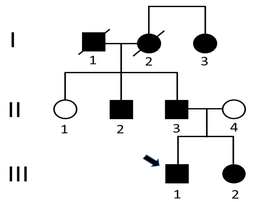
Evidence that non-syndromic familial tall stature has and oligo genic origin including ciliary genes

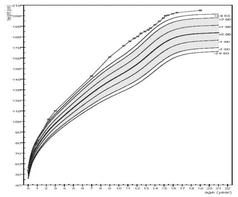
Birgit Weiss (1), Birgit Eberle (1), Ralph Roeth (1), Christiaan de Bruin (2), Julian C. Lui (3), Nagarajan Paramasivam (4), Katrin Hinderhofer (5), Hermine A. van Duyvenvoorde (6), Jeffrey Baron (3), Jan M Wit (2)*, Gudrun A. Rappold (1)*

[1]Department of Human Molecular Genetics, Institute of Human Genetics, Ruprecht Karls University Heidelberg, Heidelberg, Germany; [2]Department of Pediatrics, Leiden University Medical Center, Leiden, The Netherlands; [3]Section on Growth and Development, National Institute of Health, Bethesda, MD, United States; [4]Computational Oncology Group, Molecular Diagnostics Program at the National Center for Tumor Diseases (NCT) and German Cancer Research Center (DKFZ), Heidelberg, Germany; [5]Institute of Human Genetics, Ruprecht Karls University Heidelberg, Heidelberg, Germany; [6]Department of Clinical Genetics, University of Leiden, The Netherlands.

Human growth is a complex trait. A considerable number of gene defects have been shown to cause short stature, but there are only few examples of genetic causes of non-syndromic tall stature. Besides rare variants with large effects and common risk alleles with small effect size, oligogenic effects may contribute to this phenotype. Exome sequencing was carried out in a tall male and his parents. Filtered damaging variants with high CADD scores were validated by Sanger sequencing in the trio and three other affected and one unaffected family members. Network analysis was carried out to assess links between the candidate genes, and the transcriptome of murine growth plate was analyzed by microarray as well as RNA Seq. Heterozygous gene variants in *CEP104, CROCC, NEK1, TOM1L2* and *TSTD2* predicted as damaging were found to be shared between the four tall family members. Three of the five genes (*CEP104, CROCC* and *NEK1*) belong to the ciliary gene family. All genes are expressed in mouse growth plate. Pathway and network analysis indicated close functional connections. Together, these data expand the spectrum of genes with a role in linear growth and tall stature phenotypes.

Pedigree and Growth Analysis





Growth chart of the index case; epiphysiodesis was performed at 15.1 years

Exome sequencing was carried out in the index case with tall stature (height 3.5 SDS; arrow), his father (3.2 SDS) and mother with normal height (0.9 SDS). After bioinformatics filtering, all rare, predicted damaging variants found to be shared between the affected father and his son but not present in the mother with normal height, with a CADD score above 20 and not rated as polymorphisms in gnomAD, were considered. Twenty-eight selected rare gene variants were followed up and analyzed in three closely related family members by Sanger sequencing: two individuals with tall stature (II.2 and III.2) and one with normal height (II.1) (**see pedigree)**. The four family members with tall stature shared identical heterozygous variants, which were not present in the family members with normal stature, in the following genes: Centrosomal Protein 104 (*CEP104),* Rootletin *(CROCC),* Serine/threonine-protein kinase *(NEK1),* Thiosulfate Sulfurtransferase-like Domain containing 2 (*TSTD2)* and Target Of Myb1-Like 2 *(TOM1L2),* suggesting that these were potentially relevant for the phenotype.

Shared Gene Variants in Tall Individuals

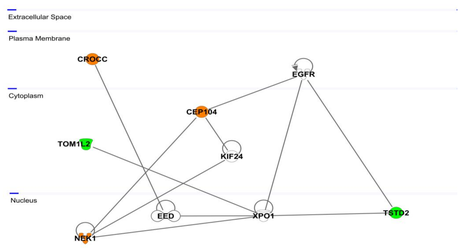
	Trio					
	Mother II.4	**Father II.3**	**Index P. III.1**	**Sister III.2**	**Brother/p II.2**	**Sister/p II.1**
Height (adjSDS)	**0.9** ○	**3.2** ■	**3.5** ■	**2.6** ●	**2.9** ■	**1.4** ○
CEP104 (27)	w	e	e	e	e	w
CROCC (25)	w	e	e	e	e	w
NEK1 (25)	w	e	e	e	e	w
TOM1L2 (26)	w	e	e	e	e	w
TSTD2 (45)	w	e	e	e	e	w

Gene variants shared between the four tall family members and not present in family members with normal hight (in grey color). Exome sequenced trio is indicated in light green; orange color indicates ciliary genes. A nonsense variant was identified in TSTD2; the others are missense variants. CADD scores are given in numbers behind the gene symbol; adjSDS, stands for standard deviation score adjusted for secular trend in the population;/p, indicates paternal side; circle, square male; e, heterozygous; w, wild type.

Evolutionary Conservation and Network Analysis

	CEP104	CROCC	NEK1	TOM1L2
Human	G	R	L	N
Chimp.	G	R	L	N
Rhesus.	G	R	L	N
Mouse	G	R	L	N
Chicken	G	-	-	N
Xenopus	G	R	L	N
Fugu	-	R	L	N
Zebrafish	N	-	L	N
Drosophila	N	K	L	N
C. elegans	-	-	-	N

Evolutionary conservation of the affected amino acids by missense mutations; –, no ortholog; a nonsense variant was identified in TSTD2.



Network Analysis by Ingenuity pathway analysis. Identified candidate gene products are highlighted in color; orange, proteins encoded by ciliary genes; green, proteins encoded by non-ciliary genes.

Network analysis of the five candidate genes for tall stature by the Ingenuity software revealed as "top diseases" *connective tissue and developmental disorders* ($1.32x10^{-3}$ - $2.20x10^{-4}$). Top hit for molecular and cellular function was *cell morphology* ($1.49x10^{-5}$- $1.49x10^{-5}$). Using the Ingenuity pathway analysis, close network links between the highlighted genes could be revealed. Direct interaction between NEK1 and CEP104 was previously experimentally demonstrated. Three of our candidate genes, *NEK1, TSTD2* and *TOM1L1* directly bind to Exportin 1 (*XPO1*), a cell cycle-regulated gene which mediates the nuclear transport of cellular proteins to the cytoplasm. XPO1 furthermore interacts via the Embryonic Ectoderm Development (EED) Polycomb protein (causative for an overgrowth syndrome similar to Weaver syndrome,) with the other candidate gene, *CROCC.* Together, these data highlight experimentally-based close network connections between the five genes.

Summary

Our genetic analysis in this family with non-syndromic extreme tall stature argues for an oligogenic origin. Five variants in *CEP104, CROCC, NEK1, TOM1L2* and *TSTD2* were found to be shared in all tall individuals of the family. All identified affected amino acids were highly conserved between species. Three genes, *CROCC, TOM1L2* and*TSTD2* were previously found to be associated with height in GWAS studies. All five candidate genes were found to be expressed in mouse growth plate. Three of the five genes (*CEP104, CROCC* and *NEK1)* are members of the ciliary gene family. Pathway and network analysis showed close interactions, indicating functionally connected genes.

In summary, we have shown that the identified genes encode mechanistically distinct proteins but their function converges on shared pathways and growth plate-related symptoms. They may therefore represent candidates for unsolved human skeletal disorders.

Reference:

Birgit Weiss[1] , Birgit Eberle[1], Ralph Roeth[1], Christiaan de Bruin[2], Julian C. Lui [3], Nagarajan Paramasivam[4], Katrin Hinderhofer[5], Hermine A. van Duyvenvoorde[6], Jeffrey Baron[3], Jan M. Wit[2]* and Gudrun A. Rappold[1]*

Frontier Endocrinol, Jun4; 12:660731, 2021