Idiopathic scoliosis (IS) is a spine deformity that affects approximately 3% of the population. The underlying causes of IS are not well understood, although there is clear evidence that there is a genetic component to the disease. Genetic mapping studies suggest high genetic heterogeneity, but no IS disease-causing gene has yet been identified. Here, genetic linkage analyses combined with exome sequencing identified a rare missense variant (p.A446T) in the centriolar protein gene POCS that cosegregated with the disease in a large family with multiple members affected with IS. Subsequently, the p.A446T variant was found in an additional set of families with IS and in an additional 3 cases of IS. Moreover, POCS variant p.A455P was present and linked to IS in one family and another rare POCS variant (p.A429V) was identified in an additional 5 cases of IS. In a zebrafish model, expression of any of the 3 human IS-associated POCS variant mRNAs resulted in spine deformity, without affecting other skeletal structures. Together, these findings indicate that mutations in the POCS gene contribute to the occurrence of IS.
Functional variants of POC5 identified in patients with idiopathic scoliosis

Shunmoogam A. Patten,1,2 Patricia Margaritte-Jeannin,3 Jean-Claude Bernard,4 Eudeline Alix,5 Audrey Labalme,5 Alicia Besson,6 Simon L. Girard,7 Khaled Fendri,8 Nicolas Fraisse,9 Bernard Biot,10 Coline Poizat,5 Amandine Campan-Fournier,6 Karim Anbel-Genovesio,5 Vincent Cunin,8 Charlotte Zaouter,9 Meijiang Liao,2 Raphaelle Lamy,9 Gaetan Lesca,5,6 Rita Menassa,10 Charles Marcaillou,11 Melanie Letexier,11 Damien Sanlaville,5,6 Jerome Berard,4 Guy A. Rouleau,7 Françoise Clerget-Darpoux,12 Pierre Drapeau,2 Florina Moldovan,1 and Patrick Edery5,6

1CHU Sainte Justine Research Center and Department of Stomatology, Faculty of Dentistry, and 2Department of Neuroscience, Faculty of Medicine and CRCHUM, University of Montréal, Montréal, Québec, Canada. 3INSERM, UMR-946, Université Paris Diderot, Institut Universitaire d’Hématologie, Paris, France. 4Croix-Rouge Française, CMCR des Massues, Lyon, France. 5Genetics Service, Hospices Civils de Lyon, Lyon, France. 6INSERM U1028, CNRS UMR5292, University Lyon 1, Lyon Neuroscience Research Center, TIGER Team, Lyon, France. 7Montreal Neurological Institute and Hospital, McGill University, Montreal, Québec, Canada. 8Service d’Orthopédie Pédiatrique, 9Service de Génétique Moléculaire et Clinique U21437, and 10Laboratoire d’Endocrinologie Moléculaire et Maladies Rares, Hospices Civils de Lyon, Lyon, France. 11IntegraGen, F-91030, Evry, France. 12Department of Genetics, Unité INSERM U-781, IJU Imagine et Université Paris Descartes Hôpital Necker-Enfants Malades, Paris, France.

The Journal of Clinical Investigation

Brief report

Idiopathic scoliosis (IS) is a spine deformity that affects approximately 3% of the population. The underlying causes of IS are not well understood, although there is clear evidence that there is a genetic component to the disease. Genetic mapping studies suggest high genetic heterogeneity, but no IS disease-causing gene has yet been identified. Here, genetic linkage analyses combined with exome sequencing identified a rare missense variant (p.A446T) in the centriloar protein gene POC5 that cosegregated with the disease in a large family with multiple members affected with IS. Subsequently, the p.A446T variant was found in an additional set of families with IS and in an additional 3 cases of IS. Moreover, POC5 variant p.A455P was present and linked to IS in one family and another rare POC5 variant (p.A429V) was identified in an additional 5 cases of IS. In a zebrafish model, expression of any of the 3 human IS-associated POC5 variant mRNAs resulted in spine deformity, without affecting other skeletal structures. Together, these findings indicate that mutations in the POC5 gene contribute to the occurrence of IS.

Introduction

Idiopathic scoliosis (IS) is a 3D spinal deformity characterized by lateral curvature of the spine of 10° or more, as measured by Cobb’s method, and rotation of the vertebral bodies. The etiology of IS remains unknown, but several studies strongly suggest a contribution from genetic factors (1, 2). Familial aggregation of IS cases was first observed many decades ago (3–5), and the disease is familial in approximately 40% of cases (6, 7). In some of these cases, transmission of the trait appears to be autosomal dominant. However, the results of genome-wide linkage studies in multiplex families were inconclusive, suggesting that IS is genetically heterogeneous (8–10). GWAS in case/control samples were recently used to identify candidate loci for IS susceptibility (11–13); however, functional variants unambiguously causing IS have not yet been identified. In the present study, we sought to refine the search for IS susceptibility loci and to identify the disease-causing gene in a large family with a high prevalence of IS.

Results and Discussion

Using the affected-only method in a large multiplex family with 11 affected individuals and 1 obligate carrier (family F2), we previously mapped a causative IS gene to either chromosome 5q13.3 or 3q12.3 (8). To identify the IS disease-causing gene, we refined the candidate regions for family F2 to chromosome 5q13.3: 73,905,694–79,488,191 (NCBI hg19; Supplemental Figure 1; supplemental material available online with this article; doi:10.1172/JCI77262DS1) and chromosome 3q12.3: 95,083,093–107,153,338 (NCBI hg19; data not shown) and performed whole-exome sequencing on 3 affected individuals from this family.

This provided a sequencing depth per target base of 134 times on average, with at least 10 times for 98.6% of bases (Supplemental Figure 2 and Supplemental Table 1). Variant calling revealed 825 changes (599 single nucleotide variants [SNVs] and 226 indels) in the 3q12.3 and 5q13.3 critical chromosomal regions for these 3 subjects. To identify potentially pathogenic variants, we first excluded synonymous variants or variants located in introns, except for those affecting consensus splice sites. We then parsed a total of 172 variants (Supplemental Table 2) to keep only those that were novel or rare (minor allele frequency [MAF] <5%), which yielded 2 candidate SNVs, one in GPR128 and the other in POC5 (Table 1). We excluded the GPR128 SNV because it did not cosegregate with IS in family F2. The other SNV, c.G1336A, is located in the POC5 gene (NM_001099271; Supplemental Table 3), where it results in a single amino acid change, p.A446T. The presence of this SNP was then confirmed by Sanger sequencing in all affected members of family F2 (Figure 1). The c.G1336A POC5 SNP was also observed in 3 probands from 40 additional multiplex IS families and cosegregated with
Exome sequencing identifies a rare missense SNV in POC5

POC5 is a gene involved in the development of zebrafish. To study the role of POC5 in human IS, researchers compared the sequence of POC5 in IS patients to that of unaffected individuals. They found two rare missense SNVs in POC5: c.G1336A (p.A446T) and c.C1286T (p.A429V). These SNVs were found more frequently in IS probands compared to controls. A previous study showed that the c.G1336A (p.A446T) SNV was found in all IS patients of a study conducted in Europe with similar ancestry (15, 16). The researchers also observed that other rare SNVs were present in IS cases but not in control individuals. They hypothesized that these SNVs might be linked to IS.

To investigate the role of these SNVs in IS, the researchers performed experiments with zebrafish. They found that the c.G1336A (p.A446T) SNV caused a curly-tailed phenotype, while the c.C1286T (p.A429V) SNV did not. They also observed that the c.G1336A (p.A446T) SNV was more frequent in IS families compared to controls. These findings suggest a potential role for these SNVs in the development of IS.

The researchers concluded that the c.G1336A (p.A446T) and c.C1286T (p.A429V) SNVs in POC5 may contribute to the development of IS. Further studies are needed to confirm these findings and to understand the mechanism by which these SNVs affect IS.

Table 1: Criteria used for filtering whole-exome sequencing SNVs and indels

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total variants in all 3 affected individuals combined</td>
<td>272,492</td>
</tr>
<tr>
<td>Variants after base quality filtering</td>
<td>237,947</td>
</tr>
<tr>
<td>Variants in 3q12.3 and 5q13.3 linkage regions</td>
<td>825 (599 SNVs + 226 indels)</td>
</tr>
</tbody>
</table>

Figure 1: Identification of a rare POC5 SNV in family F2. Exome sequencing identifies a rare missense SNV in POC5, c.1336G>A, listed in dbSNP138 (rs34678567). Sample chromatograms for a healthy individual without this SNV (upper left panel) and for an affected individual with this SNV (lower left panel) are shown. The pedigree of family F2 is shown. Cobb’s angle and genotype at the mutated position are indicated below symbols corresponding to individuals. U, uncertain status; ND, status not determined.
observed in IS patients. Importantly, no other skeletal or body malformations were observed, including misshapen, missing, or fused vertebrae. The severity of larval phenotypes upon overexpression of mutated \( POC5 \) resulted in spine defects of similar magnitude in juvenile fish.

\((P < 0.001; \text{Figure 3A})\). When mut-\( POC5 \) fish reached juvenile stages (50–60 days post-fertilization [dpf]), striking curvature of the fully mineralized vertebral column was observed (Figure 3B), in some instances combined with vertebral rotation (Figure 3B and Supplemental Video 1) reminiscent of the 3D deformation observed in IS patients. Importantly, no other skeletal or body malformations were observed, including misshapen, missing, or fused vertebrae. The severity of larval phenotypes upon overexpression of mutated \( POC5 \) resulted in spine defects of similar magnitude in juvenile fish.

**Figure 2. IS families and cases with exon 10 POC5 SNVs.** Forty additional IS families and 150 IS cases were studied. (A) IS families F19, F35, and F41 and IS cases C39, C58, and C83 show the same c.1336G>A \( POC5 \) rare missense SNV listed in dbSNP138 (rs34678567). Sample chromatogram for an IS patient with this SNV is shown. The position of the SNV is indicated by an arrow above the chromatogram. Pedigrees of families and cases are shown. (B) IS family F31 harboring the c.1363G>C \( POC5 \) novel missense SNV and corresponding chromatogram. (C) IS cases C1, C77, C137, C143, and C150 harboring the c.1286C>T \( POC5 \) rare missense SNV listed in dbSNP138 (rs146984380) and corresponding sequence chromatograms. Cobb’s angle and genotype at the mutated position are indicated below symbols corresponding to individuals.
Although phenotypic information regarding IS is not available for individuals in the control cohort of French descent, the penetrance of the disease associated with these SNVs in humans appears to be low for c.G1336A (p.A446T) and c.C1286T (p.A429V), which are present in this cohort at a frequency of 0.75% and 0.34%, respectively. This makes it difficult to explain the high number of affected individuals in families F2 and F19 among carriers of the c.G1336A (p.A446T) SNV (16/20, Figure 1 and Figure 2A). However, a risk-modifying allele, which amplifies disease expression, could also cosegregate in these families. Such a modifier allele may be located on 3q12.3 in family F2, where the same chromosome 3q12.3 haplotype segregates in all 11 affected patients, but not in the unaffected c.G1336A (p.A446T) POC5 SNV carrier. We are currently attempting to identify this modifier allele in family F2. However, in IS family F19, the putative risk-modifying allele is not located on chromosome 3q12.3, because affected members of this family do not share a common haplotype in the critical 3q12.3 IS interval (data not shown). These data suggest that the modifier allele may be different in the different families where the c.G1336A (p.A446T) POC5 SNV segregates. If we assume that at least 2 genes are required for disease expression, the combination of 2 rare variants with low marginal penetrance could confer high disease penetrance in carriers. This type of model could explain both the phenotypic and genetic observations in families F2 and F19 and why large multiplex IS families are so rare, as the disease recurrence risk in the sibship of a patient would be less than or equal to one-fourth.

To define the spatiotemporal expression of Poc5 in developing embryos and larvae, we performed whole-mount in situ hybridization (Supplemental Figure 10). Poc5 was expressed ubiquitously during early somitogenesis. Its expression became restricted to the head and bud region by 24 hpf. By 48 hpf and 72 hpf, its expression became even more confined, to the telencephalon, midbrain, and midbrain-hindbrain boundary. At 72 hpf, expression was particularly strong at the midbrain-hindbrain boundary, a crucial center organizing brain patterning (17). Given that the poc5 gene is ubiquitously expressed during somitogenesis and that injected embryos show defects in the formation of axial structures, poc5 could thus play an important role in early aspects of anterior-posterior axis development. These axial defects likely persist and later affect spine formation.

Very little is currently known about the role of POC5. In humans, POC5 localizes to the distal portion of centrioles and is recruited to procentrioles for full centriolar maturation and normal cell-cycle processing (14). This centrosomal protein interacts with centrin (18) and inversin, both involved in cell division, polarity, and motility. Thus, the function of primary cilia and left-right axis determination may be somehow impaired in patients and connected to IS. Interestingly, the 3 functional POC5 variants result in the switch of an alanine residue (alanine to tyrosine, valine, or proline) and are located in the same region of the gene (NM_001099271, exon 10), suggesting that this region may be functionally significant.

Although phenotypic information regarding IS is not available for individuals in the control cohort of French descent, the penetrance of the disease associated with these SNVs in humans appears to be low for c.G1336A (p.A446T) and c.C1286T (p.A429V), which are present in this cohort at a frequency of 0.75% and 0.34%, respectively. This makes it difficult to explain the high number of affected individuals in families F2 and F19 among carriers of the c.G1336A (p.A446T) SNV (16/20, Figure 1 and Figure 2A). However, a risk-modifying allele, which amplifies disease expression, could also cosegregate in these families. Such a modifier allele may be located on 3q12.3 in family F2, where the same chromosome 3q12.3 haplotype segregates in all 11 affected patients, but not in the unaffected c.G1336A (p.A446T) POC5 SNV carrier. We are currently attempting to identify this modifier allele in family F2. However, in IS family F19, the putative risk-modifying allele is not located on chromosome 3q12.3, because affected members of this family do not share a common haplotype in the critical 3q12.3 IS interval (data not shown). These data suggest that the modifier allele may be different in the different families where the c.G1336A (p.A446T) POC5 SNV segregates. If we assume that at least 2 genes are required for disease expression, the combination of 2 rare variants with low marginal penetrance could confer high disease penetrance in carriers. This type of model could explain both the phenotypic and genetic observations in families F2 and F19 and why large multiplex IS families are so rare, as the disease recurrence risk in the sibship of a patient would be less than or equal to one-fourth.
Finally, in the context of this highly heterogeneous disorder with possible digenic inheritance for a subgroup of patients, identifying POC5 as what we believe to be the first IS-causing gene is a major step toward deciphering the genetic causes of IS. This crucial step will pave the way for future studies to identify additional genes or pathways involved in IS. Further studies of large cohorts of various ethnicities are needed to determine the contribution of POC5 genetic variants to IS in other populations and to establish the prevalence of these variants in the general population. In the future, drawing a complete picture of the genetic events leading to IS will help with devising preventive strategies and, it is hoped, therapies in this complex disorder.

Study approval. Written informed consent was received from all patients prior to enrollment in the study. Collection and use of patient samples for this study were approved by the ethics committee of the Hospices Civils de Lyon and CHU Sainte-Justine. All zebrafish work was carried out according to Canadian Council on Animal Care (CCAC) guidelines and approved by the local animal care committee.

Further details are available in the Supplemental Methods.

Acknowledgments

We express our heartfelt thanks to the patients and families who participated in this study. We thank Faiza Senni and Facilah Gence for establishing the IS cohort and Abdelhamid Bekri for running the automated in situ hybridization system provided by Flogentec. We are grateful to Jean Dubouset for enthusiastic and interesting discussions. We thank M. Gallagher-Gambarelli for her editorial assistance and Alexandre Dionne-Laporte for his bioinformatics assistance. S.A. Patten was supported by CHU Sainte-Justine, Foundation of Stars, and Canadian Institutes of Health Research (CIHR) fellowships. This work was supported by the Fondation Yves Cotrel-Institut de France (by grants to F. Moldovan, P. Edery, and S.A. Patten); the Faculté de Médecine Dentaire, the Université de Montréal, Fonds Ernest Charron and Réseau de Recherche en Santé Buccodentaire et Osseuse (RSBO) (by grants to F. Moldovan); CIHR (by grants to P. Drapeau and F. Moldovan); the French Ministry of Health (Plan Maladies Rares 2004), the Fondation Maladies Rares, the Hospices Civils de Lyon; Inserm; and CNRS and Lyon 1 University (by grants to P. Edery).

Address correspondence to: Florina Moldovan, CHU Sainte-Justine, 3175 Côte Sainte-Catherine 4th floor, Room 4731, Montréal, Québec H3T 1C5, Canada. Phone: 15143454931, ext. 5746; E-mail: florina.moldovan@umontreal.ca. Or to: Patrick Edery, Hospices Civils de Lyon, HFME 6th floor, 59 Boulevard Pinel, Bron Cedex 69677, France. Phone: 33472129698; E-mail: patrick.edery@chu-lyon.fr.