# Whole Genome Sequencing in Adolescent Idiopathic Scoliosis Cohort Indicates Polygenic Disease Involving Multiple Biological Pathways

Islam Oguz Tuncay[1], Eun Kyoung Lee[1,2], Yoonsuh Lee[1], June-Young Koh[1], Wonchul Lee[1], Anxhela Gjyshi Gustafson[3], Sangmoon Lee[1], Kamran Shazand[3]

[1]Inocras, San Diego, CA, USA, [2]Department of Ophthalmology, Seoul National University Hospital, Seoul, Republic of Korea, [3]Shriners Genomics Institute, Tampa, FL, USA

## Introduction

- Adolescent idiopathic scoliosis (AIS) is the most common nondegenerative spinal abnormality with a prevalence of 1−4%.
- Research suggests that AIS is a highly complex polygenic disease that results from the interaction of multiple gene loci and the environment.
- Here we utilized whole-genome sequencing (WGS) to comprehensively explore the genetic landscape of 119 AIS patients from 103 families. We implemented an automated WGS analysis pipeline powered by RareVision™.

## Methods

- This is a retrospective analysis of patients with AIS at the Research for Precision Medicine study at Shriners Hospitals for Children – The Genome Institute.
- Genome sequencing was performed by the Shriners Genomics Institute using the Illumina NovaSeq6000 platform with an average depth of coverage of 30X. Sequences analyzed by Inocras using RareVision™ for mapping, filtration, variant calling, and annotation.
- Candidate variants were prioritized by considering previously reported associations between the gene and AIS, predicted effect of the mutation, and familial segregation.

## Results

- The cohort comprised 204 samples from 103 families (16 singletons, 73 duos, and 14 trios), including 119 AIS patients (103 probands and 16 affected mothers). Among the 103 probands, 81 (78.6%) were female and 22 (21.4%) were male.
- PCA (figure below) showed that the majority of participants had European and/or Admixed American ancestry.

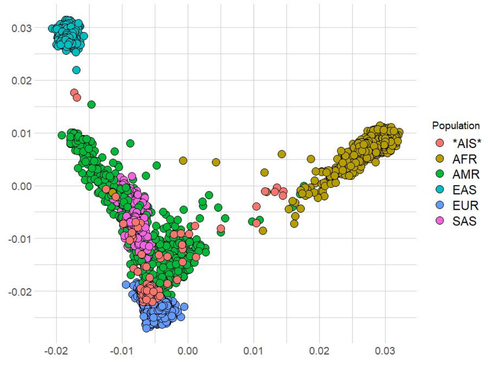


**Figure 1.** PCA of the AIS cohort with 1000 Genomes Project populations. First two principal components are plotted.

**Table 1.** A selection of strong candidate variants identified in the AIS cohort.

| Patient | Gene/ Locus | Variant Information | OMIM Phenotype |
|---|---|---|---|
| 14-0105 | NSD2 | Frameshift (p.G1344fs) | Rauch-Steindl syndrome (#619695) |
| 14-0356 | 22q11.2 | Deletion (LCR22-A/D) | DiGeorge syndrome (#188400) |
| 22-0066 | SPRED1 | Missense (p.R117L) | Legius syndrome (#611431) |
| 18-0033 | GJB2 | Frameshift (p.G12Vfs) | Keratitis-ichthyosis-deafness (#148210) |
| 15-0137 | SMARCA4 | Splicing (c.4912-2A>C) | RTPS2 (#613325) |
| 15-0107 | KIF7 | Stop gain (p.R21X) | Hydrolethalus 2 (#614120) |
| 15-0085 | HSPG2 | Frameshift (p.S2669fs) | Schwartz-Jampel type 1 (#255800) |
| 14-0387 | GJB2 | Frameshift (p.L56fs) | Keratitis-ichthyosis-deafness (#148210) |
| 14-0339 | SPATA5 | Deletion (c.1812_1714del) | NEDHSB (#616577) |
| 14-0205 | FAM111B | Stop gain (p.N286X) | POIKTMP (#615704) |
| 14-0123 | FBN2 | Splicing (c.7712-11T>A) | Contractural arachnodactyly (#121050) |
| 14-0074 | IFIH1 | Deletion (c.2363_1096del) | Singleton-Merten 1 (#182250) |
| 07-0044 | CUL7 | Missense (p.L1588P) | 3-M syndrome 1 (#273750) |
| 07-0021 | ARID1B | Frameshift (p.G434fs) | Coffin-Siris syndrome 1 (#135900) |
| 02-0074 | TNPO3 | Stop gain (p.R571X) | LGMDD2 (#608423) |

- We identified 93 genetic variants likely or highly likely to account for disease presentation in 94 patients (91.3%).
- Candidate genes included previously AIS-associated genes (e.g. *CHD7*, *COL11A1/2*, *FBN1/2*, *HSPG2*, *KIF7*), as well as genes associated with other syndromes where scoliosis is a known phenotype (e.g. *RYR1*, *GJB2*, *MYH2*, *MYH7*).
- Gene set enrichment analysis identified GO terms "extracellular matrix organization" ($p=1.12\times10^{-8}$) and "muscle contraction" ($p=2.28\times10^{-8}$), and KEGG term "ECM-receptor interaction" ($p=3.02\times10^{-4}$).

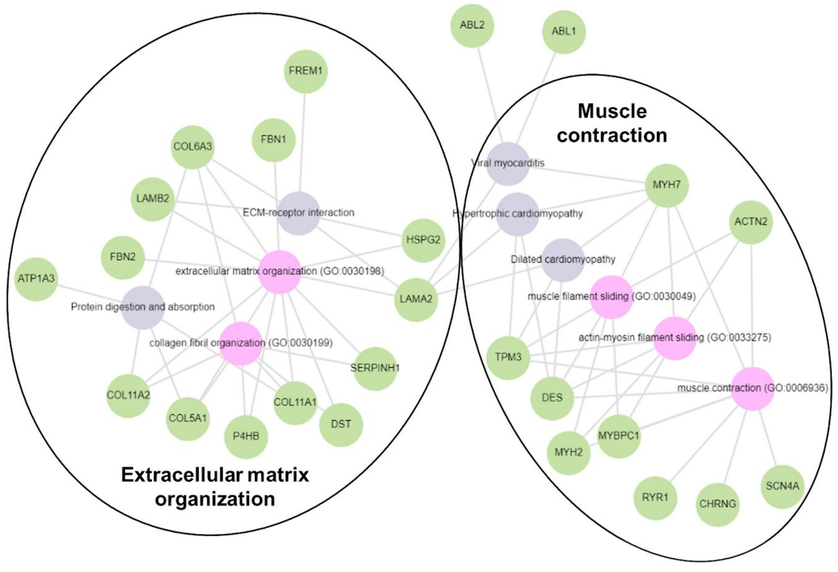


**Figure 2.** Visual representation of gene set enrichment analysis, performed using Enrichr.

- Candidate variants included coding and noncoding point mutations, and structural variants, showing the utility of WGS.

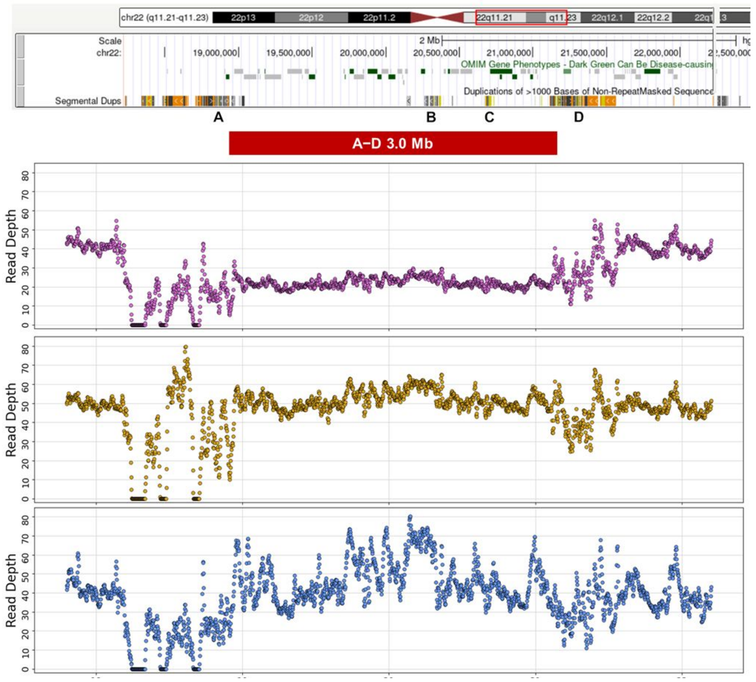


**Figure 3.** *De novo* deletion of 22q11.2 LCR22-A/D identified in patient 14-0356. Top: proband, middle: mother, bottom: father.

- We performed allelic tests of association for 15 SNPs previously associated with AIS. To improve the power of the analysis, we only included patients with European and/or Admixed American ancestry.
- We identified significant associations for 4 SNPs; rs12946942 ($p=3.68\times10^{-4}$), located between *SOX9* and *KCNJ2*, rs10756785 ($p=7.56\times10^{-4}$) and rs3904778 ($p=1.66\times10^{-3}$) of *BNC2*, and rs7633294 ($p=2.22\times10^{-3}$) of *MAGI1*.

## Conclusions

- To the best of our knowledge, this is the largest study to explore the genomic landscape of an AIS cohort using WGS.
- Most candidate genes were associated with other complex syndromes indicating that some AIS cases may be a milder presentation of more severe syndromes.
- Our genetic findings support previous research which suggest that genes involved in skeletal muscle contraction and ECM organization may play an important role in the development of the abnormal spinal curvature.
- This study revealed that the genetic etiology of AIS is highly polygenic and that WGS has a greater potential to detect a wider spectrum of genetic variants.

IP84-0924