al., *Susceptibility Genes for Schizophrenia: Mutant Models, Endophenotypes and Psychobiology*. Curr Top Behav Neurosci, 2012.
73. Haeffel, G.J., M. Eastman, and E.L. Grigorenko, *Using a cognitive endophenotype to identify risk genes for depression*. Neurosci Lett, 2012. **510**(1): p. 10-3.

74. Turetsky, B.I., et al., *Neurophysiological endophenotypes of schizophrenia: the viability of selected candidate measures*. Schizophr Bull, 2007. **33**(1): p. 69-94.
75. Akoume, M.Y., et al., *Cell-based screening test for idiopathic scoliosis using cellular dielectric spectroscopy*. Spine (Phila Pa 1976), 2010. **35**(13): p. E601-8.
76. Irizarry, R.A., et al., *Summaries of Affymetrix GeneChip probe level data*. Nucleic Acids Res, 2003. **31**(4): p. e15.
77. Blazejczyk M, M.M., Nadon R, *FlexArray: A statistical data analysis software for gene expression microarrays*. 2007.
78. Hu, J.X., H. Zhao, and H.H. Zhou, *False Discovery Rate Control With Groups*. J Am Stat Assoc, 2010. **105**(491): p. 1215-1227.
79. Dennis, G., Jr., et al., *DAVID: Database for Annotation, Visualization, and Integrated Discovery*. Genome Biol, 2003. **4**(5): p. P3.
80. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method*. Methods, 2001. **25**(4): p. 402-8.
81. Xi, G., et al., *Expression of the human beta3 integrin subunit in mouse smooth muscle cells enhances IGF-I-stimulated signaling and proliferation*. J Cell Physiol, 2008. **214**(2): p. 306-15.
82. Yoshimoto, N., et al., *Genetic and environmental predictors, endogenous hormones and growth factors, and risk of estrogen receptor-positive breast cancer in Japanese women*. Cancer Sci, 2011. **102**(11): p. 2065-72.
83. Cui, R., et al., *Abrogation of the interaction between osteopontin and alphavbeta3 integrin reduces tumor growth of human lung cancer cells in mice*. Lung Cancer, 2007. **57**(3): p. 302-10.
84. Kim, T., et al., *Identification of LRRc17 as a negative regulator of receptor activator of NF-kappaB ligand (RANKL)-induced osteoclast differentiation*. J Biol Chem, 2009. **284**(22): p. 15308-16.

85. Moreau, A., et al., *Melatonin signaling dysfunction in adolescent idiopathic scoliosis*. Spine (Phila Pa 1976), 2004. **29**(16): p. 1772-81.
86. Azeddine, B., et al., *Molecular determinants of melatonin signaling dysfunction in adolescent idiopathic scoliosis*. Clin Orthop Relat Res, 2007. **462**: p. 45-52.
87. Moreau, A., et al., [Molecular and genetic aspects of idiopathic scoliosis. Blood test for idiopathic scoliosis]. Orthopade, 2009. **38**(2): p. 114-6, 118-21.
88. Austin, C.E., et al., *Stable expression of the human kinin B1 receptor in Chinese hamster ovary cells. Characterization of ligand binding and effector pathways*. J Biol Chem, 1997. **272**(17): p. 11420-5.
89. Gutowski, S., et al., *Antibodies to the alpha q subfamily of guanine nucleotide-binding regulatory protein alpha subunits attenuate activation of phosphatidylinositol 4,5-bisphosphate hydrolysis by hormones*. J Biol Chem, 1991. **266**(30): p. 20519-24.
90. Kang, D.S. and L.M. Leeb-Lundberg, *Negative and positive regulatory epitopes in the C-terminal domains of the human B1 and B2 bradykinin receptor subtypes determine receptor coupling efficacy to G(q/11)-mediated [correction of G(9/11)-mediated] phospholipase C $\beta$  activity*. Mol Pharmacol, 2002. **62**(2): p. 281-8.
91. Seifert, R. and K. Wenzel-Seifert, *Constitutive activity of G-protein-coupled receptors: cause of disease and common property of wild-type receptors*. Naunyn Schmiedebergs Arch Pharmacol, 2002. **366**(5): p. 381-416.
92. Tilakaratne, N. and P.M. Sexton, *G-Protein-coupled receptor-protein interactions: basis for new concepts on receptor structure and function*. Clin Exp Pharmacol Physiol, 2005. **32**(11): p. 979-87.
93. Luttrell, L.M., *Transmembrane signaling by G protein-coupled receptors*. Methods Mol Biol, 2006. **332**: p. 3-49.
94. Tuteja, N., *Signaling through G protein coupled receptors*. Plant Signal Behav, 2009. **4**(10): p. 942-7.

95. Ritter, S.L. and R.A. Hall, *Fine-tuning of GPCR activity by receptor-interacting proteins*. Nat Rev Mol Cell Biol, 2009. **10**(12): p. 819-30.
96. Schweickart, V.L., et al., *CCR11 is a functional receptor for the monocyte chemoattractant protein family of chemokines*. J Biol Chem, 2000. **275**(13): p. 9550-6.
97. Liu, Y., et al., *Characterization of C-C chemokine receptor subfamily in teleost fish*. Mol Immunol, 2009. **46**(3): p. 498-504.
98. Gorman, K.F. and F. Breden, *Teleosts as models for human vertebral stability and deformity*. Comp Biochem Physiol C Toxicol Pharmacol, 2007. **145**(1): p. 28-38.
99. Abera, A.B., et al., *EP2 receptor mediated cAMP release is augmented by PGF 2 alpha activation of the FP receptor via the calcium-calmodulin pathway*. Cell Signal, 2010. **22**(1): p. 71-9.
100. Goldberg, M., et al., *Fibromodulin-deficient mice display impaired collagen fibrillogenesis in predentin as well as altered dentin mineralization and enamel formation*. J Histochem Cytochem, 2006. **54**(5): p. 525-37.
101. Embree, M., et al., *Role of subchondral bone during early-stage experimental TMJ osteoarthritis*. J Dent Res, 2011. **90**(11): p. 1331-8.
102. Nakanishi, R., et al., *Secreted frizzled-related protein 4 is a negative regulator of peak BMD in SAMP6 mice*. J Bone Miner Res, 2006. **21**(11): p. 1713-21.
103. Nakanishi, R., et al., *Osteoblast-targeted expression of Sfrp4 in mice results in low bone mass*. J Bone Miner Res, 2008. **23**(2): p. 271-7.
104. Cho, H.Y., et al., *Transgenic mice overexpressing secreted frizzled-related proteins (sFRP)4 under the control of serum amyloid P promoter exhibit low bone mass but did not result in disturbed phosphate homeostasis*. Bone, 2010. **47**(2): p. 263-71.
105. Fendri, K. and F. Moldovan, *Potential role of COMP as a biomarker for adolescent idiopathic scoliosis*. Med Hypotheses, 2011. **76**(5): p. 762-3.

106. Rock, M.J., et al., *Cartilage oligomeric matrix protein promotes cell attachment via two independent mechanisms involving CD47 and alphaVbeta3 integrin*. Mol Cell Biochem, 2010. **338**(1-2): p. 215-24.
107. Schedel, J., et al., *Differential adherence of osteoarthritis and rheumatoid arthritis synovial fibroblasts to cartilage and bone matrix proteins and its implication for osteoarthritis pathogenesis*. Scand J Immunol, 2004. **60**(5): p. 514-23.
108. Xiong, D.H., et al., *Genome-wide association and follow-up replication studies identified ADAMTS18 and TGFBR3 as bone mass candidate genes in different ethnic groups*. Am J Hum Genet, 2009. **84**(3): p. 388-98.
109. Dastych, M., J. Cienciala, and M. Krbec, *Changes of selenium, copper, and zinc content in hair and serum of patients with idiopathic scoliosis*. J Orthop Res, 2008. **26**(9): p. 1279-82.
110. Zhu, Y., et al., *Comparative analysis of serum proteomes of degenerative scoliosis*. J Orthop Res, 2011. **29**(12): p. 1896-903.
111. Tsuji, T., et al., *Analysis of chondrex (YKL-40, HC gp-39) in the cerebrospinal fluid of patients with spine disease*. Spine (Phila Pa 1976), 2002. **27**(7): p. 732-5.
112. Pozzuoli, A., et al., *YKL-40 in human lumbar herniated disc and its relationships with nitric oxide and cyclooxygenase-2*. Clin Exp Rheumatol, 2007. **25**(3): p. 453-6.
113. Morocz, M., et al., *Association study of BMP4, IL6, Leptin, MMP3, and MTNR1B gene promoter polymorphisms and adolescent idiopathic scoliosis*. Spine (Phila Pa 1976), 2011. **36**(2): p. E123-30.
114. Liu, Z., et al., *Lack of association between the promoter polymorphisms of MMP-3 and IL-6 genes and adolescent idiopathic scoliosis: a case-control study in a Chinese Han population*. Spine (Phila Pa 1976), 2010. **35**(18): p. 1701-5.
115. Aulisa, L., et al., *Association between IL-6 and MMP-3 gene polymorphisms and adolescent idiopathic scoliosis: a case-control study*. Spine (Phila Pa 1976), 2007. **32**(24): p. 2700-2.

116. Crean, J.K., et al., *Matrix metalloproteinases in the human intervertebral disc: role in disc degeneration and scoliosis*. Spine (Phila Pa 1976), 1997. **22**(24): p. 2877-84.
117. Grivas, T.B., et al., *Expression of matrix metalloproteinase-1 (MMP-1) in Wistar rat's intervertebral disc after experimentally induced scoliotic deformity*. Scoliosis, 2011. **6**(1): p. 9.
118. Haglund, L., J. Ouellet, and P. Roughley, *Variation in chondroadherin abundance and fragmentation in the human scoliotic disc*. Spine (Phila Pa 1976), 2009. **34**(14): p. 1513-8.
119. Kang, J.A., et al., *Osteopontin regulates actin cytoskeleton and contributes to cell proliferation in primary erythroblasts*. J Biol Chem, 2008. **283**(11): p. 6997-7006.
120. Nagao, M., et al., *Sympathetic control of bone mass regulated by osteopontin*. Proc Natl Acad Sci U S A, 2011. **108**(43): p. 17767-72.
121. Chang, I.C., et al., *Increased serum osteopontin is a risk factor for osteoporosis in menopausal women*. Osteoporos Int, 2010. **21**(8): p. 1401-9.
122. Sakai, K., et al., *Microarray hybridization with fractionated cDNA: enhanced identification of differentially expressed genes*. Anal Biochem, 2000. **287**(1): p. 32-7.
123. Tusher, V.G., R. Tibshirani, and G. Chu, *Significance analysis of microarrays applied to the ionizing radiation response*. Proc Natl Acad Sci U S A, 2001. **98**(9): p. 5116-21.
124. Pavlidis, P., *Using ANOVA for gene selection from microarray studies of the nervous system*. Methods, 2003. **31**(4): p. 282-9.
125. Pavlidis, P. and W.S. Noble, *Analysis of strain and regional variation in gene expression in mouse brain*. Genome Biol, 2001. **2**(10): p. RESEARCH0042.
126. Hochberg, Y. and Y. Benjamini, *More powerful procedures for multiple significance testing*. Stat Med, 1990. **9**(7): p. 811-8.
127. Kane, R., C. Godson, and C. O'Brien, *Chordin-like 1, a bone morphogenetic protein-4 antagonist, is upregulated by hypoxia in human retinal pericytes and plays a role in regulating angiogenesis*. Mol Vis, 2008. **14**: p. 1138-48.

128. Branam, A.M., et al., *Zebrafish chordin-like and chordin are functionally redundant in regulating patterning of the dorsoventral axis*. Dev Biol, 2010. **341**(2): p. 444-58.
129. Videtic, A., et al., *Association study of seven polymorphisms in four serotonin receptor genes on suicide victims*. Am J Med Genet B Neuropsychiatr Genet, 2006. **141B**(6): p. 669-72.
130. Anguelova, M., C. Benkelfat, and G. Turecki, *A systematic review of association studies investigating genes coding for serotonin receptors and the serotonin transporter: II. Suicidal behavior*. Mol Psychiatry, 2003. **8**(7): p. 646-53.
131. Arango, V., et al., *Genetics of the serotonergic system in suicidal behavior*. J Psychiatr Res, 2003. **37**(5): p. 375-86.
132. Turecki, G., et al., *Suicide and serotonin: study of variation at seven serotonin receptor genes in suicide completers*. Am J Med Genet B Neuropsychiatr Genet, 2003. **118B**(1): p. 36-40.
133. Peters, H., et al., *Pax1 and Pax9 synergistically regulate vertebral column development*. Development, 1999. **126**(23): p. 5399-408.
134. Dastych, M. and O. Vlach, *Zinc status in patients with idiopathic scoliosis*. Spine (Phila Pa 1976), 1990. **15**(2): p. 65-6.

## **CHAPITRE 3 : ARTICLE II**

## **Introduction sur le deuxième article : Gene expression analysis of monozygotic twins discordant for idiopathic scoliosis**

La scoliose idiopathique est une maladie génétique complexe, dont la variabilité phénotypique ajoute à la difficulté d'en définir les bases génétiques. Nous avons tenté via une étude pilote d'utiliser des jumeaux monozygotes discordants pour la scoliose idiopathique pour identifier les gènes associés à cette maladie. Puisque ces jumeaux sont identiques, seuls les gènes dont l'expression est altérée pourraient jouer un rôle dans la SIA. Nous avons choisi les cellules sanguines (PBMCs) puisque celles-ci présentent également un défaut de signalisation et représentent aussi une méthode minimalement invasive.

Ici, nous avons comparé les profils d'expression du génome entier en utilisant la technique de biopuces à ADN sur une paire de jumelles monozygotes dont les phénotypes ont été bien caractérisés au niveau clinique, biochimique, et qui présentent des endophénotypes distincts définis par le dysfonctionnement de la signalisation systémique. Cette approche a facilité l'identification de nouveaux gènes candidats de prédisposition et aussi de gènes secondaires liés possiblement à la progression de la courbure.

### **La contribution des auteurs :**

Qing Yuan a réalisé l'étude expérimentale, a contribué à la préparation du manuscrit et aux analyses fonctionnelles des gènes. Kristen Fay Gorman a évalué et interprété les données et a aidé à la préparation du manuscrit. Cédric Julien a participé aux analyses fonctionnelles des gènes et à la validation des gènes candidats. Anita Franco a aidé dans les expériences et a isolé les PBMCs des patients scoliotiques et des sujets sains. Saadallah Bouhanik a aidé dans la préparation des échantillons et à réaliser les analyses dans le programme de Flexarray.

Alain Moreau a conçu l'étude, a supervisé l'ensemble du projet, a évalué et interprété les données, et aidé à la préparation du manuscrit.

Tous les auteurs ont lu et approuvé le manuscrit final.

# **Gene expression analysis of monozygotic twins discordant for idiopathic scoliosis**

Qing Yuan<sup>1</sup>, Kristen Fay Gorman<sup>1</sup>, Saadallah Bouhnik<sup>1</sup>, Cédric Julien<sup>1</sup>, Anita Franco<sup>1</sup>, Alain Moreau<sup>1,2</sup> 

<sup>1</sup>*Viscogliosi Laboratory in Molecular Genetics of Musculoskeletal Diseases, Sainte-Justine University Hospital Research Center; Department of Biochemistry, Faculty of Medicine, Université de Montréal, Québec, Canada.*

<sup>2</sup>*Department of Stomatology, Faculty of Dentistry, Université de Montréal, Québec, Canada.*



Phone: 514-345-4931 ext: 3476

Fax: 514-345-4801

## **Abstract**

Idiopathic scoliosis is a multifactorial disease, which a genetic disorder is considered to be a critical factor of IS. Its phenotypic variability adds to the difficulties of defining the genetic bases. In this pilot study, we are using the model of monozygotic discordant twins in order to identify genes associated with this disease. Discordant monozygotic twins (MZ) are ideal for the study of genetic expression related to variability in complex traits like IS because of gene and environmental background sharing. Wherefore only the genes whose expression is altered may play a role in the disease of AIS. We describe the first application of genome-wide expression profile to IS using a pair of female monozygotic twins whose phenotypes have been well characterized at clinical and also biochemical level, according to the distinct endophenotypes defined by the associated systemic signaling dysfunction systems. We use Peripheral Blood Mononuclear Cells (PBMCs) as a minimum invasive method in this study. We perform genome-wide expression microarray technology in order to assess gene expression patterns that are associated with IS. This is the first gene expression study in IS relates genetic, epigenetic, and environmental influences. This approach has facilitated the identification of novel candidate predisposing genes and also possibly related to the secondary progression of the curvature in IS patients.

## **Introduction**

Idiopathic scoliosis (IS) is a complex genetic disease affecting on average 4% of the global pediatric population. It is manifest as an abnormal structural curvature of the spine that involves changes in the frontal, sagittal, and transverse planes. Because these changes occur during growth, IS is considered a complex three-dimensional deformity. A genetic basis is supported from twin studies and from analysis of inheritance in multigenerational families (reviewed in [1] [2]). Phenotypic variability for factors such as age of onset, curve morphology (*i.e.*, curve magnitude and curve shape), and the propensity for curve progression has complicated efforts to understand the genetic basis (*e.g.* [3]). There is evidence that IS is a polygenic disorder [4-6]. Furthermore, genetic heterogeneity among different populations is presumed from the limited consensus for loci identified in linkage analysis studies.

Discordant monozygotic twins (MZ) are ideal for the study of genetic expression related to variability in complex traits because of gene sharing, and a common family and background environment (reviewed in [7]). Reports regarding concordance among MZ twins with IS are varied, depending on sample sizes and the nature of the study. Here concordance refers specifically to whether both twins are affected, in that curvature is approached as a binary trait: presence or absence of curvature. Those studies that report a high pairwise concordance (0.76-0.92) have used small sets of twins [8, 9], compared to population-based studies that report lower concordance rates (0.11-0.13 for pairwise; 0.17-0.25 for proband wise) [10, 11]. Generally, low concordance rates suggest that there are substantial environmental and/or epigenetic effects. A recent study of concordance among monozygotic and dizygotic twins from the Swedish twin registry estimates that additive genetic effects explain 38% of the observed phenotypic variance for IS, leaving 62% to unique environmental effects [11]. Furthermore,

among affected MZ twins, there are varying reports regarding concordance for curve severity and/or curve morphology [3, 8, 9].

The etiology of IS has been defined as multi-factorial, in that many parts of the body have been incriminated in the etiopathogenesis. Examples of systems hypothesized to be involved in the etiology and/or pathogenesis are: the nervous and vestibular systems [12-15] (*e.g.* balance and gait studies) [16-20]; bone and connective tissues [21-26]; muscle [27-32]; platelets [33, 34]; the endocrine system (*i.e.* melatonin, estrogen, testosterone, growth hormone, insulin growth factor) [35] [[36-39]]; and molecular signaling (reviewed in [40]). As a consequence of this, many different tissues have been considered as affected. In order to reveal IS relevant genes, not much expression work has been done because there is a dilemma about which tissue would be most appropriate. We previously identified a systemic Gi protein signaling pathway dysfunction in IS patients using cellular dielectric spectroscopy (CDS) to measure the functional activity of G proteins in multiple cell types [41]. Importantly, because the signaling dysfunction is systemic, we avoid issues related to which tissue type is most suitable for expression studies in IS.

Gene expression levels are a result of genetic (transcriptional and translational), epigenetic, and environmental influences. Here we describe the first application of genome-wide expression analysis to IS using a pair of female MZ twins whose phenotypes were characterized clinically and also biochemically, according to the distinct endophenotypes defined by the IS associated systemic signaling dysfunction [42]. This approach and the stringent analysis employed has facilitated the identification of important novel candidates for IS.

## **Materials and Methods**

### *Patients and phenotypes*

The female monozygotic twins of this study were born in May, 1995, following a full-term pregnancy. The Institutional Review Board of The Sainte-Justine Hospital approved this study. Parents gave written informed consent, and minors gave their assent. The patients were examined by 1 of the 6 orthopedic surgeons (H.L., B.P., C.-H.R., G.G., J.O., and S.P.) at Sainte-Justine Hospital. Each twin was subject to examinations with radiographs at multiple time points (Table 1), and blood for the microarray was drawn at the age of 15.6 and immediately processed. In addition to clinical evaluations, both twins were classified according to endophenotypes (*i.e.*, functional subgroups) described by Moreau, et al [42], using cellular dielectric spectroscopy (CDS) (MDS Analytical Technologies, CA). This CDS functional test classifies IS patients based on a *Gi* protein signaling dysfunction, relative to cells from non-scoliosis individuals [41].

### *Isolation of Human Peripheral Blood Mononuclear Cells*

We isolated Peripheral Blood Mononuclear Cells (PBMCs) from blood: blood was drawn from the coccygeal vein and collected into two 6 ml EDTA-treated tubes (BD Biosciences, NJ, USA) and immediately processed. Approximately 10 ml of blood was collected between the two tubes and subjected to Ficoll-Paque Plus gradient (1.077 g/ml) centrifugation to isolate the PBMCs. Briefly, the blood was transferred to a 50ml centrifuge culture tube and an equal volume (1:1 ratio) of balanced salt buffer was added. The blood solution was then slowly and carefully deposited on top of the Ficoll-Paque solution, respecting a 1:1.5 ratio of Ficoll to blood. These tubes were then centrifuged at 400×g for 30 min at room temperature, allowing the separation of phases. Approximately, 2-3ml of the cloudy lymphocyte layer was collected, and washed in

20ml of RPMI medium complemented with 10% FBS and 1% antibiotic-antimicotic, and centrifuged at 288×g for 7 min. The supernatant was subsequently removed, and the pellet was gently resuspended in 1ml of Trizol reagent (Invitrogen, CA, USA). After gently disrupting the pellet by vortexing, this homogenate was placed at -80°C until RNA isolation.

#### *Extraction of RNA*

Total RNA from PBMCs was extracted using Trizol (Invitrogen, CA, USA) according to manufacturer's instructions. Briefly, cells were lysed in 1ml of Trizol and homogenized with pipette and transferred to clean RNA-free tube. Samples were layered with 100 µl of BCP (1-bromo-3-chloropropane) (MRC, OH, USA) inverted 15 times, and incubated 3 min at room temperature. Lysates were centrifuged at 4°C for 15 min at 12,000 g and the aqueous phase transferred to a clean tube. RNA was precipitated with equal volumes of isopropanol for 10 min at room temperature. Precipitates were spun for 10 min at 4°C at 12,000 g. Pellets were washed with 70% Ethanol, and dried at room temperature. RNA was resuspended in 50µl RNase free water (Qiagen, Canada). An aliquot of total RNA diluted in RNase free water was set aside to estimate RNA quantity and integrity, and the remaining sample was stored at -80°C until gene expression analysis. Total RNA quantity and integrity was assessed using the spectrophotometry and samples maintained a 28S/18S rRNA ratio of 1.5 or greater. RNA integrity was confirmed by RNA Quality Control provided by Genome Québec, according to specified instructions.

#### *Microarray and Analysis of Data*

Experiments were performed in triplicate using the Affymetrix, GeneChip® Human Exon 1.0 ST array, at the Centre d'Innovation de Génome Québec in Montréal. All protocols were conducted as described in the Affymetrix GeneChip Expression Analysis technical manuals. The raw data were normalized by the RMA method (Robust Multi-array Analysis). The log2

(treatment/control) ratio for each probe set was calculated and *p*-values from one-way ANOVA followed by false recovery rate (FDR)-correction were calculated according to Benjamini and Hochberg [43]. Probes were considered for further evaluation if their FDR-corrected *p*-value for the ANOVA was significant at  $p \leq 0.05$ . For candidate gene selection, probe sets were filtered by expression change being at least  $\pm 2$  fold. We used Wright and Simons (EB) method [44] in the FlexArray program [45] to identify genes whose expression differed significantly between the twins.

## Results

To detect molecular pathways related to the etiopathogenesis of IS we conducted microarray analysis in a pair of monozygotic twins discordant for their diagnosis and phenotype. We used the Affymetrix, GeneChip® Human Exon 1.0 ST array with 54,000 probe sets analyze the expression of over 47,000 transcripts and variants from over 30,000 well-characterized human genes. At an expected false discovery rate of 5%, our analysis revealed 26 genes that were significantly differentially expressed by at least 2-fold (Figure 1). Of the 18 overexpressed genes, there are 3 clusters based on location (defined by two or more genes in the same locus), one of which is a putative locus for adolescent idiopathic scoliosis (AIS) [46]. There are 2 location-based clusters among the 8 underexpressed genes. Results regarding the locations of genes that are differentially expressed between the twins are summarized in Table 2 and discussed below.

We used functional annotation tools (DAVID Bioinformatics Resources) [47] to analyze the genes according to their functions. This analysis produced seven gene ontology clusters

(summarized in Table 3). Below we discuss how some of the gene functions relate to hypotheses regarding the etiopathogenesis for IS.

## Discussion

With regard to IS, there are two important levels of discordance to be considered. The first is in regard to the presence of curvature. In prior twin studies, discordance for the presence of curvature is rare in small, confirmed datasets [9, 48, 49]. The second level of discordance to be considered for IS twin studies is with regard to the pathogenesis of curvature. Prior studies demonstrate that discordance among monozygotic twins for curve morphology and curve magnitude is common [3, 50]. Here we present a gene expression analysis study for a set of monozygotic twins that are clinically discordant for IS (Figure 2). Although from a clinical perspective the twins in this study are discordant, because one twin has AIS and the other does not, but considering that they both present *some* spinal curvature, one twin has a 34-43 degrees Cobb angle with a rTITL curve type, other twin has a 14 degree Cobb angle with a LL curve type, these make it difficult to be certain that they do not share a genetic predisposition (see Table 1). It is more conceivable that changes detected between those twins could rather reflect their genetics predisposition to spinal deformity progression. Certainly these twins are discordant for curve pathogenesis as well as the morphology and magnitude of curvature. However, considering that little is understood about predisposing factors, it cannot be ruled out that the non-IS twin had inherited predisposing factors but did not develop the syndrome due to environmental, epigenetic, and/or behavioral differences (differences in their respective lifestyle). Furthermore, according to our cell-based screening test using cellular dielectric spectroscopy (CDS), both twins belong to the same biochemical functional subgroup: their specific Gi-coupled receptor signaling defect endophenotype is shared. Considering that we hypothesize that this signaling defect could be one of the predisposing factors for IS, this evidence further suggests that such factors are shared

between the twins. Conversely, if the twins are indeed discordant for the genetic predisposition for IS, we still may not have detected their differential expression in this study. It is unknown whether genes for curve predisposition are expressed early in life only (*e.g* during embryogenesis, or early childhood), or constitutively throughout development. If the genes were expressed early in life only, then they would not be detected by the array presented here because samples were taken at age 15.6. Hence, we conservatively expected that the genes differentially expressed in this microarray relate most immediately to curve progression.

Considering that the pathogenesis of IS occurs during growth, a dynamic period with changing hormone levels and gene expression levels, it is important to keep in mind that the microarray results presented here represent a small window into the greater landscape of the disease. There is a strong relationship between curve progression and rapid growth occurring during puberty, with the increase in curve magnitude correlated to a rapid increase in spinal height [51-54]. In girls, the peak growth spurt is known to take place between 10 and 14 years in 95% of girls [54]. In this study, although the non-IS twin demonstrated a minor curve in her lumbar spine resulting from a leg length discrepancy, the curve did not progress in magnitude with growth (Table 1, Figure 1). In contrast, the twin with IS first presented a thoracolumbar curve (*i.e* double major curve) at the age of 12.5 with a Cobb angle of 24-30 degrees that showed a tendency to progress and so after 6 months she was prescribed a conservative treatment (SpineCor brace). At 11 months past curve onset, even with brace, the curvature continued to progress, so that at the age 15.6 the curve developed to 34-43 Cobb angle. At the time when blood was collected for the microarray, both girls were at the age of 15.6; the scoliotic twin had a progressed curve (34-43 degree, double major curve), and the other twin still had a minor lumbar curve from her diagnosed limb inequality. Of the genes identified in our array, integrin beta 3 (*ITGB3*) is known to enhance IGF-I-stimulated signaling, a known growth factor [55]. IGF1 has a strong

relationship with estrogen at different ages in women in a breast cancer study [56]. Furthermore, ITGB3 is a receptor for matrix metalloproteinase-2, osteopontin, osteomodulin, and thrombospondin (another gene that is significantly upregulated in the IS twin) [57]. Matrix metalloproteinases regulate degradation and remodeling of the extracellular matrix (ECM), and are important candidate structure proteins genes in IS. Thus, it is possible that *ITGB3* is a relevant growth gene related to IS progression. In addition to *ITGB3* being up regulated, *KIAA1324-like* and *SPRR2B* [58] are also estrogen induced and up regulated in the IS twin. Considering that estrogen receptor signaling is affected by epigenetic regulation [59], discordant estrogen levels among these monozygotic twins might explain variation among some genes regulated by estrogen. Indeed, discordant estrogen regulation might explain why the IS twin started menarche in June 2008 while the non-IS twin started in January 2009. There were no other obvious growth genes differentially expressed in our results, possibly because at the age of 15.6 both girls have already past their growth peak, and so growth genes that might have been expressed differentially during the period would not be detected by our array.

Variation among monozygotic twins (MZ) demonstrates that environmental and/or epigenetic factors are important considerations for IS study. Epigenetic modifications can affect how a person's biology interacts with the environment either directly or indirectly. Epigenetic mechanisms such as methylation or acetylation are dynamic processes that are influenced by developmental stage, tissue type, environmental (and/or behavioral) factors, and stochastic events [60]. Because epigenetic modifications are dynamic and thus partially stable, there is the potential for epigenetic variation among MZ twin pairs [61]. Four genes in our study have been demonstrated to be susceptible to methylation at the gene promotor, or regulatory module: *HLA-DQBI* [62], *MSLN* [63], *TLR3* [64], *THBS1* [65]. Other genes likely to be involved in epigenetic regulation are: non-codingRNA (ENST00000391154) (reviewed in [66]), and

*HLA-DQAI* [67]. Among the differentially expressed epigenetic targets, *MSLN* and *THBS1* are of particular interest. Membrane anchored forms of *MSLN* play a role in cell adhesion. It is possible therefore that the observed down regulation of this gene in the IS twin is related to the curve pathology. *THBS1* is also involved in cell adhesion and binds fibrinogen, fibronectin, laminin, and type V collagen. However, although the IS twin removed her brace two days before blood was withdrawn, considering that *THBS1* expression is known to increase after mechanical strain [68], it difficult to rule-out that the observed up regulation in the IS twin is not secondary to long term bracing.

Chromosomal clustering of up or down regulated genes may be indicative of shared transcription factors or copy number variants (CNVs) that are important for IS pathogenesis. There are three clusters of two or more genes that are up regulated, and two chromosomal clusters among the down regulated genes. We found that among the down regulated genes, *HLA-DQB1* and *HLADQA1* and *DDR1* which are located in the MHC locus, are typically co-acetylated [69]. Histone acetylation is known to facilitate gene transcription. The receptor tyrosine kinase *DDR1* gene is an interesting candidate because its expression can be reduced significantly by excess OPN [70]. In the case of the twins studied here, the endophenotype called Group 3 is generally associated with an elevated [OPN]. That elevated OPN levels can affect epigenetic profiles to down regulate gene expression has been established *in vitro* [71]. Discoidin domain receptors (DDRs) are a novel class of receptor tyrosine kinases that bind several collagens and stimulate matrix metalloproteinase (MMP) production, and are expressed in skeletal bones [72, 73]. DDR1 activation results in increased junctional stability: the activation of DDR1 inhibits the  $\beta 1$  integrin signal that normally activates Cdc42, and this inactivation of Cdc42 promotes E-cadherin-mediated cell-cell adhesion [74].

Genetic differences between monozygotic twins can arise *de novo* through somatic CNVs. There is evidence that somatic mosaicism for pathogenic mutations affecting some disease genes is a common occurrence among MZ twins [75]. Among the genes significantly differentially expressed in this study, two that are down regulated are located at 5q13.2, a region associated with CNVs. *GTF2H2* encodes a RNA polymerase II transcription factor IIH and is a part of a 500kb inverted duplication that is prone to deletions and rearrangements (NCBI gene ID 2966). This genetically complex region comprises the SMA (Spinal Muscular Atrophy) region, which is implicated in numerous pediatric abnormalities (OMIM 253550, 253400, 253300, 271150, 600354, 603011). Also, three genes up regulated in the IS twin (*CCDC144A*, 5x; *FAM106A*, 3.1x; *CCDC144NL*, 2.2x) are located at 17p11.2, a known hotspot for CNVs that are associated with numerous disorders. Considering that these genes are thought to be important for gene regulation, their CNVs may have a strong impact. Interestingly, this is a putative IS locus (MIM607354), identified by linkage analysis [46].

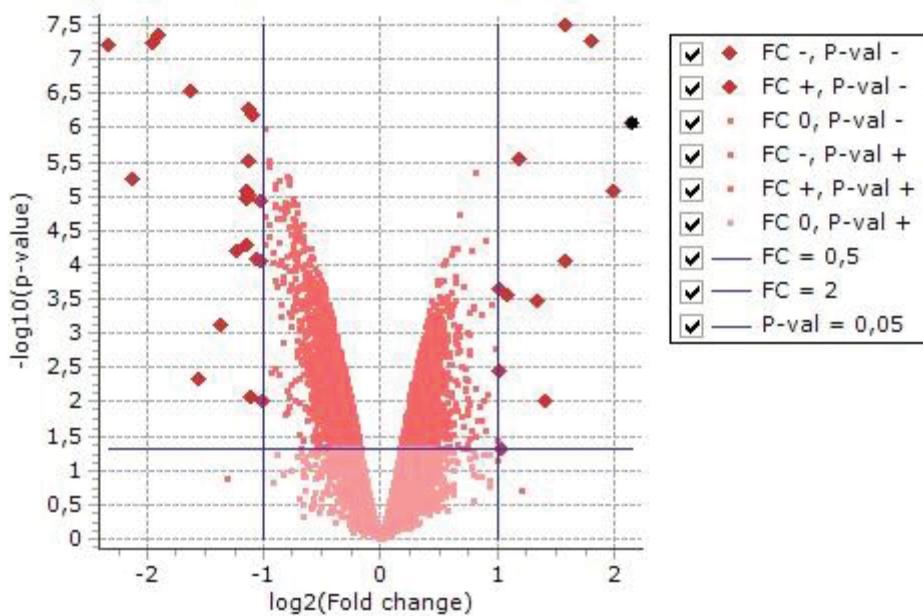
Finally, differentially expressed genes known to be involved in osteopontin (OPN) signaling are of special interest. Chemokine (C-motif) ligand 2 (*XCL2*) acts through a subset of G-protein receptors to play a fundamental role in the development, homeostasis, and function of the immune system. *XCL2* from macrophages specifically bind Wag31 as a part of the primary response to *M. tuberculosis* infection [76]. Considering that OPN gene expression is also prompted by mycobacterial infection of human macrophages [77], it is likely that these two genes are functionally linked in their immune and/or inflammatory roles.

## Conclusions

In contrast to prior genetic studies for IS (*i.e.* linkage analysis or candidate gene), gene expression technology provides us with information on the activities of all genes simultaneously therefore allowing us to apply the field of functional genomics to IS. Such an approach to a complex phenotype such as IS will allow for the discovery of novel molecular pathways not revealed by sequence analysis alone. Many genes have been examined as candidates for IS [78]. These genes are either based on hypotheses about systems thought to be involved in IS, based on clinical or experimental results, or are genes identified in intervals from linkage analysis studies. None of the candidates for IS has been confirmed. The goal of this genome-wide gene expression study in monozygotic discordant scoliosis twins was to look for candidate genes with functional significance using a non-biased approach. We identify 26 genes that are significantly differentially regulated, some with important functional considerations, and three within a putative IS locus.

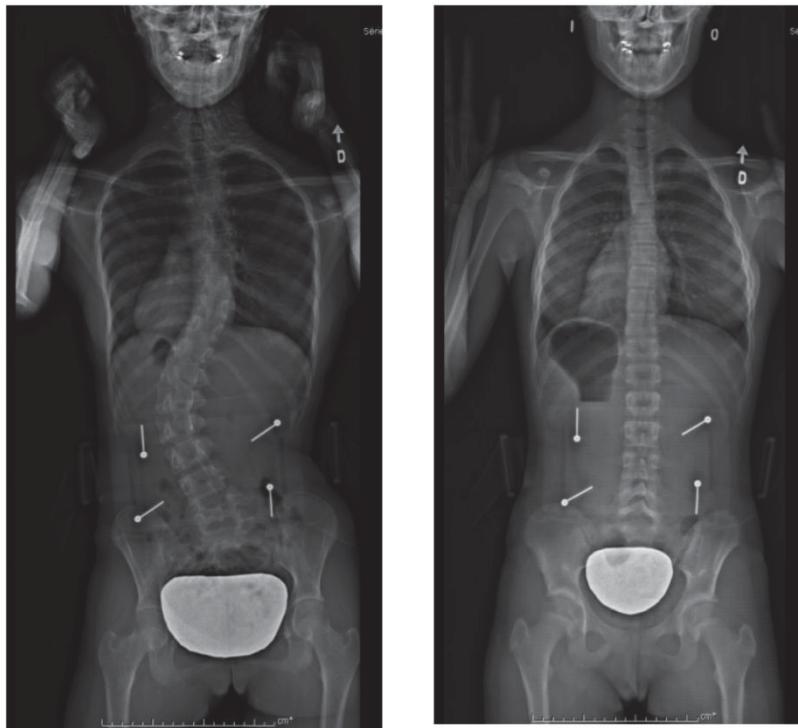
Considering the variability for the curve magnitude and morphology observed within families, it has been hypothesized that variations in the environment and /or epigenetic phenomenon may affect the gene expression in IS. [3, 79] Here we present a method for the identification of candidate genes functionally related to curve progression. A limitation with this approach is that we cannot rule-out that differential gene expression in the IS twin is not a consequence of curve progression rather than the cause. However, those genes identified are important to consider none-the-less because they suggest important molecular pathways involved in the evolution of the IS phenotype. Prior genetic studies of IS using Genome-wide association study (GWAS) or linkage analysis have the capacity to identify candidate disposition genes, but are limited in their capacity to discover genes that explain phenotypic variation.

## Tables and Figures



**Figure 1: Volcano plot of p-values EB (Wright & Simon)**

Plot shows the 26 significantly differentially expressed genes that are 2 fold under or over expressed as large diamonds. Plot created from analysis in FlexArray [45].



**Figure 2: Radiograph of monozygotic twins discordant for IS**

Twins shown here at age 15.6. The twin on the left, random number assignment 340, has a progressive double major curve (here with Cobb angles of thoracic 34° and lumbar 43°) and has been diagnosed with IS. The twin on the right, random number assignment 3230, has a non-progressive slight lumbar curve (here with a Cobb angle of 14°) from a diagnosed limb length inequality.

**Table 1. Patient phenotypic information**

| Twin ID | Age  | Diagnosis       | Cobb angle <sup>+</sup> | Curve Type     |
|---------|------|-----------------|-------------------------|----------------|
| 340     | 12.5 | AIS             | 24-30                   | rTITL          |
| 340     | 13.0 | AIS             | 23-30*                  | rTITL          |
| 340     | 13.4 | AIS             | 29-34*                  | rTITL          |
| 340     | 13.8 | AIS             | 24-39*                  | rTITL          |
| 340     | 14.2 | AIS             | 30-38*                  | rTITL          |
| 340     | 14.7 | AIS             | 28-39*                  | rTITL          |
| 340     | 15.2 | AIS             | 32-41*                  | rTITL          |
| 340     | 15.6 | AIS             | 34-43                   | rTITL          |
| <hr/>   |      |                 |                         |                |
| 3230    | 14.7 | Limb inequality | 12                      | IL             |
| 3230    | 15.2 | Limb inequality | No radiograph           | Not Applicable |
| 3230    | 15.6 | Limb inequality | 14                      | IL             |

Twin 340 was examined by an orthopaedic doctor eight times between the ages of 12 and 16 years old.

Diagnostic values from T0-T38 (38 months past diagnosis) are shown. AIS= Adolescent idiopathic scoliosis; rTITL= right thoracic, left thoracic, lumbar curve type; <sup>+</sup>numbers represent the right thoracic and left thoracic curves, respectively; \*indicates values made from radiographs taken while the patient was in a Spinecor dynamic brace.

Twin 3230 was examined by the same orthopaedic doctor as Twin 340, but did not have a spinal curvature until the age of 14.7, when she had a left lumbar (IL) curve, associated with a leg length inequality. The curve did not progress and her file was closed after three visits, 11 months after her first visit.

**Table 2. Significant genes differentially expressed between MZ twins with and without IS**

\* Significance was defined as p-value < 0.05

| Location                    | Gene Symbol      | Title   | Fold change |
|-----------------------------|------------------|---|-------------|
| <b>OVER-EXPRESSED GENES</b> |                  |   |             |
| 17p11.2                     | <i>CCDC144A</i>  | Coiled-coil domain containing 144A                        | 5           |
| 1p13.3                      | <i>GSTM4</i>     | Glutathione S-transferase mu 4                            | 3.8         |
| 1q24.2                      | <i>XCL2</i>      | Chemokine (C motif) ligand 2                              | 3.7         |
| 17p11.2                     | <i>FAM106A</i>   | Family with sequence similarity 106, member A             | 3.1         |
| 12p13.31                    | <i>FAM90A1</i>   | Family with sequence similarity 90, member A1             | 2.6         |
| 19q13.42                    | <i>KIR3DX1</i>   | Killer immunoglobulin-like receptor                       | 2.3         |
| 2p12                        | <i>IGKC</i>      | Immunoglobulin kappa constant                             | 2.2         |
| 10p12.1                     | <i>ENKUR</i>     | Enkurin, TRPC channel interacting protein                 | 2.2         |
| 4q27                        | <i>PDE5A</i>     | Phosphodiesterase 5A                                      | 2.2         |
| 17q21                       | <i>ITGB3</i>     | Integrin, beta 3  | 2.2         |
| 7q21.12                     | <i>KIAA1324L</i> | KIAA1324-like protein                                     | 2.2         |
| 17p11.2                     | <i>CCDC144NL</i> | Coiled-coil domain containing 144 family, N-terminal like | 2.2         |
| 2q11.1                      | <i>ANKRD20B</i>  | Ankyrin repeat domain 20B                                 | 2.2         |
| 15q15                       | <i>THBS1</i>     | Thrombospondin 1  | 2.1         |
| 12p13.31                    | <i>KLRF1</i>     | Killer cell lectin-like receptor subfamily F              | 2.1         |

---

|                              |                     |  |      |
|------------------------------|---------------------|--|------|
| 19q13.4                      | <i>KIR2DL2</i>      | Killer cell immunoglobulin-like receptor         | 2.0  |
| 4q35                         | <i>TLR3</i>         | Toll-like receptor 3                             | 2.0  |
| 1q21.3                       | <i>SPRR2B</i>       | Small proline-rich protein 2B                    | 2.0  |
| <b>UNDER-EXPRESSED GENES</b> |                     |  |      |
| 1q21.2                       | <i>snRNA</i>        | U1 spliceosomal RNA                              | -4.4 |
| 6p21.3                       | <i>HLA-DQBI</i>     | Histocompatibility complex, class II, DQ beta 1  | -4.0 |
| 5q13.2                       | <i>RP11-589F5.3</i> | Pseudogene                                       | -3.5 |
| 12                           | <i>TAS2R45</i>      | Taste receptor, type 2, member 45                | -3.0 |
| 6p21.3                       | <i>DDRI</i>         | Discoidin domain receptor tyrosine kinase 1      | -2.7 |
| 6p21.1                       | <i>HLA-DQA1</i>     | Histocompatibility complex, class II, DQ alpha 1 | -2.5 |
| 6p13.3                       | <i>MSLN</i>         | Mesothelin                                       | -2.1 |
| 5q13.2                       | <i>GTF2H2</i>       | General transcription factor IIH, polypeptide 2  | -2.0 |

---

Fold change reflects gene expression in the twin with IS relative to her sister without IS.

The table shows significant genes differentially expressed between monozygotic twins with or without idiopathic scoliosis. With 2 large groups which are over-expressed and under-expressed genes, each shows by its location on the whole genome, the gene symbol, the full tile and the fold change.

**Table 3. Summary of functional annotation of candidate genes in twin with IS**

\* Significance was defined as p-value < 0.05

| Locatio                           | GENE                       | Title   | Fold  | P-value |
|-----------------------------------|----------------------------|---|-------|---------|
| <b>Response for stimulus</b>      |                            |   |       | 0.002   |
| chr17                             | <i>ITGB3</i>               | Integrin, beta 3                                | 2.2   |         |
| chr19                             | <i>KIR2DL2</i>             | Killer cell immunoglobulin-like receptor        | 2.0   |         |
| chr15                             | <i>THBS1</i>               | Thrombospondin 1                                | 2.1   |         |
| chr4                              | <i>TLR3</i>                | Toll-like receptor 3                            | 2.0   |         |
| chr5                              | <i>GTF2H2</i>              | General transcription factor IIH, polypeptide 2 | -2.0  |         |
| chr6                              | <i>DDR1</i>                | Discoidin domain receptor tyrosine kinase 1     | -2.7  |         |
| chr1                              | <i>XCL2</i>                | Chemokine (C motif) ligand 2                    | 3.7   |         |
| chr4                              | <i>PDE5A</i>               | Phosphodiesterase 5A                            | 2.2   |         |
| chr6                              | <i>HLA-DQ<sub>A1</sub></i> | Histocompatibility complex, class II, DQ        | -2.5  |         |
| --                                | <i>TAS2R45</i>             | Taste receptor, type 2, member 45               | -2.98 |         |
| <b>Cell communication</b>         |                            |   |       | 0.0025  |
| chr1                              | <i>XCL2</i>                | Chemokine (C motif) ligand 2                    | 3.7   |         |
| chr6                              | <i>DDR1</i>                | Discoidin domain receptor tyrosine kinase 1     | -2.7  |         |
| chr17                             | <i>ITGB3</i>               | Integrin, beta 3                                | 2.2   |         |
| chr19                             | <i>KIR3DX1</i>             | Killer immunoglobulin-like receptor             | 2.33  |         |
| chr19                             | <i>KIR2DL2</i>             | Killer cell immunoglobulin-like receptor        | 2.0   |         |
| chr16                             | <i>MSLN</i>                | Mesothelin                                      | -2.13 |         |
| <b>Immune response</b>            |                            |   |       | 0.0033  |
| chr1                              | <i>XCL2</i>                | Chemokine (C motif) ligand 2                    | 3.7   |         |
| chr15                             | <i>THBS1</i>               | Thrombospondin 1                                | 2.1   |         |
| chr19                             | <i>KIR2DL2</i>             | Killer cell immunoglobulin-like receptor        | 2.0   |         |
| chr4                              | <i>TLR3</i>                | Toll-like receptor 3                            | 2.0   |         |
| chr6                              | <i>HLA-DQ<sub>A1</sub></i> | Histocompatibility complex, class II, DQ        | -2.5  |         |
| <b>Signal transducer activity</b> |                            |   |       | 0.0036  |
| chr7                              | <i>KIAA1324L</i>           | KIAA1324-like protein                           | 2.17  |         |
| chr6                              | <i>DDR1</i>                | Discoidin domain receptor tyrosine kinase 1     | -2.7  |         |
| chr17                             | <i>ITGB3</i>               | Integrin, beta 3                                | 2.2   |         |
| chr19                             | <i>KIR2DL2</i>             | Killer cell immunoglobulin-like receptor        | 2.0   |         |
| chr6                              | <i>HLA-DQ<sub>A1</sub></i> | Histocompatibility complex, class II, DQ        | -2.5  |         |
| --                                | <i>TAS2R45</i>             | Taste receptor, type 2, member 45               | -2.98 |         |

|                                  |                  |  |       |
|----------------------------------|------------------|--|-------|
| chr4                             | <i>TLR3</i>      | Toll-like receptor 3                         | 2.0   |
| chr15                            | <i>THBS1</i>     | Thrombospondin 1                             | 2.1   |
| <b>Ligand-mediated signaling</b> |                  |  | 0.005 |
| chr1                             | <i>XCL2</i>      | Chemotactic activity for lymphocytes         | 3.7   |
| chr19                            | <i>KIR3DX1</i>   | Killer immunoglobulin-like receptor          | 2.33  |
| chr19                            | <i>KIR2DL2</i>   | Inhibits the activity of NK cells thus       | 2.0   |
| chr16                            | <i>MSLN</i>      | Mesothelin                                   | -2.13 |
| <b>Cell adhesion</b>             |                  |  | 0.027 |
| chr6                             | <i>DDR1</i>      | Discoidin domain receptor tyrosine kinase 1  | -2.7  |
| chr17                            | <i>ITGB3</i>     | Integrin, beta 3                             | 2.2   |
| chr16                            | <i>MSLN</i>      | Mesothelin                                   | -2.13 |
| chr15                            | <i>THBS1</i>     | Thrombospondin 1                             | 2.1   |
| <b>Receptor activity</b>         |                  |  | 0.001 |
| chr7                             | <i>KIAA1324L</i> | KIAA1324-like protein                        | 2.17  |
| chr6                             | <i>DDR1</i>      | Discoidin domain receptor tyrosine kinase 1  | -2.7  |
| chr17                            | <i>ITGB3</i>     | Integrin, beta 3                             | 2.2   |
| chr19                            | <i>KIR2DL2</i>   | Killer cell immunoglobulin-like receptor     | 2.0   |
| chr12                            | <i>KLRF1</i>     | Killer cell lectin-like receptor subfamily F | 2.09  |
| chr6                             | <i>HLA-DQ41</i>  | Histocompatibility complex, class II, DQ     | -2.5  |
| --                               | <i>TAS2R45</i>   | Taste receptor, type 2, member 45            | -2.98 |
| chr4                             | <i>TLR3</i>      | Toll-like receptor 3                         | 2.0   |

Functional annotation of the differentially expressed genes by DAVID Bioinformatics shows 7 distinct functional clusters.

## References

1. Miller, N.H., *Genetics of familial idiopathic scoliosis*. Clin Orthop Relat Res, 2007. **462**: p. 6-10.
2. Wise, C.A., et al., *Understanding genetic factors in idiopathic scoliosis, a complex disease of childhood*. Curr Genomics, 2008. **9**(1): p. 51-9.
3. van Rhijn, L.W., et al., *Curve characteristics in monozygotic twins with adolescent idiopathic scoliosis: 3 new twin pairs and a review of the literature*. Acta Orthop Scand, 2001. **72**(6): p. 621-5.
4. Filho NA, T.M., *Genetic studies in scoliosis*. J Bone Joint Surg [Am], 1971. **53**: p. 199.
5. Riseborough, E.J. and R. Wynne-Davies, *A genetic survey of idiopathic scoliosis in Boston, Massachusetts*. J Bone Joint Surg Am, 1973. **55**(5): p. 974-82.
6. Ward, K., et al., *Polygenic inheritance of adolescent idiopathic scoliosis: a study of extended families in Utah*. Am J Med Genet A, 2010. **152A**(5): p. 1178-88.
7. MacGregor, A.J., et al., *Twins. Novel uses to study complex traits and genetic diseases*. Trends Genet, 2000. **16**(3): p. 131-4.
8. Kesling, K.L. and K.A. Reinker, *Scoliosis in twins. A meta-analysis of the literature and report of six cases*. Spine (Phila Pa 1976), 1997. **22**(17): p. 2009-14; discussion 2015.
9. Inoue, M., et al., *Idiopathic scoliosis in twins studied by DNA fingerprinting: the incidence and type of scoliosis*. J Bone Joint Surg Br, 1998. **80**(2): p. 212-7.
10. Andersen, M.O., K. Thomsen, and K.O. Kyvik, *Adolescent idiopathic scoliosis in twins: a population-based survey*. Spine (Phila Pa 1976), 2007. **32**(8): p. 927-30.
11. Grauers, A., I. Rahman, and P. Gerdhem, *Heritability of scoliosis*. Eur Spine J, 2011.
12. Pincott, J.R., J.S. Davies, and L.F. Taffs, *Scoliosis caused by section of dorsal spinal nerve roots*. J Bone Joint Surg Br, 1984. **66**(1): p. 27-9.

13. Fisk, J.R. and W.H. Bunch, *Scoliosis in neuromuscular disease*. Orthop Clin North Am, 1979. **10**(4): p. 863-75.
14. Lonstein, J.E. and T.S. Renshaw, *Neuromuscular spine deformities*. Instr Course Lect, 1987. **36**: p. 285-304.
15. Manzoni, D. and F. Miele, *Vestibular mechanisms involved in idiopathic scoliosis*. Arch Ital Biol, 2002. **140**(1): p. 67-80.
16. Herman, R., et al., *Idiopathic scoliosis and the central nervous system: a motor control problem*. *The Harrington lecture, 1983. Scoliosis Research Society*. Spine (Phila Pa 1976), 1985. **10**(1): p. 1-14.
17. Guo, X., et al., *Balance control in adolescents with idiopathic scoliosis and disturbed somatosensory function*. Spine (Phila Pa 1976), 2006. **31**(14): p. E437-40.
18. Lao, M.L., et al., *Impaired dynamic balance control in adolescents with idiopathic scoliosis and abnormal somatosensory evoked potentials*. J Pediatr Orthop, 2008. **28**(8): p. 846-9.
19. Mahaudens, P., et al., *Gait in adolescent idiopathic scoliosis: kinematics and electromyographic analysis*. Eur Spine J, 2009. **18**(4): p. 512-21.
20. Domenech, J., et al., *Motor cortical hyperexcitability in idiopathic scoliosis: could focal dystonia be a subclinical etiological factor?* Eur Spine J, 2010. **19**(2): p. 223-30.
21. Zorab, P.A., et al., *Bone collagen turnover in idiopathic scoliosis estimated from total hydroxyproline excretion*. Arch Dis Child, 1971. **46**(250): p. 828-32.
22. Shapiro, J.R., et al., *Osteoporosis and familial idiopathic scoliosis: association with an abnormal alpha 2(I) collagen*. Connect Tissue Res, 1989. **21**(1-4): p. 117-23; discussion 124.
23. Miller, N.H., et al., *Genetic analysis of structural elastic fiber and collagen genes in familial adolescent idiopathic scoliosis*. J Orthop Res, 1996. **14**(6): p. 994-9.

24. Guo, X., et al., *Relative anterior spinal overgrowth in adolescent idiopathic scoliosis--result of disproportionate endochondral-membranous bone growth? Summary of an electronic focus group debate of the IBSE*. Eur Spine J, 2005. **14**(9): p. 862-73.
25. Man, G.C., et al., *Abnormal melatonin receptor 1B expression in osteoblasts from girls with adolescent idiopathic scoliosis*. J Pineal Res, 2011. **50**(4): p. 395-402.
26. Wang, W.J., et al., *Top theories for the etiopathogenesis of adolescent idiopathic scoliosis*. J Pediatr Orthop, 2011. **31**(1 Suppl): p. S14-27.
27. Fidler, M.W. and R.L. Jowett, *Muscle imbalance in the aetiology of scoliosis*. J Bone Joint Surg Br, 1976. **58**(2): p. 200-1.
28. Reuber, M., et al., *Trunk muscle myoelectric activities in idiopathic scoliosis*. Spine (Phila Pa 1976), 1983. **8**(5): p. 447-56.
29. Gonyea, W.J., et al., *An evaluation of muscle pathology in idiopathic scoliosis*. J Pediatr Orthop, 1985. **5**(3): p. 323-9.
30. Avikainen, V.J., A. Rezasoltani, and H.A. Kauhanen, *Asymmetry of paraspinal EMG-time characteristics in idiopathic scoliosis*. J Spinal Disord, 1999. **12**(1): p. 61-7.
31. Cheung, J., et al., *A preliminary study on electromyographic analysis of the paraspinal musculature in idiopathic scoliosis*. Eur Spine J, 2005. **14**(2): p. 130-7.
32. Martinez-Llorens, J., et al., *Muscle dysfunction and exercise limitation in adolescent idiopathic scoliosis*. Eur Respir J, 2010. **36**(2): p. 393-400.
33. Kahmann, R.D., et al., *Platelet function in adolescent idiopathic scoliosis*. Spine (Phila Pa 1976), 1992. **17**(2): p. 145-8.
34. Lowe, T., et al., *Platelet calmodulin levels in adolescent idiopathic scoliosis: do the levels correlate with curve progression and severity?* Spine (Phila Pa 1976), 2002. **27**(7): p. 768-75.
35. Misol, S., et al., *Growth hormone blood levels in patients with idiopathic scoliosis*. Clin Orthop Relat Res, 1971. **81**: p. 122-5.

36. Skogland, L.B. and J.A. Miller, *Growth related hormones in idiopathic scoliosis. An endocrine basis for accelerated growth.* Acta Orthop Scand, 1980. **51**(5): p. 779-80.
37. Fagan, A.B., D.J. Kennaway, and A.D. Sutherland, *Total 24-hour melatonin secretion in adolescent idiopathic scoliosis. A case-control study.* Spine (Phila Pa 1976), 1998. **23**(1): p. 41-6.
38. Brodner, W., et al., *Melatonin and adolescent idiopathic scoliosis.* J Bone Joint Surg Br, 2000. **82**(3): p. 399-403.
39. Barrios, C., et al., *Anthropometry and body composition profile of girls with nonsurgically treated adolescent idiopathic scoliosis.* Spine (Phila Pa 1976), 2011. **36**(18): p. 1470-7.
40. Lombardi, G., et al., *Biochemistry of adolescent idiopathic scoliosis.* Adv Clin Chem, 2011. **54**: p. 165-82.
41. Akoume, M.Y., et al., *Cell-based screening test for idiopathic scoliosis using cellular dielectric spectroscopy.* Spine (Phila Pa 1976), 2010. **35**(13): p. E601-8.
42. Moreau, A., et al., *Melatonin signaling dysfunction in adolescent idiopathic scoliosis.* Spine (Phila Pa 1976), 2004. **29**(16): p. 1772-81.
43. Hu, J.X., H. Zhao, and H.H. Zhou, *False Discovery Rate Control With Groups.* J Am Stat Assoc, 2010. **105**(491): p. 1215-1227.
44. Irizarry, R.A., et al., *Summaries of Affymetrix GeneChip probe level data.* Nucleic Acids Res, 2003. **31**(4): p. e15.
45. Blazejczyk M, M.M., Nadon R, *FlexArray: A statistical data analysis software for gene expression microarrays.* 2007.
46. Salehi, L.B., et al., *Assignment of a locus for autosomal dominant idiopathic scoliosis (IS) to human chromosome 17p11.* Hum Genet, 2002. **111**(4-5): p. 401-4.
47. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.

48. Weiss, H.R., *Idiopathic scoliosis: how much of a genetic disorder? Report of five pairs of monozygotic twins*. Dev Neurorehabil, 2007. **10**(1): p. 67-73.
49. Potoupnis, M.E., et al., *The role of exercising in a pair of female monozygotic (high-class athletes) twins discordant for adolescent idiopathic scoliosis*. Spine (Phila Pa 1976), 2008. **33**(17): p. E607-10.
50. Hermus, J.P., L.W. van Rhijn, and A. van Ooij, *Non-genetic expression of adolescent idiopathic scoliosis: a case report and review of the literature*. Eur Spine J, 2007. **16 Suppl 3**: p. 338-41.
51. Charles, Y.P., et al., *Progression risk of idiopathic juvenile scoliosis during pubertal growth*. Spine (Phila Pa 1976), 2006. **31**(17): p. 1933-42.
52. Loncar-Dusek, M., M. Pecina, and Z. Prebeg, *A longitudinal study of growth velocity and development of secondary gender characteristics versus onset of idiopathic scoliosis*. Clin Orthop Relat Res, 1991(270): p. 278-82.
53. Wever, D.J., K.A. Tonseth, and A.G. Veldhuizen, *Curve progression and spinal growth in brace treated idiopathic scoliosis*. Stud Health Technol Inform, 2002. **91**: p. 387-92.
54. Busscher, I., F.H. Wapstra, and A.G. Veldhuizen, *Predicting growth and curve progression in the individual patient with adolescent idiopathic scoliosis: design of a prospective longitudinal cohort study*. BMC Musculoskelet Disord, 2010. **11**: p. 93.
55. Xi, G., et al., *Expression of the human beta3 integrin subunit in mouse smooth muscle cells enhances IGF-I-stimulated signaling and proliferation*. J Cell Physiol, 2008. **214**(2): p. 306-15.
56. Yoshimoto, N., et al., *Genetic and environmental predictors, endogenous hormones and growth factors, and risk of estrogen receptor-positive breast cancer in Japanese women*. Cancer Sci, 2011. **102**(11): p. 2065-72.
57. Cui, R., et al., *Abrogation of the interaction between osteopontin and alphavbeta3 integrin reduces tumor growth of human lung cancer cells in mice*. Lung Cancer, 2007. **57**(3): p. 302-10.

58. Hong, S.H., et al., *Estrogen regulates the expression of the small proline-rich 2 gene family in the mouse uterus*. Mol Cells, 2004. **17**(3): p. 477-84.
59. Mann, M., V. Cortez, and R.K. Vadlamudi, *Epigenetics of Estrogen Receptor Signaling: Role in Hormonal Cancer Progression and Therapy*. Cancers (Basel), 2011. **3**(3): p. 1691-1707.
60. Petronis, A., et al., *Monozygotic twins exhibit numerous epigenetic differences: clues to twin discordance?* Schizophr Bull, 2003. **29**(1): p. 169-78.
61. Wong, A.H., Gottesman, II, and A. Petronis, *Phenotypic differences in genetically identical organisms: the epigenetic perspective*. Hum Mol Genet, 2005. **14 Spec No 1**: p. R11-8.
62. Majumder, P. and J.M. Boss, *DNA methylation dysregulates and silences the HLA-DQ locus by altering chromatin architecture*. Genes Immun, 2011. **12**(4): p. 291-9.
63. Nelson, H.H., et al., *The relationship between tumor MSLN methylation and serum mesothelin (SMRP) in mesothelioma*. Epigenetics, 2011. **6**(8): p. 1029-34.
64. Porras, A., et al., *Developmental and epigenetic regulation of the human TLR3 gene*. Mol Immunol, 2008. **46**(1): p. 27-36.
65. Yang, Q., et al., *Epigenetic alterations differ in phenotypically distinct human neuroblastoma cell lines*. BMC Cancer, 2010. **10**: p. 286.
66. Malecova, B. and K.V. Morris, *Transcriptional gene silencing through epigenetic changes mediated by non-coding RNAs*. Curr Opin Mol Ther, 2010. **12**(2): p. 214-22.
67. Sasaki, T., et al., *Preferential transmission of maternal allele with DQA1\*0301-DQB1\*0302 haplotype to affected offspring in families with type 1 diabetes*. J Hum Genet, 1999. **44**(5): p. 318-22.
68. Tran, M.D. and J.T. Neary, *Purinergic signaling induces thrombospondin-1 expression in astrocytes*. Proc Natl Acad Sci U S A, 2006. **103**(24): p. 9321-6.
69. Ottaviani, D., et al., *Reconfiguration of genomic anchors upon transcriptional activation of the human major histocompatibility complex*. Genome Res, 2008. **18**(11): p. 1778-86.

70. Cook, A.C., et al., *Osteopontin induces multiple changes in gene expression that reflect the six "hallmarks of cancer" in a model of breast cancer progression*. Mol Carcinog, 2005. **43**(4): p. 225-36.
71. Metge, B.J., et al., *Elevated osteopontin levels in metastatic melanoma correlate with epigenetic silencing of breast cancer metastasis suppressor 1*. Oncology, 2010. **78**(1): p. 75-86.
72. Hou, G., W. Vogel, and M.P. Bendeck, *The discoidin domain receptor tyrosine kinase DDR1 in arterial wound repair*. J Clin Invest, 2001. **107**(6): p. 727-35.
73. Vogel, W.F., et al., *Discoidin domain receptor 1 tyrosine kinase has an essential role in mammary gland development*. Mol Cell Biol, 2001. **21**(8): p. 2906-17.
74. Yeh, Y.C., et al., *DDR1 triggers epithelial cell differentiation by promoting cell adhesion through stabilization of E-cadherin*. Mol Biol Cell, 2011. **22**(7): p. 940-53.
75. Bruder, C.E., et al., *Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles*. Am J Hum Genet, 2008. **82**(3): p. 763-71.
76. Cao, W., et al., *Mycobacterium tuberculosis antigen Wag31 induces expression of C-chemokine XCL2 in macrophages*. Curr Microbiol, 2008. **57**(3): p. 189-94.
77. Nau, G.J., et al., *A chemoattractant cytokine associated with granulomas in tuberculosis and silicosis*. Proc Natl Acad Sci U S A, 1997. **94**(12): p. 6414-9.
78. Gorman, K.F., C. Julien, and A. Moreau, *The genetic epidemiology of idiopathic scoliosis*. Eur Spine J, 2012.
79. Patwari, P.P., et al., *Monozygotic Twins Discordant for ROHHAD Phenotype*. Pediatrics, 2011.

## **CHAPITRE 4 : DISCUSSION**

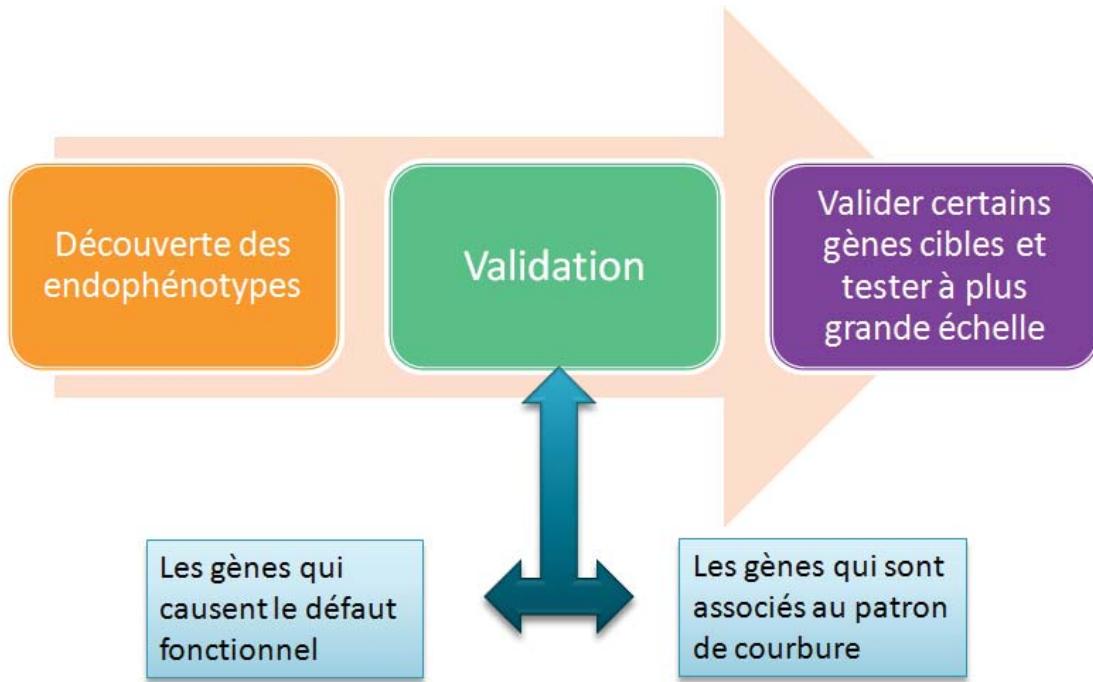
Des progrès considérables ont été accomplis dans les dernières années dans la compréhension de l'étiopathogénie de la scoliose idiopathique de l'adolescent. Cependant, les connaissances actuelles sont encore fragmentées et nous sommes encore loin de comprendre l'ensemble des mécanismes jouant un rôle dans l'initiation de la scoliose *versus* ceux responsables de la progression des déformations rachidiennes. De plus, nous ignorons également la nature des gènes et les interactions nécessaires à l'élaboration des patrons de courbure dans la SIA. La scoliose affecte un groupe de patients très hétérogène comme en témoigne la grande variabilité des phénotypes. Il a donc été nécessaire de trouver d'autres approches pour étudier cette maladie. Devant une telle complexité, nous avons concentré nos efforts dans le cadre de ce projet de maîtrise sur les gènes possiblement impliqués dans la prédisposition à développer une scoliose et/ou jouant un rôle dans la progression des déformations rachidiennes dans la SIA. Nous avons utilisé le concept d'endophénotype afin de stratifier les patients sur une base biochimique et moléculaire plutôt que sur le phénotype de la déformation. Ces endophénotypes représentent des phénomènes plus élémentaires nous permettant d'identifier une voie biologique impliquée dans le pathomécanisme de la SIA. Le nombre de gènes nécessaires à la production d'un endophénotype peut être plus limité que celui impliqué dans la production de la SIA comme entité clinique. Ainsi, les endophénotypes que nous avons identifiés à l'aide de la signalisation de la mélatonine, nous avons conduits à catégoriser les patients SIA en trois groupes qui correspondent à des endophénotypes spécifiques.

De plus, la pertinence clinique de cette classification est soutenue par le fait que des travaux en cours au laboratoire du Dr Moreau ont pu démontrer que dans les familles où plusieurs individus sont atteints de SIA et ce sur plusieurs générations, les sujets scoliotiques appartiennent tous au même endophénotype nonobstant leur type de courbure et la magnitude de leur déformation qui sont des paramètres cliniques toujours très variables (*Alkoume et al., 2013 manuscrit en*

*préparation*). Ces endophénotypes pourraient être les vrais liens génétiques et pourraient nous permettre d'identifier les commutateurs «switches» moléculaires sous-jacents à la pathologie de la SIA.

Nous avons donc appliqué cette stratégie et sélectionné parmi les trois endophénotypes observés des patients SIA présentant une scoliose double majeure (c.-à-d. une courbure thoracique dont l'apex est à droite et une autre au niveau lombaire dont l'apex est à gauche) et thoracique droite afin d'augmenter encore plus l'homogénéité de nos échantillons. Ces deux types de courbures sont de loin les plus fréquents et le fait de sélectionner que des patientes SIA présentant ces patrons de courbures réduit le risque de trouver des gènes qui ne seraient impliqués que dans l'élaboration d'un unique patron de courbure *versus* ceux régulant l'initiation ou la progression des déformations rachidiennes dans leur ensemble. D'ailleurs, il est anticipé que ces gènes, prédisposant à l'élaboration d'un patron de courbure spécifique, puissent être également présents dans la population normale mais qu'en absence d'activation de gènes causant la scoliose il n'y a pas de phénotype observable.

Par conséquent, notre objectif premier était d'examiner les profils moléculaires des patients SIA, afin d'identifier les marqueurs moléculaires responsables des trois différents défauts de signalisation présents chez les patients SIA. Ce projet de maîtrise est donc un premier jalon visant à élucider les mécanismes moléculaires sous-jacents au développement des déformations rachidiennes dans la SIA.



**Figure 13. Les endophénotypes, des outils pour étudier les maladies complexes.**

Le projet est d'identifier les gènes qui sont associés aux trois endophénotypes et de valider certains gènes cibles par RT-QPCR avec un plus grand nombre d'individus par cohortes. Les profils d'expression génique qui sont associés aux patrons de courbure vont être analysés dans le cadre d'études ultérieures.

Dans le premier article, nous avons utilisé la technique de biopuces d'ADN pour étudier l'expression du génome entier à partir d'ostéoblastes dérivés de patients SIA. Cette approche a permis d'identifier des gènes différemment exprimés dans les trois endophénotypes. Nous avons privilégié l'étude des ostéoblastes parce que le tissu osseux est atteint dans la SIA comme en témoigne l'ostéopénie (diminution de la masse osseuse) survenant chez environ 30 % des patients atteints de SIA (*Hans, Biot et al. 1996; Cheng, Guo et al. 1999; Courtois, Collet et al. 1999*). La pertinence clinique du tissu osseux dans la SIA s'explique également parce que l'os répond aussi à la mélatonine *in vitro* et *in vivo* (*Nakade, Koyama et al. 1999; Roth, Kim et al. 1999; Cardinali, Ladizesky et al. 2003*). Toutefois, il est peu probable que l'os soit le seul tissu atteint dans la SIA compte tenu que les récepteurs à la mélatonine sont présents dans la majorité des tissus du corps humain. Il est donc concevable que ce défaut puisse être systémique et que dans ce contexte, le profilage moléculaire des patients SIA selon les trois endophénotypes puissent permettre d'identifier de nouveaux gènes selon le tissu sélectionné. Ceci explique aussi les travaux entrepris dans l'étude des profils moléculaires de jumelles identiques discordantes pour la SIA, en utilisant des cellules sanguines, une approche minimalement invasive et éthiquement acceptable.

#### **4.1 Profilage moléculaire des patients SIA en fonction de leur endophénotype**

Dans le chapitre II, nous avons étudié le changement d'expression génique dans les ostéoblastes de patients SIA en fonction des trois endophénotypes observés chez ces patients. Parmi les limitations expérimentales rencontrées, nous avons décidé d'exclure les ostéoblastes des sujets sains, car ils provenaient de régions anatomiques (bras, jambes) autres que les vertèbres et cela introduisait certains biais expérimental dans nos analyses. Par conséquent, les ostéoblastes obtenus de patients SIA ont été comparés entre eux, car du point de vue clinique et

éthique, il est impossible d'obtenir des spécimens osseux chez les enfants victimes d'un traumatisme de la colonne vertébrale. De plus, ce genre de cas, ne se présente que de façon exceptionnelle au CHU Sainte-Justine. Parmi les 402 gènes identifiés au total, 116 gènes se sont avérés communs aux trois groupes fonctionnels suggérant un rôle dans le risque de progression de la maladie ou dans la formation d'une courbure thoracique droite, seul point en commun entre tous ces patients. Nous avons aussi identifié seulement 21 gènes dont l'expression étaient affectées similairement entre les endophénotypes G1 et G2, mais qui n'étaient pas présents dans les ostéoblastes des patients SIA classés dans le troisième endophénotype (G3). En outre, nous avons identifié 79 gènes exclusifs au profil G3 *vs.* G1, 12 gènes exclusifs au profil G2 *vs.* G1 et 28 gènes exclusifs au profil G2 *vs.* G3. Les gènes exclusifs à chaque profil sont des candidats potentiels pouvant possiblement expliquer les différences moléculaires entre ces trois groupes.

Nous avons constaté que le profil d'expression chez les patients du G1 est distinct de ceux des G2 et G3 alors que ces derniers ont des profils très similaires. Ces différences dans les profils d'expression génique reflètent des activités biochimiques et des mécanismes de signalisation distincts tels que le G1 est résistant à l'ostéopontine (OPN) et les G2 et G3 sont sensibles à l'OPN. Ainsi, certains gènes spécifiques sont associés à un groupe donné, alors que d'autres sont associés aux deux groupes. Par exemple, certains gènes sont impliqués dans la signalisation de l'OPN, vont être soit spécifiquement associé à une résistance à l'OPN (G1), soit vont y être sensibles (G2 et G3). Une caractérisation plus poussée de ces trois endophénotypes a permis d'élucider un rôle central de l'OPN dans l'initiation de la scoliose et la progression des déformations rachidiennes (données non publiées).

Nous avons analysé les données de Microarray avec une autre méthode plus rigoureuse, ANOVA, qui a permis de révéler que 25 gènes étaient différentiellement exprimés, de façon significative, dans les trois groupes. Nous nous attendions à ce qu'ils soient associés à une réponse biologique

envers l'OPN (résistance *vs.* sensibilité). Par la suite, la validation des gènes d'intérêt par la technique de qPCR a été réalisée dans une cohorte plus large. Par cette approche, les gènes *MKX*, *SLPI* et *CLGN* ont été confirmés et s'expriment de façon différentielle selon chaque endophénotype. Soit, *MKX* augmente 6.5 fois, *SLPI* augmente 15 fois et *CLGN* augmente 20 fois dans le groupe résistant à l'OPN (G1) comparé aux groupes sensibles à l'OPN (G2 et G3), suggérant leur implication dans la régulation signalétique induite par l'OPN.

#### **4.1.1 Contribution des résultats du profilage moléculaire dans la compréhension des aspects cliniques de la SIA.**

Les gènes impliqués dans la SIA sont multiples et hétérogènes, par conséquent, avec les méthodes traditionnelles comme les analyses de liaisons génétiques ou les études d'association, il est très difficile de trouver les gènes responsables de cette maladie complexe. Dans nos études, nous avons d'abord introduit le concept d'endophénotypes biologiques, les variations génétiques étant partitionnées sur des bases biochimiques ce qui, nous croyons, augmente les chances de détecter des gènes d'intérêt associés au pathomécanisme de la maladie. Deuxièmement, on a émis l'hypothèse que ces endophénotypes sont causés par des ensembles distincts de gènes (G1 *vs* G2 et G3) ou assez similaires (G2 *vs* G3) permettant grâce aux profils d'expression de regrouper biochimiquement les patients en utilisant leur défaut différentiel de signalisation des protéines Gi normalement couplées aux récepteurs à la mélatonine, MT1 et MT2. Troisièmement, cette approche d'endophénotypes a permis d'aller au-delà de la variabilité phénotypique présente chez les patients SIA, en utilisant la technique de microarray (bio-puce d'ADN Affymetrix) ce qui a permis d'identifier simultanément un ensemble de gènes d'intérêt. Grâce à l'analyse des profils d'expression des trois endophénotypes, nous avons identifié des gènes spécifiquement associés à chacun d'entre eux. Les études fonctionnelles en cours de réalisation de ces gènes et

l'identification éventuelle de leurs partenaires moléculaires pourraient nous aider à identifier les voies moléculaires et les gènes candidats associés à la SIA et/ou certaines co-morbidités comme l'ostéoporose, l'anxiété et la dépression.

#### **4.1.2 Identification des gènes impliqués dans le métabolisme osseux**

Une des co-morbidités associées à la scoliose est une faible masse osseuse ou ostéopénie (*Cheng and Guo 1997; Hung, Qin et al. 2005; Chiru 2011; Dede, Akel et al. 2011*). Le mécanisme précis et les causes de la perte osseuse chez les patients SIA n'ont à ce jour pas encore été identifiés. Dans notre étude, nous avons identifié des gènes possiblement associés à la pathologie de l'ostéopénie (une diminution de la masse osseuse), notamment au niveau des gènes de la famille des MMPs (Matrix Metalloprotease) comme *MMP1*, *MMP3* et *MMP13* (*Bolton, Stone et al. 2009; Bellido, Lugo et al. 2011; Jemtland, Holden et al. 2011*). Ces métallo-protéases régulent la dégradation et le remodelage de la matrice extracellulaire et sont essentielles au bon fonctionnement du métabolisme osseux (*Inoue, Mikuni-Takagaki et al. 2006*). Dans notre analyse, les gènes *MMP1*, *MMP3* et *MMP13* sont tous sous-exprimés dans les ostéoblastes des patients SIA G3 lorsque comparés aux niveaux d'expressions mesurés pour les patients SIA du G1 et G2. Le gène *LRRC17* (Leucine Rich Repeat containing 17) a également retenu notre attention compte tenu de son rôle comme régulateur négatif de la différenciation des ostéoclastes induite par RANKL (NF-kappaB ligand) (*Kim, Kim et al. 2009*). Ce gène est sur-exprimé dans les ostéoblastes des patients du G2 et G3. Nous avons également identifié le gène *CHRD1* (Chordin-like 1) qui joue un rôle dans le développement du système squelettique, ainsi que dans l'ossification via la voie de signalisation des BMPs (Bone Morphogenetic Protein) (*Kane, Godson et al. 2008*). Il est à noter que *CHRD1* est sous-exprimé dans les ostéoblastes

des patients du G2 et G3 lorsque comparés aux ostéoblastes des patients du G1. D'autre part, le gène *ITGB3* code pour l'intégrine  $\beta 3$  qui joue un rôle dans l'activation de la signalisation du facteur de croissance IGF-1 (*Xi, Maile et al. 2008*). Et l'intégrine interagit au niveau de la membrane avec l'OPN et la métalloprotéase MMP2 (*Cui, Takahashi et al. 2007*). Lessey et al. ont démontré dans une étude que la protéine intégrine bêta 3 est régulée négativement par les œstrogènes dans les lignes cellulaires d'endométriales. En concurrence contre l'œstrogènes avec différents modulateurs sélectifs des récepteurs aux œstrogènes (SERMs) et les agonistes et antagonistes de récepteurs aux œstrogènes, l'inhibition de l'expression de l'intégrine bêta 3 par les œstrogènes peuvent être atténués (*Lninger, et al. 2006*).

Tous ces gènes sont impliqués dans le métabolisme osseux et sont exprimés différemment dans le G3, suggérant ainsi que les patients classés dans ce sous-groupe pourraient être plus à risque de développer une ostéopénie.

#### **4.1.3 Concordance des résultats de profilage moléculaire avec les études publiées par d'autres chercheurs.**

Les gènes qui sont différemment exprimés dans les trois groupes définis par nos endophénotypes sont de nouveaux gènes candidats et ne corrèlent pas aux gènes candidats identifiés dans de précédentes études. Ceci n'est pas vraiment surprenant puisque ces études classiques ne prennent en compte que le phénotype (la courbure principale) qui est un paramètre peu fiable car très hétérogène entre les membres scoliotiques d'une même famille. Toutefois, il est difficile de préciser à ce stade de notre étude si les changements observés dans l'expression de nos gènes d'intérêt est la résultante du défaut de signalisation observé chez les patients SIA ou la cause.

#### **4.1.4 Discrimination des gènes dont l'expression contribue au pathomécanisme de la SIA versus ceux dont l'expression est modifiée en conséquence de la maladie**

Effectivement, cette étude de profilage d'expression présente pour la première fois une analyse de l'expression génique sur le génome entier de patients scoliotiques. Toutefois, c'est l'analyse fonctionnelle de certains gènes qui nous permettra de discriminer si les gènes identifiés dans cette étude jouent un rôle dans la pathogenèse de la SIA ou sont simplement altérés en réponse à la maladie. Dans cette étude, nous avons identifié des points communs et distinctifs chez nos trois endophénotypes retrouvés chez les patients SIA. Il est conceivable que certains gènes impliqués dans la signalisation des protéines Gi, puissent être considérées comme des gènes contributeurs au développement de cette maladie. Mais pour évaluer plus précisément, cette question, on doit combiner différentes approches expérimentales : génotypique (l'analyse de liaisons génétiques, SNPs, les études d'association du génome entier); protéomique (ELISA, Western Blot, Immunochimie); métabolique (Cellular Dielectric Spectroscopy, G-proteins Assay), et l'utilisation de modèles animaux.

Nous avons initialement sélectionné 22 gènes pour la phase de validation par qPCR dans un plus grand nombre de patients SIA. Parmi les gènes candidats retenus quatre gènes ont été re-confirmé à ce jour par qPCR et sont d'intérêt car impliqués dans la formation de la matrice extracellulaire : *MMP1*, *CHI3L1*, *SPP1* et *COMP*.

*MMP1* est impliqué dans la dégradation de la matrice extracellulaire observée dans certaine maladies telles que l'arthrite (*Hui, Litherland et al. 2012*) et aussi dans la SIA (*Aulisa, Papaleo et al. 2007; Liu, Tang et al. 2010; Grivas, Vasiliadis et al. 2011*). *CHI3L1* semble jouer un rôle dans le processus d'inflammation et du remodelage tissulaire (*Zivanovic, Petrovic-Rackov et al. 2010*). Le gène *SPP1* code quant à lui pour l'OPN, une cytokine sécrétée, qui est impliquée dans

plusieurs processus physiologiques et pathologiques (*Liu, Tam et al. 2012*). Le gène *COMP* code pour une protéine de la matrice extracellulaire appelée cartilage oligomeric matrix protein et est d'intérêt car il est localisé dans un locus de susceptibilité de la SI (Chr19p13, identifié par Alden *et al.*). Les autres gènes dont nous avons choisi de valider l'expression génique par qPCR ont été choisis sur leur niveau d'expression génétique, soit hautement exprimé ou faiblement en comparaison des endophénotypes. Bien entendu, c'est le point de départ de notre projet de validation et nous sommes en train de valider un plus grand nombre de gènes avec plus de patients.

#### **4.1.5 Contributions possibles d'autres tissus ou types de cellules dans la SIA**

Parmi les limitations de la présente étude, nous avons ciblé jusqu'ici que le tissu osseux et certaines cellules sanguines alors que la SIA affecte d'autres tissus comme les muscles, les nerfs périphériques et même le cerveau. Dans ce contexte, des études similaires pourraient être envisageables par exemple à partir de biopsies de la musculature para-spinale.

#### **4.2 Études de l'expression génique chez jumelles identiques discordantes**

Dans le chapitre III, nous avons comparé les profils d'expression au niveau d'une paire de jumelles monozygotiques discordantes pour la scoliose. Les jumeaux monozygotes discordants représentent un modèle expérimental intéressant pour l'étude de maladies complexes comme la scoliose idiopathique, car au départ il partage le même bagage génétique (*MacGregor, Snieder et al. 2000*). Les jumelles monozygotes discordantes sont génétiquement identiques et seraient potentiellement utile dans l'étude de l'expression génique pour les maladies complexes,

parce qu'il y aurait beaucoup moins de variation dans l'expression génique par rapport aux individus non apparentés. Les études d'expression génétique chez les jumelles monozygotiques ne sont pas nombreuses (*Wilson, Kent et al. 2000; Cheung, Conlin et al. 2003; Mak, Hampson et al. 2004*).

Une recherche basée sur une autre maladie complexe, l'ostéoporose, par l'équipe du Dr Tim D. Spector a révélé avec succès des gènes impliqués dans la physiologie de l'os (*Mak, Hampson et al. 2004*). Nous avons donc exploré cette approche, preuve du potentiel de ce concept, compte tenu de la grande hétérogénéité de la SIA.

Dans ce contexte, nous avons eu l'opportunité d'étudier une première paire de jumelles discordantes avec les limitations suivantes : premièrement, la discordance observée n'est pas complète dans le sens que les deux jumelles présentaient une scoliose au moment de l'atteinte de leur maturité squelettique. Néanmoins, il y avait discordance quant à la progression et le patron de courbure (34° de courbure en thoracique et 43° au niveau lombaire pour la première versus seulement 14° au niveau lombaire pour la seconde). Assumant une marge d'erreur possible et reconnue de  $\pm 5^\circ$  dans ce genre de mesure radiologique, il est même possible que la seconde jumelle ne soit pas considérée comme scoliotique si on enlève cette fluctuation de 5° (car il faut au moins 10° d'angulation pour être considéré atteint d'une scoliose).

Deuxièmement, pour des raisons éthiques et cliniques, nous ne pouvions réaliser une biopsie osseuse et/ou musculaire et nous avons préféré tester les cellules sanguines (PBMCs). Ce choix s'explique par le fait que les PBMCs des patients SIA présentent tout comme les ostéoblastes le même défaut de signalisation permettant de les classer dans l'un ou l'autre des trois endophénotypes (G1, G2 ou G3). Ce résultat sous-entend que le même processus causant le

défaut de signalisation impliquant les protéines Gi survient autant au niveau des cellules sanguines (PBMCs) que des ostéoblastes dérivés de ces mêmes patients SIA. Par conséquent, une variation des niveaux d'expression génique entre ces deux jumelles ne pourrait s'expliquer que par des mécanismes épigénétiques en lien avec l'environnement.

Par ailleurs, les changements transcriptionnels observés avec les PBMCs sont aussi la résultante des états physiologiques ou pathologiques présents au moment du prélèvement sanguin (*Whitney, Diehn et al. 2003; Chon, Gaillard et al. 2004; Lampe, Stepaniants et al. 2004*). Néanmoins, les réponses d'expression génique des PBMCs peuvent potentiellement fournir une alerte précoce pour chaque individu, en les combinant avec la technique de puces à ADN. Ainsi, il est possible de mesurer les niveaux d'expression de milliers de gènes dans de petits échantillons biologiques. Dans le contexte de la SIA, des anomalies sanguines ont aussi été observées chez les patients SIA (*Yarom, Meyer et al. 1982; Kindsfater, Lowe et al. 1994*).

Avec des analyses stringentes, nous avons identifié 26 gènes ayant une expression différentielle, dont trois associés à des loci de susceptibilité pour la SIA.

La variation de la magnitude et de la morphologie entre les jumelles suggèrent une contribution potentielle de facteurs environnementaux et/ou épigénétiques pouvant influencer les niveaux d'expression de certains gènes dans la SIA. Dans le cadre de la validation de ce concept (article 2), nous avons identifié quatre gènes d'intérêt soit : *HLA-DQB1*, *MSLN*, *TLR3* et *THBS1*. Ces gènes sont susceptibles à la méthylation au niveau de leur promoteur. Les gènes *HLA-DQA1* et *DDR* sont sur le locus MHC, et sont souvent co-acétylés. Une des limitations dans l'utilisation de cellules sanguines (PBMCs) est la possibilité d'introduire une variabilité dans les profils d'expression en réponse à une infection transitoire ou même à une médication modulant la

réponse immunitaire comme lors d'un rhume ou suite à une allergie. Il est donc important de contrôler ces paramètres lors de la prise de sang afin de s'assurer que les jumelles soient exemptes depuis au moins 15 jours de tels symptômes. Sur le plan analytique, l'expression de certains gènes est tributaire du tissu analysé. Il est donc tout à fait normal de ne pas retrouver dans certains cas la vaste majorité des gènes identifiés dans les ostéoblastes d'où l'importance d'analyser plusieurs tissus provenant dans la mesure du possible des mêmes patients. Cette approche est déjà en cours dans le laboratoire du Dr Moreau.

En résumé, cette étude offre un appui solide pour utilisation de PBMCs isolés du sang périphérique dans les études d'expression génique, ceci en tant que base pour la détection et le diagnostic de la maladie chez patients.

Dans le cadre des deux études présentées dans ce mémoire, nous avons identifié très peu de gènes communs aux trois endophénotypes. Le seul gène répondant à ce critère est *ITGB3* (intégrine β3). Cette intégrine est impliquée dans plusieurs processus biologiques notamment dans la régulation de la croissance, une problématique pertinente à la SIA puisque les patients scoliotiques présentent une croissance accélérée qui coïncide avec l'évolution de leur déformation rachidienne.

#### **4.2.1 Travaux futur**

À ce stade de la recherche, nous avons seulement travaillé avec une paire de jumelles discordantes pour la SIA. Nous comptons recruter d'autres paires de jumelles SIA monozygotiques discordantes, afin d'avoir une image plus complète des changements observés dans les profils d'expression et d'identifier les gènes dont l'expression est altérée par la maladie et/ou le degré de progression.

Comme nous avons discuté précédemment, il y a des limitations expérimentales reliées au fait d'utiliser seulement les outils de profilage d'expression génique pour identifier des gènes candidats. En effet, il est difficile de juger si les changements d'expression observés sont en réalité la source du problème, ou plutôt une conséquence de la maladie. Cette situation est toujours un dilemme dans les études génétiques. Donc en combinant d'autres approches expérimentales, on pourrait confirmer ces résultats notamment par des approches fonctionnelles comme le gain ou la perte de fonction de ces gènes en utilisant des méthodes d'atténuation de gènes par siRNA au niveau cellulaire ou chez le poisson zèbre. De plus, il est crucial de valider certains gènes dans une cohorte plus large par qPCR provenant de notre population, mais aussi dans d'autres populations.

#### **4.2.2 Importance des méthodes d'analyses statistiques**

Les puces à ADN ou microarrays permettent d'étudier l'expression de plusieurs milliers de gènes simultanément. C'est un outil très puissant à utiliser, mais de grandes quantités de données (milliers de gènes étudiés en même temps) sont générées, donc les méthodes d'analyses statistiques deviennent très importantes dans ce genre d'étude.

Les gènes différentiellement exprimés sont identifiés en calculant l'expression pour chaque gène avec un gène de référence. Dans les études de microarray, il y a beaucoup de faux positifs, c'est-à-dire des gènes qui sont déclarés différentiellement exprimés alors qu'ils ne le sont pas. Les données brutes sont donc normalisées en utilisant la méthode de Robust Multi-array Analysis, analysées par la méthode statistique EB développée par Wright&Simons et utilisé par le logiciel FlexArray. Par la suite la méthode de correction de Benjamini Hochberg a été utilisée pour ajuster la probabilité statistique (valeur p) pour les tests multiples et ainsi contrôler le taux de faux positifs ou FDR (False Discovery Rate). Ce processus d'analyse est essentiel et permet d'éliminer une proportion de faux positifs pour correctement choisir les gènes candidats pour les analyses subséquentes par qPCR. Pour les résultats de qPCR, un test ANOVA a été réalisé.

## **CHAPITRE 5 : CONCLUSION**

Les travaux de notre laboratoire ont démontré un défaut différentiel de la signalisation des récepteurs de la mélatonine qui sont couplés aux protéines Gi chez les patients SIA. Ce défaut affecte toutes les protéines Gi chez les patients SIA et est causé par une phosphorylation accrue des protéines Gi au niveau de certaines sérines. Les essais fonctionnels développés par l'équipe du Dr Moreau ont démontré trois types de réponses chez les cellules de patients SIA permettant ainsi leur classification en trois endophénotypes. Les résultats de ce projet permettront de mettre en évidence les fondements moléculaires de la méthode de classification des patients SIA, et d'identifier de nouveaux gènes et leurs fonctionnalités associées à chacun des endophénotypes de la SIA. Nos études ont révélé des gènes importants dans la signalisation des protéines G et d'autres régulant possiblement la voie moléculaire de l'OPN dans l'étiopathogenèse de la SIA, ainsi que des gènes impliqués dans le métabolisme osseux.

### **La pertinence scientifique et clinique de ce projet de maîtrise**

Les résultats d'expression génique nous aident à mieux définir le profilage moléculaire au sein de chaque endophénotype. Cette étude pourrait contribuer à l'utilisation de marqueurs génétiques pour distinguer les différences entre les trois endophénotypes. À cet égard, nos résultats permettront l'identification de cibles potentielles et de marqueurs moléculaires spécifiques, ainsi que de mieux comprendre l'étiopathogenèse de cette maladie. Une meilleure compréhension des bases moléculaires sous-jacentes aux trois endophénotypes identifiés devrait se traduire à plus long terme par le développement d'outils pronostiques et d'arriver à proposer des traitements plus personnalisés. Un autre impact clinique serait l'identification des gènes de prédisposition afin de faire une détection plus précoce de la maladie chez de jeunes enfants encore asymptomatiques (0-8 ans). Une bonne compréhension des bases génétiques pourrait aider à identifier les personnes à risque, et potentiellement conduire à la mise en place des

procédures préventives et/ou thérapeutiques et finalement enrayer le développement de la maladie.

## **CHAPITRE 6 : RÉFÉRENCES**

## Références

- Abu-Amro, K. (2008). "Gene symbol: ROBO3. Disease: Gaze Palsy, horizontal, with progressive scoliosis." Hum Genet **123**(5): 540.
- Abu-Amro, K. K., S. Kapoor, et al. (2011). "Horizontal gaze palsy and progressive scoliosis due to a deleterious mutation in ROBO3." Ophthalmic Genet **32**(4): 231-236.
- Akoume, M. Y., B. Azeddine, et al. (2010). "Cell-based screening test for idiopathic scoliosis using cellular dielectric spectroscopy." Spine (Phila Pa 1976) **35**(13): E601-608.
- Alden, K. J., B. Marosy, et al. (2006). "Idiopathic scoliosis: identification of candidate regions on chromosome 19p13." Spine (Phila Pa 1976) **31**(16): 1815-1819.
- Amouri, R., H. Nehdi, et al. (2009). "Allelic ROBO3 heterogeneity in Tunisian patients with horizontal gaze palsy with progressive scoliosis." J Mol Neurosci **39**(3): 337-341.
- Andersen, M. O., K. Thomsen, et al. (2007). "Adolescent idiopathic scoliosis in twins: a population-based survey." Spine (Phila Pa 1976) **32**(8): 927-930.
- Aulisa, L., P. Papaleo, et al. (2007). "Association between IL-6 and MMP-3 gene polymorphisms and adolescent idiopathic scoliosis: a case-control study." Spine (Phila Pa 1976) **32**(24): 2700-2702.
- Avikainen, V. J., A. Rezasoltani, et al. (1999). "Asymmetry of paraspinal EMG-time characteristics in idiopathic scoliosis." J Spinal Disord **12**(1): 61-67.
- Azeddine, B., K. Letellier, et al. (2007). "Molecular determinants of melatonin signaling dysfunction in adolescent idiopathic scoliosis." Clin Orthop Relat Res **462**: 45-52.
- Bagnall, K., V. J. Raso, et al. (1999). "The effects of melatonin therapy on the development of scoliosis after pinealectomy in the chicken." J Bone Joint Surg Am **81**(2): 191-199.
- Bagnall, K. M., M. Beuerlein, et al. (2001). "Pineal transplantation after pinealectomy in young chickens has no effect on the development of scoliosis." Spine (Phila Pa 1976) **26**(9): 1022-1027.

- Bagnall, K. M., V. J. Raso, et al. (1996). "Melatonin levels in idiopathic scoliosis. Diurnal and nocturnal serum melatonin levels in girls with adolescent idiopathic scoliosis." Spine (Phila Pa 1976) **21**(17): 1974-1978.
- Barrios, C., S. Cortes, et al. (2011). "Anthropometry and body composition profile of girls with nonsurgically treated adolescent idiopathic scoliosis." Spine (Phila Pa 1976) **36**(18): 1470-1477.
- Bashiardes, S., R. Veile, et al. (2004). "SNTG1, the gene encoding gamma 1-syntrophin: a candidate gene for idiopathic scoliosis." Hum Genet **115**(1): 81-89.
- Beals, R. K. (1973). "Nosologic and genetic aspects of scoliosis." Clin Orthop Relat Res(93): 23-32.
- Bell, M. and A. S. Teebi (1995). "Autosomal dominant idiopathic scoliosis?" Am J Med Genet **55**(1): 112.
- Bellido, M., L. Lugo, et al. (2011). "Improving subchondral bone integrity reduces progression of cartilage damage in experimental osteoarthritis preceded by osteoporosis." Osteoarthritis Cartilage **19**(10): 1228-1236.
- Bjornsson, H. T., M. D. Fallin, et al. (2004). "An integrated epigenetic and genetic approach to common human disease." Trends Genet **20**(8): 350-358.
- Blank, R. D., C. L. Raggio, et al. (1999). "A genomic approach to scoliosis pathogenesis." Lupus **8**(5): 356-360.
- Bolton, C. E., M. D. Stone, et al. (2009). "Circulating matrix metalloproteinase-9 and osteoporosis in patients with chronic obstructive pulmonary disease." Chron Respir Dis **6**(2): 81-87.
- Bozenna Olszanska, et al. (2003) "Bidirectional Communication Between The Pineal Gland And The Immune System. ..." Canadian Journal Of Physiology & Pharmacology **81**(4): 342.
- Bredoux, R., E. Corvazier, et al. (2006). "Human platelet Ca<sup>2+</sup>-ATPases: new markers of cell differentiation as illustrated in idiopathic scoliosis." Platelets **17**(6): 421-433.

Brodner, W., P. Krepler, et al. (2000). "Melatonin and adolescent idiopathic scoliosis." J Bone Joint Surg Br **82**(3): 399-403.

Burwell, R. G. (2003). "Aetiology of idiopathic scoliosis: current concepts." Pediatr Rehabil **6**(3-4): 137-170.

Burwell, R. G. and P. H. Dangerfield (2002). "Etiologic theories of idiopathic scoliosis: neurodevelopmental concepts to be evaluated." Stud Health Technol Inform **91**: 15-19.

Burwell, R. G., P. H. Dangerfield, et al. (2011). "Adolescent idiopathic scoliosis (AIS), environment, exposome and epigenetics: a molecular perspective of postnatal normal spinal growth and the etiopathogenesis of AIS with consideration of a network approach and possible implications for medical therapy." Scoliosis **6**(1): 26.

Burwell, R. G., N. J. James, et al. (1983). "Standardised trunk asymmetry scores. A study of back contour in healthy school children." J Bone Joint Surg Br **65**(4): 452-463.

Cajochen, C., K. Kräuchi, and A. Wirz-Justice. (2003) "Role Of Melatonin In The Regulation Of Human Circadian Rhythms And Sleep." Journal Of Neuroendocrinology **15**(4): 432-437.

Cardinali, D. P., M. G. Ladizesky, et al. (2003). "Melatonin effects on bone: experimental facts and clinical perspectives." J Pineal Res **34**(2): 81-87.

Cardinali, D. P., et al. (2008) "Physiological Effects Of Melatonin: Role Of Melatonin Receptors And Signal Transduction Pathways." Progress In Neurobiology **85**.3: 335-353

Chen, Z., N. L. Tang, et al. (2009). "Promoter polymorphism of matrilin-1 gene predisposes to adolescent idiopathic scoliosis in a Chinese population." Eur J Hum Genet **17**(4): 525-532.

Cheng, J. C. and X. Guo (1997). "Osteopenia in adolescent idiopathic scoliosis. A primary problem or secondary to the spinal deformity?" Spine (Phila Pa 1976) **22**(15): 1716-1721.

Cheng, J. C., X. Guo, et al. (1999). "Persistent osteopenia in adolescent idiopathic scoliosis. A longitudinal follow up study." Spine (Phila Pa 1976) **24**(12): 1218-1222.

Cheng, J. C., N. L. Tang, et al. (2007). "Genetic association of complex traits: using idiopathic scoliosis as an example." Clin Orthop Relat Res **462**: 38-44.

Cheung, C. S., W. T. Lee, et al. (2006). "Generalized osteopenia in adolescent idiopathic scoliosis--association with abnormal pubertal growth, bone turnover, and calcium intake?" Spine (Phila Pa 1976) **31**(3): 330-338.

Cheung, J., J. P. Halbertsma, et al. (2005). "A preliminary study on electromyographic analysis of the paraspinal musculature in idiopathic scoliosis." Eur Spine J **14**(2): 130-137.

Cheung, K. M., T. Wang, et al. (2005). "The effect of pinealectomy on scoliosis development in young nonhuman primates." Spine (Phila Pa 1976) **30**(18): 2009-2013.

Cheung, V. G., L. K. Conlin, et al. (2003). "Natural variation in human gene expression assessed in lymphoblastoid cells." Nat Genet **33**(3): 422-425.

Chiru, M. (2011). "Adolescent idiopathic scoliosis and osteopenia)." Maedica (Buchar) **6**(1): 17-22.

Chon, H., C. A. Gaillard, et al. (2004). "Broadly altered gene expression in blood leukocytes in essential hypertension is absent during treatment." Hypertension **43**(5): 947-951.

Cook, S. D., A. F. Harding, et al. (1987). "Trabecular bone mineral density in idiopathic scoliosis." J Pediatr Orthop **7**(2): 168-174.

Courtois, I., P. Collet, et al. (1999). "Bone mineral density at the femur and lumbar spine in a population of young women treated for scoliosis in adolescence." Rev Rhum Engl Ed **66**(12): 705-710.

Cowell, H. R. (1970). "Genetic aspects of orthopedic diseases." Am J Nurs **70**(4): 763-767.

Cowell, H. R., J. N. Hall, et al. (1972). "Genetic aspects of idiopathic scoliosis. A Nicholas Andry Award essay, 1970." Clin Orthop Relat Res **86**: 121-131.

Cui, R., F. Takahashi, et al. (2007). "Abrogation of the interaction between osteopontin and alphavbeta3 integrin reduces tumor growth of human lung cancer cells in mice." Lung Cancer **57**(3): 302-310.

Deacon, P., B. M. Flood, et al. (1984). "Idiopathic scoliosis in three dimensions. A radiographic and morphometric analysis." J Bone Joint Surg Br **66**(4): 509-512.

Dede, O., I. Akel, et al. (2011). "Is decreased bone mineral density associated with development of scoliosis? A bipedal osteopenic rat model." Scoliosis **6**(1): 24.

Delagrange, P., et al. (2008) "Melatonin Receptors, Heterodimerization, Signal Transduction And Binding Sites: What's New?." British Journal Of Pharmacology **154**(6): 1182-1195.

Dickson, R. ( 2001). "Early-onset idiopathic scoliosis." In The Pediatric Spine, 2nd edition, Stewart L. Weinstein, ed. Lippincott Williams & Wilkins, Philadelphia, PA, USA.

Dickson, R. A. (1983). "Scoliosis in the community." Br Med J (Clin Res Ed) **286**(6379): 1745.

Dickson, R. A. (1992). "The etiology and pathogenesis of idiopathic scoliosis." Acta Orthop Belg **58 Suppl 1**: 21-25.

Dubousset, J. (1983). "[Orthopedic disorders of the spine]." Soins Chir(32): 10-15.

Dubousset, J. and M. Machida (2001). "[Possible role of the pineal gland in the pathogenesis of idiopathic scoliosis. Experimental and clinical studies]." Bull Acad Natl Med **185**(3): 593-602; discussion 602-594.

Fagan, A. B., D. J. Kennaway, et al. (2009). "Pinealectomy in the chicken: a good model of scoliosis?" Eur Spine J **18**(8): 1154-1159.

Fagan, A. B., D. J. Kennaway, et al. (1998). "Total 24-hour melatonin secretion in adolescent idiopathic scoliosis. A case-control study." Spine (Phila Pa 1976) **23**(1): 41-46.

Farce, A., A. O. Chugunov, et al. (2008). "Homology modeling of MT1 and MT2 receptors." Eur J Med Chem **43**(9): 1926-1944.

Faria Poloni, Joice, Bruno Feltes, and Diego Bonatto. (2011) "Melatonin As A Central Molecule Connecting Neural Development And Calcium Signaling." Functional & Integrative Genomics **11**(3): 383-388.

Federman, A. D., B. R. Conklin, K. A. Schrader, R. R. Reed and H. R. Bourne (1992). "Hormonal stimulation of adenylyl cyclase through Gi-protein beta gamma subunits." *Nature* **356**(6365): 159-61.

Feinberg, A. P. (2008). "Epigenetics at the epicenter of modern medicine." *JAMA* **299**(11): 1345-1350.

Feinberg, A. P. (2010). "Genome-scale approaches to the epigenetics of common human disease." *Virchows Arch* **456**(1): 13-21.

Fidler, M. W. and R. L. Jowett (1976). "Muscle imbalance in the aetiology of scoliosis." *J Bone Joint Surg Br* **58**(2): 200-201.

Fisk, J. R. and W. H. Bunch (1979). "Scoliosis in neuromuscular disease." *Orthop Clin North Am* **10**(4): 863-875.

Fong, D. Y., C. F. Lee, et al. (2010). "A meta-analysis of the clinical effectiveness of school scoliosis screening." *Spine (Phila Pa 1976)* **35**(10): 1061-1071.

Gao, X., D. Gordon, et al. (2007). "CHD7 gene polymorphisms are associated with susceptibility to idiopathic scoliosis." *Am J Hum Genet* **80**(5): 957-965.

Garcia-Perganeda, A., D. Pozo, et al. (1997). "Signal transduction for melatonin in human lymphocytes: involvement of a pertussis toxin-sensitive G protein." *J Immunol* **159**(8): 3774-3781.

Garland, H. G. (1934). "Hereditary Scoliosis." *Br Med J* **1**(3816): 328.

Goldberg, C. J., F. E. Dowling, et al. (1993). "Adolescent idiopathic scoliosis: is rising growth rate the triggering factor in progression?" *Eur Spine J* **2**(1): 29-36.

Gonyea, W. J., C. Moore-Woodard, et al. (1985). "An evaluation of muscle pathology in idiopathic scoliosis." *J Pediatr Orthop* **5**(3): 323-329.

Gorman, K. F., C. Julien, et al. (2012). "The genetic epidemiology of idiopathic scoliosis." *Eur Spine J*.

Grivas, T. B., E. S. Vasiliadis, et al. (2011). "Expression of matrix metalloproteinase-1 (MMP-1) in Wistar rat's intervertebral disc after experimentally induced scoliotic deformity." Scoliosis **6**(1): 9.

Guo, X., W. W. Chau, et al. (2005). "Relative anterior spinal overgrowth in adolescent idiopathic scoliosis--result of disproportionate endochondral-membranous bone growth? Summary of an electronic focus group debate of the IBSE." Eur Spine J **14**(9): 862-873.

Gurnett, C. A., F. Alaee, et al. (2009). "Genetic linkage localizes an adolescent idiopathic scoliosis and pectus excavatum gene to chromosome 18 q." Spine (Phila Pa 1976) **34**(2): E94-100.

Hagglund, G., J. Karlberg, et al. (1992). "Growth in girls with adolescent idiopathic scoliosis." Spine (Phila Pa 1976) **17**(1): 108-111.

Hans, D., B. Biot, et al. (1996). "No diffuse osteoporosis in lumbar scoliosis but lower femoral bone density on the convexity." Bone **18**(1): 15-17.

Harrington, P. R. (1977). "The etiology of idiopathic scoliosis." Clin Orthop Relat Res(126): 17-25.

Hilibrand, A. S., L. C. Blakemore, et al. (1996). "The role of melatonin in the pathogenesis of adolescent idiopathic scoliosis." Spine (Phila Pa 1976) **21**(10): 1140-1146.

Hou, Y., I. Azpiazu, A. Smrcka and N. Gautam (2000). "Selective role of G protein gamma subunits in receptor interaction." J Biol Chem **275**(50): 38961-4.

Hui, W., G. J. Litherland, et al. (2012). "Leptin produced by joint white adipose tissue induces cartilage degradation via upregulation and activation of matrix metalloproteinases." Ann Rheum Dis **71**(3): 455-462.

Hung, V. W., L. Qin, et al. (2005). "Osteopenia: a new prognostic factor of curve progression in adolescent idiopathic scoliosis." J Bone Joint Surg Am **87**(12): 2709-2716.

Ikeda, S. R. (1996). "Voltage-dependent modulation of N-type calcium channels by G protein beta gamma subunits." Nature **380**(6571): 255-8.

Inoue, K., Y. Mikuni-Takagaki, et al. (2006). "A crucial role for matrix metalloproteinase 2 in osteocytic canalicular formation and bone metabolism." J Biol Chem **281**(44): 33814-33824.

Inoue, M., S. Minami, et al. (2002). "Association between estrogen receptor gene polymorphisms and curve severity of idiopathic scoliosis." Spine (Phila Pa 1976) **27**(21): 2357-2362.

James, J. I. (1954). "Idiopathic scoliosis; the prognosis, diagnosis, and operative indications related to curve patterns and the age at onset." J Bone Joint Surg Br **36-B**(1): 36-49.

Jemtland, R., M. Holden, et al. (2011). "Molecular disease map of bone characterizing the postmenopausal osteoporosis phenotype." J Bone Miner Res **26**(8): 1793-1801.

Jockers, R., P. Maurice, et al. (2008). "Melatonin receptors, heterodimerization, signal transduction and binding sites: what's new?" Br J Pharmacol **154**(6): 1182-1195.

Justice, C. M., N. H. Miller, et al. (2003). "Familial idiopathic scoliosis: evidence of an X-linked susceptibility locus." Spine (Phila Pa 1976) **28**(6): 589-594.

Kahmann, R. D., J. M. Donohue, et al. (1992). "Platelet function in adolescent idiopathic scoliosis." Spine (Phila Pa 1976) **17**(2): 145-148.

Kane, R., C. Godson, et al. (2008). "Chordin-like 1, a bone morphogenetic protein-4 antagonist, is upregulated by hypoxia in human retinal pericytes and plays a role in regulating angiogenesis." Mol Vis **14**: 1138-1148.

Katz, A., D. Wu and M. I. Simon (1992). "Subunits beta gamma of heterotrimeric G protein activate beta 2 isoform of phospholipase C." Nature **360**(6405): 686-9.

Kesling, K. L. and K. A. Reinker (1997). "Scoliosis in twins. A meta-analysis of the literature and report of six cases." Spine (Phila Pa 1976) **22**(17): 2009-2014; discussion 2015.

Kim, D., D. L. Lewis, L. Graziadei, E. J. Neer, D. Bar-Sagi and D. E. Clapham (1989). "G protein beta gamma-subunits activate the cardiac muscarinic K<sup>+</sup>-channel via phospholipase A2." Nature **337**(6207): 557-60.

Kim, T., K. Kim, et al. (2009). "Identification of LRRc17 as a negative regulator of receptor activator of NF-kappaB ligand (RANKL)-induced osteoclast differentiation." J Biol Chem **284**(22): 15308-15316.

Kindsfater, K., T. Lowe, et al. (1994). "Levels of platelet calmodulin for the prediction of progression and severity of adolescent idiopathic scoliosis." J Bone Joint Surg Am **76**(8): 1186-1192.

Lampe, J. W., S. B. Stepaniants, et al. (2004). "Signatures of environmental exposures using peripheral leukocyte gene expression: tobacco smoke." Cancer Epidemiol Biomarkers Prev **13**(3): 445-453.

Lee, W. T., C. S. Cheung, et al. (2005). "Generalized low bone mass of girls with adolescent idiopathic scoliosis is related to inadequate calcium intake and weight bearing physical activity in peripubertal period." Osteoporos Int **16**(9): 1024-1035.

Lee, W. T., S. S. Leung, et al. (1995). "A randomized double-blind controlled calcium supplementation trial, and bone and height acquisition in children." Br J Nutr **74**(1): 125-139.

Lee, W. T., S. S. Leung, et al. (1997). "Bone mineral acquisition in low calcium intake children following the withdrawal of calcium supplement." Acta Paediatr **86**(6): 570-576.

Lee, W. T., S. S. Leung, et al. (1994). "Double-blind, controlled calcium supplementation and bone mineral accretion in children accustomed to a low-calcium diet." Am J Clin Nutr **60**(5): 744-750.

Letellier, K., B. Azeddine, et al. (2007). "[Etiopathogenesis of adolescent idiopathic scoliosis and new molecular concepts]." Med Sci (Paris) **23**(11): 910-916.

Letellier, K., B. Azeddine, et al. (2008). "Estrogen cross-talk with the melatonin signaling pathway in human osteoblasts derived from adolescent idiopathic scoliosis patients." J Pineal Res **45**(4): 383-393.

Levoye, A., R. Jockers, et al. (2006). "Are G protein-coupled receptor heterodimers of physiological relevance?--Focus on melatonin receptors." Chronobiol Int **23**(1-2): 419-426.

Lininger, Ruth A., et al. (2006). "Estrogen Receptor-Alpha (ER-Alpha) And Defects In Uterine Receptivity In Women." Reproductive Biology & Endocrinology **4**: S9-10.

Liu, Z., E. M. Tam, et al. (2012). "Abnormal leptin bioavailability in adolescent idiopathic scoliosis: an important new finding." Spine (Phila Pa 1976) **37**(7): 599-604.

Liu, Z., N. L. Tang, et al. (2010). "Lack of association between the promoter polymorphisms of MMP-3 and IL-6 genes and adolescent idiopathic scoliosis: a case-control study in a Chinese Han population." Spine (Phila Pa 1976) **35**(18): 1701-1705.

Lombardi, G., M. Y. Akoume, et al. (2011). "Biochemistry of adolescent idiopathic scoliosis." Adv Clin Chem **54**: 165-182.

Lonstein, J. E. and T. S. Renshaw (1987). "Neuromuscular spine deformities." Instr Course Lect **36**: 285-304.

Lowe, T., D. Lawellin, et al. (2002). "Platelet calmodulin levels in adolescent idiopathic scoliosis: do the levels correlate with curve progression and severity?" Spine (Phila Pa 1976) **27**(7): 768-775.

Lowe, T. G., M. Edgar, et al. (2000). "Etiology of idiopathic scoliosis: current trends in research." J Bone Joint Surg Am **82-A**(8): 1157-1168.

MacDonald, J. T. (2006). "Mutation in ROBO3 gene recently documented in patients with horizontal gaze palsy with progressive scoliosis." J Child Neurol **21**(6): 545.

MacGregor, A. J., H. Snieder, et al. (2000). "Twins. Novel uses to study complex traits and genetic diseases." Trends Genet **16**(3): 131-134.

Machida, M. (1999). "Cause of idiopathic scoliosis." Spine (Phila Pa 1976) **24**(24): 2576-2583.

Machida, M., J. Dubousset, et al. (1993). "An experimental study in chickens for the pathogenesis of idiopathic scoliosis." Spine (Phila Pa 1976) **18**(12): 1609-1615.

Machida, M., J. Dubousset, et al. (1995). "Role of melatonin deficiency in the development of scoliosis in pinealectomised chickens." J Bone Joint Surg Br **77**(1): 134-138.

Machida, M., J. Dubousset, et al. (1996). "Melatonin. A possible role in pathogenesis of adolescent idiopathic scoliosis." Spine (Phila Pa 1976) **21**(10): 1147-1152.

Machida, M., J. Dubousset, et al. (2006). "Experimental scoliosis in melatonin-deficient C57BL/6J mice without pinealectomy." J Pineal Res **41**(1): 1-7.

MacKenzie, R. S., M. A. Melan, et al. (2002). "Dual coupling of MT(1) and MT(2) melatonin receptors to cyclic AMP and phosphoinositide signal transduction cascades and their regulation following melatonin exposure." Biochem Pharmacol **63**(4): 587-595.

Mak, Y. T., G. Hampson, et al. (2004). "Variations in genome-wide gene expression in identical twins - a study of primary osteoblast-like culture from female twins discordant for osteoporosis." BMC Genet **5**: 14.

Man, G. C., J. H. Wong, et al. (2011). "Abnormal melatonin receptor 1B expression in osteoblasts from girls with adolescent idiopathic scoliosis." J Pineal Res **50**(4): 395-402.

Manzoni, D. and F. Miele (2002). "Vestibular mechanisms involved in idiopathic scoliosis." Arch Ital Biol **140**(1): 67-80.

Martinez-Llorens, J., M. Ramirez, et al. (2010). "Muscle dysfunction and exercise limitation in adolescent idiopathic scoliosis." Eur Respir J **36**(2): 393-400.

McMaster M, L. A., Burwell RG. (2006). "Physical activities of patients with adolescent idiopathic scoliosis (AIS) compared with a control group: implications for etiology and possible prevention" J Bone Joint Surg Br, **88-B(Supp II)**: 225.

Medina, I., G. Krapivinsky, S. Arnold, P. Kovoor, L. Krapivinsky and D. E. Clapham (2000). "A switch mechanism for G beta gamma activation of I(KACh)." J Biol Chem **275**(38): 29709-16.

Millard, S. M., A. M. Louie, et al. (2011). "Blockade of receptor-activated G(i) signaling in osteoblasts in vivo leads to site-specific increases in cortical and cancellous bone formation." J Bone Miner Res **26**(4): 822-832.

- Miller, N. H. (2002). "Genetics of familial idiopathic scoliosis." Clin Orthop Relat Res(401): 60-64.
- Miller, N. H. (2007). "Genetics of familial idiopathic scoliosis." Clin Orthop Relat Res **462**: 6-10.
- Miller, N. H., C. M. Justice, et al. (2005). "Identification of candidate regions for familial idiopathic scoliosis." Spine (Phila Pa 1976) **30**(10): 1181-1187.
- Miller, N. H., B. Mims, et al. (1996). "Genetic analysis of structural elastic fiber and collagen genes in familial adolescent idiopathic scoliosis." J Orthop Res **14**(6): 994-999.
- Miller, N. H., D. L. Schwab, et al. (2001). "Characterization of idiopathic scoliosis in a clinically well-defined population." Clin Orthop Relat Res(392): 349-357.
- Misol, S., I. V. Ponseti, et al. (1971). "Growth hormone blood levels in patients with idiopathic scoliosis." Clin Orthop Relat Res **81**: 122-125.
- Montanaro, L., P. Parisini, et al. (2006). "Evidence of a linkage between matrilin-1 gene (MATN1) and idiopathic scoliosis." Scoliosis **1**: 21.
- Moreau, A., M. Y. Akoume Ndong, et al. (2009). "[Molecular and genetic aspects of idiopathic scoliosis. Blood test for idiopathic scoliosis]." Orthopade **38**(2): 114-116, 118-121.
- Moreau, A., D. S. Wang, et al. (2004). "Melatonin signaling dysfunction in adolescent idiopathic scoliosis." Spine (Phila Pa 1976) **29**(16): 1772-1781.
- Nakade, O., H. Koyama, et al. (1999). "Melatonin stimulates proliferation and type I collagen synthesis in human bone cells in vitro." J Pineal Res **27**(2): 106-110.
- Nelissen, E. C., A. P. van Montfoort, et al. (2011). "Epigenetics and the placenta." Hum Reprod Update **17**(3): 397-417.
- Nissinen, M., M. Heliovaara, et al. (1993). "Trunk asymmetry and screening for scoliosis: a longitudinal cohort study of pubertal schoolchildren." Acta Paediatr **82**(1): 77-82.
- Nordwall, A. and S. Willner (1975). "A study of skeletal age and height in girls with idiopathic scoliosis." Clin Orthop Relat Res(110): 6-10.

Peng, J., M. Bencsik, et al. (2008). "Conditional expression of a Gi-coupled receptor in osteoblasts results in trabecular osteopenia." Endocrinology **149**(3): 1329-1337.

Perdriolle, R. and J. Vidal (1987). "Morphology of scoliosis: three-dimensional evolution." Orthopedics **10**(6): 909-915.

Petronis, A., Gottesman, II, et al. (2003). "Monozygotic twins exhibit numerous epigenetic differences: clues to twin discordance?" Schizophr Bull **29**(1): 169-178.

Pincott, J. R., J. S. Davies, et al. (1984). "Scoliosis caused by section of dorsal spinal nerve roots." J Bone Joint Surg Br **66**(1): 27-29.

Qiu, X. S., N. L. Tang, et al. (2007). "Melatonin receptor 1B (MTNR1B) gene polymorphism is associated with the occurrence of adolescent idiopathic scoliosis." Spine (Phila Pa 1976) **32**(16): 1748-1753.

Radio, N. M., J. S. Doctor, et al. (2006). "Melatonin enhances alkaline phosphatase activity in differentiating human adult mesenchymal stem cells grown in osteogenic medium via MT2 melatonin receptors and the MEK/ERK (1/2) signaling cascade." J Pineal Res **40**(4): 332-342.

Raggio, C. L., P. F. Giampietro, et al. (2009). "A novel locus for adolescent idiopathic scoliosis on chromosome 12p." J Orthop Res **27**(10): 1366-1372.

Reuber, M., A. Schultz, et al. (1983). "Trunk muscle myoelectric activities in idiopathic scoliosis." Spine (Phila Pa 1976) **8**(5): 447-456.

Robin, G. C. and T. Cohen (1975). "Familial scoliosis. A clinical report." J Bone Joint Surg Br **57**(2): 146-148.

Rogala, E. J., D. S. Drummond, et al. (1978). "Scoliosis: incidence and natural history. A prospective epidemiological study." J Bone Joint Surg Am **60**(2): 173-176.

Roth, J. A., B. G. Kim, et al. (1999). "Melatonin promotes osteoblast differentiation and bone formation." J Biol Chem **274**(31): 22041-22047.

Sadat-Ali, M., I. al-Habdan, et al. (2000). "Adolescent idiopathic scoliosis. Is low melatonin a cause?"

Joint Bone Spine **67**(1): 62-64.

Salehi, L. B., M. Mangino, et al. (2002). "Assignment of a locus for autosomal dominant idiopathic scoliosis (IS) to human chromosome 17p11." Hum Genet **111**(4-5): 401-404.

Sanders, J. O., R. H. Browne, et al. (2006). "Correlates of the peak height velocity in girls with idiopathic scoliosis." Spine (Phila Pa 1976) **31**(20): 2289-2295.

Shands, A. R., Jr. and H. B. Eisberg (1955). "The incidence of scoliosis in the state of Delaware; a study of 50,000 minifilms of the chest made during a survey for tuberculosis." J Bone Joint Surg Am **37-A**(6): 1243-1249.

Shapiro, J. R., V. E. Burn, et al. (1989). "Osteoporosis and familial idiopathic scoliosis: association with an abnormal alpha 2(I) collagen." Connect Tissue Res **21**(1-4): 117-123; discussion 124.

Sharma, S., X. Gao, et al. (2011). "Genome-wide association studies of adolescent idiopathic scoliosis suggest candidate susceptibility genes." Hum Mol Genet **20**(7): 1456-1466.

Skogland, L. B. and J. A. Miller (1980). "Growth related hormones in idiopathic scoliosis. An endocrine basis for accelerated growth." Acta Orthop Scand **51**(5): 779-780.

Smrcka, A. V. (2008) "G Protein [Beta][Gamma] Subunits: Central Mediators of G Protein-Coupled Receptor Signaling." Cellular and Molecular Life Sciences **65**(14): 2191-214.

Thillard, M. J. (1959). "[Vertebral column deformities following epiphysectomy in the chick]." C R Hebd Seances Acad Sci **248**(8): 1238-1240.

Thomas, K. A., S. D. Cook, et al. (1992). "Lumbar spine and femoral neck bone mineral density in idiopathic scoliosis: a follow-up study." J Pediatr Orthop **12**(2): 235-240.

Trobisch, P., O. Suess, et al. (2010). "Idiopathic scoliosis." Dtsch Arztebl Int **107**(49): 875-883; quiz 884.

United States Bone and Joint Initiative: *The Burden of Musculoskeletal Diseases in the United States*, Second Edition. Rosemont, IL: American Academy of Orthopaedic Surgeons; 2011.

Verdonk, E., K. Johnson, et al. (2006). "Cellular dielectric spectroscopy: a label-free comprehensive platform for functional evaluation of endogenous receptors." Assay Drug Dev Technol **4**(5): 609-619.

Von Gall, Jö Stehle, and David R. Weaver. "Mammalian Melatonin Receptors: Molecular Biology and Signal Transduction." Cell and tissue research **309**.1 (2002): 151-62.

Wang, H., Z. Wu, et al. (2008). "Association study of tryptophan hydroxylase 1 and arylalkylamine N-acetyltransferase polymorphisms with adolescent idiopathic scoliosis in Han Chinese." Spine (Phila Pa 1976) **33**(20): 2199-2203.

Wang, J., J. Zhou, et al. (1999). "Igf1 promotes longitudinal bone growth by insulin-like actions augmenting chondrocyte hypertrophy." FASEB J **13**(14): 1985-1990.

Wang, W. J., H. Y. Yeung, et al. (2011). "Top theories for the etiopathogenesis of adolescent idiopathic scoliosis." J Pediatr Orthop **31**(1 Suppl): S14-27.

Wang, X., H. Jiang, et al. (1997). "Characterization of the scoliosis that develops after pinealectomy in the chicken and comparison with adolescent idiopathic scoliosis in humans." Spine (Phila Pa 1976) **22**(22): 2626-2635.

Warner, W. (2001). "Juvenile idiopathic scoliosis." In The Pediatric Spine, 2 edition, Stewart L. Weinstein, ed. Lippincott Williams & Wilkins, Philadelphia, PA, USA.

Weinstein, S. L., L. A. Dolan, et al. (2008). "Adolescent idiopathic scoliosis." Lancet **371**(9623): 1527-1537.

Whitney, A. R., M. Diehn, et al. (2003). "Individuality and variation in gene expression patterns in human blood." Proc Natl Acad Sci U S A **100**(4): 1896-1901.

Willner, S. (1982). "A comparative study of the efficiency of different types of school screening for scoliosis." Acta Orthop Scand **53**(5): 769-774.

Wilson, S. B., S. C. Kent, et al. (2000). "Multiple differences in gene expression in regulatory Valpha 24Jalpha Q T cells from identical twins discordant for type I diabetes." Proc Natl Acad Sci U S A **97**(13): 7411-7416.

Wise, C. A., R. Barnes, et al. (2000). "Localization of susceptibility to familial idiopathic scoliosis." Spine (Phila Pa 1976) **25**(18): 2372-2380.

Wise, C. A., X. Gao, et al. (2008). "Understanding genetic factors in idiopathic scoliosis, a complex disease of childhood." Curr Genomics **9**(1): 51-59.

Witt-Enderby, P. A., J. Bennett, et al. (2003). "Melatonin receptors and their regulation: biochemical and structural mechanisms." Life Sci **72**(20): 2183-2198.

Worthington, V. and P. Shambaugh (1993). "Nutrition as an environmental factor in the etiology of idiopathic scoliosis." J Manipulative Physiol Ther **16**(3): 169-173.

Wynne-Davies, R. (1968). "Familial (idiopathic) scoliosis. A family survey." J Bone Joint Surg Br **50**(1): 24-30.

Xi, G., L. A. Maile, et al. (2008). "Expression of the human beta3 integrin subunit in mouse smooth muscle cells enhances IGF-I-stimulated signaling and proliferation." J Cell Physiol **214**(2): 306-315.

Yang, Y., Z. Wu, et al. (2009). "Adolescent idiopathic scoliosis and the single-nucleotide polymorphism of the growth hormone receptor and IGF-1 genes." Orthopedics **32**(6): 411.

Yarom, R., S. Meyer, et al. (1982). "Metal impregnation abnormalities in platelets of patients with idiopathic scoliosis." Haemostasis **12**(3): 282-288.

Yeung, H. Y., N. L. Tang, et al. (2006). "Genetic association study of insulin-like growth factor-I (IGF-I) gene with curve severity and osteopenia in adolescent idiopathic scoliosis." Stud Health Technol Inform **123**: 18-24.

Yoshimoto, N., T. Nishiyama, et al. (2011). "Genetic and environmental predictors, endogenous hormones and growth factors, and risk of estrogen receptor-positive breast cancer in Japanese women." Cancer Sci **102**(11): 2065-2072.

Zhang, H. Q., S. J. Lu, et al. (2009). "Association of estrogen receptor beta gene polymorphisms with susceptibility to adolescent idiopathic scoliosis." Spine (Phila Pa 1976) **34**(8): 760-764.

Zivanovic, S., L. Petrovic-Rackov, et al. (2010). "[Detection of bone erosions in knee osteoarthritis by serum biomarkers]." Srp Arh Celok Lek **138**(1-2): 62-66.

Zorab, P. A., S. Clark, et al. (1971). "Bone collagen turnover in idiopathic scoliosis estimated from total hydroxyproline excretion." Arch Dis Child **46**(250): 828-832.