CeGaT

CeGaT GmbH | Paul-Ehrlich-Str. 23 | D-72076 Tübingen | Germany

Dr. Jane Doe
Paul-Ehrlich-Str. 23
72076 TÜBINGEN
GERMANY

Patient	XXX, XX (*DD.MM.YYYY)
Sex	Male
Patient-ID	#
Sample receipt	xxx
Material	DNA
Report date	xxx

Genetic analysis report – XXX, XX (*DD.MM.YYYY)

Indication	Suspicion of spondyloepiphyseal dysplasia; significant spine shortening, limb shortening, height: -6 SD; head circumference and weight preserved
Order	Panel Diagnostics: Spondylometaphyseal dysplasia and Spondylo-epi-(meta)-physeal dysplasia (whole exome enrichment)

Dear Dr. Doe,

Thank you for your request for molecular genetic analysis.

RESULTS

- **Detection of a pathogenic variant in gene *COL2A1*, which is causative for spondyloepi-physeal dysplasia in your patient.**
- Based upon current scientific knowledge, we did not identify any reportable copy number variants which are likely to be causative for your patient's disease.

Gene	Variant	Zygosity	Heredity	MAF (%)	*in silico* Prediction	Classification
COL2A1	**c.2006G>C; p.Gly669Ala**	**het.**	**AD**	**-**	**-**	**pathogenic**

Information for table interpretation:
AD: Describes a trait or disorder requiring only one copy of a gene variant at a particular locus in order to express an observable phenotype. Single heterozygous gene variants can only be causative for the phenotype if the disorder follows the **autosomal dominant** mode of inheritance.
AR: Describes a trait or disorder requiring the presence of two copies of a gene variant at a particular locus in order to express an observable phenotype. In **autosomal recessive** disorders, a single heterozygous variant in one gene cannot by itself be causative for the observable phenotype.
XL: X-linked mode of inheritance
mitochondrial: gene encoded in the mitochondrial DNA. Information on degree of heteroplasmy (heteropl. or homopl.), frequency and read count (variant/total reads).
MAF: The **minor allele frequency** describes the least frequent allele at a specific locus in a given population. The less frequently a variant occurs, the more likely it is pathogenic; however, the prevalence and mode of inheritance of any particular disease must be considered. For mitochondrial variants, the population frequency (MAF column) is based on the homoplasmic frequency within a reference population (gnomAD).
***in silico* Prediction:** The ACMG (American College of Medical Genetics) guidelines recommend using prediction programs to assess the possible pathogenicity of a variant. Each program calculates its predictions based upon different criteria, and the correspondence between a prediction and the actual functional effect of a variant is variable. **These predictions may therefore not serve as the sole basis for the evaluation of pathogenicity.**

Classification: Refers to the possible pathogenicity of a variant, but does not necessarily provide clear evidence of clinical significance. Variants are evaluated based upon current data and specific criteria, and are then classified as follows: pathogenic, likely pathogenic, and uncertain significance. **All variants for which clinical relevance cannot be conclusively confirmed or excluded are referred as variants of unknown clinical significance.**

INTERPRETATION

COL2A1, c.2006G>C; p.Gly669Ala (het.), NM_001844.5

OMIM / Reference	Phenotype	Heredity
108300	Stickler syndrome, type I	AD
609508	Stickler syndrome, type I, nonsyndromic ocular	AD
132450	Epiphyseal dysplasia, multiple, with myopia and deafness	AD
200610	Achondrogenesis type II or hypochondrogenesis	AD
608805	Avascular necrosis of the femoral head	AD
609162	Czech dysplasia	AD
156550	Kniest dysplasia	AD
150600	Legg-Calve-Perthes disease	AD
604864	Osteoarthritis with mild chondrodysplasia	AD
151210	Platyspondylic skeletal dysplasia Torrance type	AD
183900	**Spondyloepiphyseal dysplasia (SED) congenita**	**AD**
184250	Spondylometaepiphyseal dysplasia (SMED) Strudwick type	AD
616583	Spondyloepiphyseal dysplasia Stanescu type	AD
271700	Spondyloperipheral dysplasia	AD
ORPHA:209867	Vitreoretinopathy	AD

The gene ***COL2A1*** encodes the alpha-1 chain of type II collagen, which is the major collagen synthesized by chondrocytes but is also present in the vitreous body (GeneCards). The majority of pathogenic variants in this gene are nonsense and frameshift (mostly associated with Stickler syndrome type I), splice site changes (mainly associated with Kniest dysplasia) or missense variations affecting the glycine residues within the triple-helical region (collagen type II disorders; GeneReviews “Type II Collagen Disorders Overview”, Gregersen & Savarirayan, 2019, PMID: 31021589). However, there is no exact and distinctive genotype-phenotype correlation. The *COL2A1* gene is associated with a broad spectrum of phenotypes, including milder types of skeletal dysplasias (e.g. multiple epiphyseal dysplasia or SED congenita) as well as lethal types (e.g. achondrogenesis or hypochondrogenesis). The penetrance of *COL2A1* pathogenic variants in Stickler syndrome is assumed to be complete (GeneReviews "Stickler Syndrome", Robin et al., 2021; PMID: 20301479).

In your patient we detect the unknown missense variant **c.2006G>C; p.Gly669Ala** in the gene *COL2A1* in a heterozygous conformation. This alteration results in a glycine residue change in the triple-helical region of the gene, an established cause of disease in *COL2A1*-related disorders (Barat-Houari et al., 2016, PMID: 26626311). The Gly-Xaa-Yaa backbone of the protein is critical for folding and stability (Barat-Houari et al.,

2016; PMID: 26626311; Deng et al., 2016, PMID: 27234559). ClinVar contains another entry (Variation ID 873513) for this amino acid residue in which a different missense change (c.2005G>A; p.Gly669Ser) has been determined to be likely pathogenic. It has not been detected in population controls (gnomAD).

The identified pathogenic variant in gene *COL2A1* confirms the diagnosis of spondyloepiphyseal dysplasia in your patient.

Additional variants may be present within regions which were not analyzed (e.g. introns, promoter and enhancer regions and long repeats). A lower specificity enrichment and/or inaccurate variant calling cannot be ruled out for homologous regions with multiple genomic copies. The occurrence of low frequency somatic mosaicism cannot be reliably assessed using a pipeline optimized for germline variant detection, and may therefore remain undetected. Moreover, detection of large deletions and duplications is not guaranteed by next-generation high-throughput sequencing. Although unlikely, it is also possible that the classification of variants may change in the future due to improvements in scientific understanding.

GENETIC RELEVANCE

Your patient is heterozygous for a pathogenic variant in gene *COL2A1*. This may be of relevance for family planning and at-risk family members.

Individual variants have a 50% probability of being passed on to each respective offspring.

RECOMMENDATION

To determine whether the detected variant in gene *COL2A1* is *de novo* in your patient, testing of both parents regarding this variant is recommended.

We recommend further clinical evaluation, management, and treatment according to the current guidelines for Type II Collagen Disorders (Gregersen and Savarirayan, 2019, PMID: 31021589, GeneReviews).

According to § 10 of the German Genetic Diagnostics Legislation, appropriate genetic counseling should be offered with all diagnostic genetic testing. Following the identification of the molecular cause of a genetic disease, genetic counseling has to be offered.

Medical report written by: Scientist 1

Proofread by: Scientist 2

With kind regards,

Dr. med. Dr. rer. nat. Saskia Biskup
Dr. med. Friedmar Kreuz, M.A.

Consultant for Human Genetics

Dr. rer. nat. Heinz-Dieter Gabriel
PD Dr. biol. hum. Christiane Maier
Dr. rer. nat. Christian Wilhelm
Dr. rer. nat. Martin Ritthaler

Diagnostics

ADDITIONAL INFORMATION

Requested Regions

ACAN, ***ACP5***, ***B3GALT6***, ***B3GAT3***, ***B4GALT7***, ***BPNT2***, ***CANT1***, ***CCN6***, ***CFAP410***, ***CHST3***, ***COL11A1***, ***COL11A2***, ***COL2A1***, ***COL9A1***, ***COL9A2***, ***COL9A3***, ***CSGALNACT1***, ***DDR2***, ***DYM***, ***EIF2AK3***, ***EXTL3***, ***FLNB***, ***FN1***, ***GZF1***, ***HSPA9***, ***HSPG2***, ***IARS2***, ***INPPL1***, ***KIF22***, ***LONP1***, ***MBTPS1***, ***NANS***, ***NEPRO***, ***NKX3-2***, ***PAPSS2***, ***PCYT1A***, ***PISD***, ***POP1***, ***RAB33B***, ***RMRP***, ***RNU4ATAC***, ***RSPRY1***, ***SLC10A7***, ***SLC39A13***, ***SMARCAL1***, ***TMEM165***, ***TONSL***, ***TRAPPC2***, ***TRIP11***, ***TRPV4***, ***XYLT1*** (Spondylometaphyseal dysplasia and Spondylo-epi-(meta)-physeal dysplasia)

Methods

Sequencing: Protein-coding regions, as well as flanking intronic regions and additional disease-relevant non-coding regions, were enriched using in-solution hybridization technology, and were sequenced using the Illumina NovaSeq6000 system. In order to provide a second, independent validation, one or more rare variant(s) identified within the NGS data were additionally sequenced using conventional Sanger sequencing.

NGS based CNV-Calling: (only applicable for nuclear encoded genes) Copy number variations (CNV) were computed on uniquely mapping, non-duplicate, high-quality reads using an internally developed method based on sequencing coverage depth. Briefly, we used reference samples to create a model of the expected coverage that represents wet-lab biases as well as inter-sample variation. CNV calling was performed by computing the sample's normalized coverage profile and its deviation from the expected coverage. Genomic regions are called as variant if they deviate significantly from the expected coverage.

Please note that next generation sequencing based detection of copy number variations has lower sensitivity/specificity than a direct quantification method, e.g. MLPA. Copy-neutral structural aberrations can not be detected using this method (e.g. balanced translocations and balanced inversions). The absence of reported CNVs therefore does not ultimately guarantee the absence of CNVs.

Computational Analysis: Illumina bcl2fastq2 was used to demultiplex sequencing reads. Adapter removal was performed with Skewer. The trimmed reads were mapped to the human reference genome (hg19) using the Burrows Wheeler Aligner. Reads mapping to more than one location with identical mapping score were discarded. Read duplicates that likely result from PCR amplification were removed. The remaining high-quality sequences were used to determine sequence variants (single nucleotide changes and small insertions/deletions). The variants were annotated based on several internal as well as external databases.

Diagnostic data analysis: Variants were classified and reported based on ACMG/ACGS-2020v4.01 guidelines (Richards et al., 2015, PMID: 25741868, https://www.acgs.uk.com/quality/best-practice-guidelines/).

Only variants (SNVs/Small Indels) in the coding region and the flanking intronic regions (±8 bp) with a minor allele frequency (MAF) < 1.5% are evaluated. Known disease-causing variants (according to HGMD) are evaluated in up to ±30 bp of flanking regions and up to 5% MAF. Minor allele frequencies are taken from public databases (e.g. gnomAD) and an in-house database. If an acceptable sequencing-depth per base is not achieved by high-throughput sequencing, our quality guidelines demand local re-sequencing using classical Sanger-technology. Candidate CNV calls are evaluated manually. Potentially pathogenic findings are validated with a second method, like MLPA, on a case by case basis.

In this case, 97.91% of the targeted regions were covered by a minimum of 30 high-quality sequencing reads per base. **The evaluation of variants is dependent on available clinical information at the time of analysis.** The medical report contains all variants not classified as benign or likely benign according to current literature. Synonymous variants in mitochondrially encoded genes are classified as benign. *In silico* prediction of variants listed in the chart above is calculated on the basis of the output of the programs Mutation Taster, fathmm, Mutation Assessor, SIFT, fathmm-MKL coding, LRT, and PROVEAN according to the following criteria: 100% consensus = pathogenic/benign, ≥ 75% consensus = mostly pathogenic/benign, consensus < 75% or no prediction possible = inconsistent. SpliceAI was utilized to evaluate the consequence of variants on splicing (thresholds: 0.8-1 "splice effect", 0.6-0.8 "possible splice effect", <0.6 "no splice effect"; Jaganathan et al., 2019, PMID: 30661751). The prediction of a splice effect for missense variants is only indicated if a threshold of 0.8 is reached. This prediction can be complemented with additional *in silico* predictions in individual cases.

Variants are named according to the HGVS recommendations without any information regarding the cis or trans configuration.

The sample fulfilled our quality criteria upon arrival and during/after each processing step in the laboratory.

The procedure described above was developed and validated in-house (Laboratory developed test; LDT).

APPENDIX: ADDITIONAL VARIANTS IN ANALYZED GENES

Every individual carries many rare variants which are not necessarily causative for genetic disease. The requested analysis yielded variants of uncertain significance that, according to current scientific knowledge, are not likely to be causative for the phenotype of your patient. They have therefore not been included in the medical report. However, they are provided in the following table for transparency.

Based on the currently available scientific data, we do not encourage further investigation of these variants, should the clinical features of your patient remain unchanged.

Please note, that meaningful reevaluation of a rare variant requires additional phenotypic information or new scientific data.

Gene	Position Ref/Alt	Variant	Transcript	Zygosity	MAF (%)	rs-Number
XYLT1	chr16:17232246 T/C	c.1730A>G; p.Asn577Ser	NM_022166.4	het.	< 0.01	rs368060072

Please note that manual curation of the variants listed above has been performed without finding evidence for reliable disease relevance.

MAF: The minor allele frequency describes the least frequent allele at a specific locus in a given population (gnomAD). Het = Heterozygous, Homo = Homozygous, Hemi = Hemizygous, Heteropl = Heteroplasmic, Homopl = Homoplasmic. Positions refer to the hg19 genome build.