Published in final edited form as:
*Horm Res Paediatr*. 2014 ; 82(1): 44–52. doi:10.1159/000360857.

# Whole exome sequencing to identify genetic causes of short stature

**Michael H. Guo, BS**[1,2,3], **Yiping Shen, PhD**[4,5], **Emily C. Walvoord, MD**[6], **Timothy C. Miller, MS**[2], **Jennifer E. Moon, BA**[2], **Joel N Hirschhorn, MD, PhD**[1,2,3], and **Andrew Dauber, MD**[2,3]

[1]Department of Genetics, Harvard Medical School, Boston, MA, USA

[2]Division of Endocrinology, Boston Children's Hospital, Boston, MA, USA

[3]Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA

[4]Shanghai Children's Medical Center, Jiaotong University School of Medicine, Shanghai, China

[5]Departments of Laboratory Medicine and Pathology, Boston Children's Hospital and Harvard Medical School, Boston, MA, USA

[6]Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN, USA

## Abstract

**Background/Aims—**Short stature is a common reason for presentation to pediatric endocrinology clinics. However, for most patients, no cause for the short stature can be identified. As genetics plays a strong role in height, we sought to identify known and novel genetic causes of short stature.

**Methods—**We recruited 14 children with severe short stature of unknown etiology. We conducted whole exome sequencing of the patients and their family members. We used an analysis pipeline to identify rare nonsynonymous genetic variants that cause the short stature.

**Results—**We identified a genetic cause of short stature in 5 of the 14 patients. This included cases of Floating Harbor syndrome, Kenny-Caffey syndrome, the progeroid form of Ehlers-Danlos syndrome, as well as two cases of the 3-M syndrome. For remaining patients, we have generated lists of candidate variants.

**Conclusions—**Whole exome sequencing can help identify genetic causes of short stature in the context of defined genetic syndromes, but may be less effective in identifying novel genetic causes of short stature in individual families. Utilized in the clinic, whole exome sequencing can provide clinically relevant diagnoses for these patients. Rare syndromic causes of short stature may be under-recognized and under-diagnosed in pediatric endocrinology clinics.

### Keywords

whole exome sequencing; growth disorder; short stature; skeletal dysplasia; growth hormone

Corresponding author: Andrew Dauber, Division of Endocrinology, Boston Children's Hospital, Boston, MA 02115. Phone: 617-919-2413. Fax: 617-730-0856. andrew.dauber@childrens.harvard.edu. Reprint request to Andrew Dauber as above.

The other authors have nothing to disclose.

## Introduction

Short stature is one of the most common reasons for presentation to pediatric endocrinology clinics in the United States. Although there are many known causes of short stature, in most cases of non-familial short stature, there is no identifiable cause and the patient is classified as having idiopathic short stature (ISS) [1–3]. Without a definitive diagnosis, it can be difficult for clinicians to provide optimal treatment recommendations and to counsel patients on disease progression and recurrence risk [1].

It is well known that some cases of short stature have a single predominant genetic cause. These include Turner syndrome, mutations in the *SHOX* gene, and mutations in numerous genes in the growth hormone/insulin-like growth factor axis [1,4]. Additionally, there are currently more than 250 Mendelian syndromes or skeletal dysplasias associated with short stature that have had a causal gene identified, and this list is rapidly expanding. Studies have also revealed an increased burden of copy number variants (CNVs), particularly of rare deletions, in patients with short stature [5–7]. Thus, rare genetic sequence and copy number variants likely explain some cases of ISS, and it is important to identify these causes.

We hypothesized that rare highly penetrant nonsynonymous genetic variants could explain some cases of short stature of unknown cause. Strictly speaking, ISS refers to a condition characterized by normal birth weight, normal body proportions, no evidence of endocrine abnormalities, and no nutritional or psychosocial problems [2]. In this study, we sought to identify known or novel genetic causes of short stature in a cohort of patients with severe short stature presenting to an endocrine clinic, a third of whom met criteria for ISS. As part of a larger genetic study of patients with short stature [8], we selected a cohort of 14 patients with severe short stature and conducted whole exome sequencing and genome wide copy number assessment in the patients and their family members to identify genetic etiologies for their short stature. We successfully identified the genetic cause of short stature in 5 patients. Our report highlights the ability for whole exome sequencing to identify clinically important rare genetic disorders.

## Methods

### Patient Recruitment

This study was approved by the Institutional Review Board of Boston Children's Hospital. All participants or their legal guardians provided written informed consent. Subjects were recruited as part of a larger cohort searching for novel genetic etiologies of short stature in individuals with no known genetic etiology. Inclusion criteria for the larger cohort have previously been described [8]. Over the course of our study, we have selected individuals for exome sequencing who were predicted to have a higher likelihood of a monogenic cause of their short stature. Herein, we report the results of our exome sequencing studies of these patients. Specifically, we selected patients with severe short stature (>3 SD below mean for age and gender at study enrollment) who had received a prior standard clinical workup that was unable to identify a genetic cause of their short stature. However, subjects were allowed to have additional medical comorbidities, dysmorphic features, or other hormonal deficiencies as long as these alternate medical problems did not provide a clear explanation

for the subject's short stature. We selected patients only if DNA samples were available from both parents, and we gave preference to individuals for whom a DNA sample was available for a sibling as well. The total number of individuals selected for sequencing was based on the available budget for the project. In separate analyses, we have examined patients with more severe clinically diagnosed syndromes and were able to provide genetic diagnoses in a subset of these patients [9, 10].

## Sequencing

Whole exome sequencing of the participants was conducted from genomic DNA isolated from blood or saliva and was performed at the Broad Institute (Cambridge, MA) as previously described [11]. We included in our study only variants that passed all quality filters or received a VQSR SNP quality score above 98.50. All presumptive causal variants were verified by Sanger sequencing. As *de novo* variants are highly susceptible to sequencing and variant calling artifacts, we took several steps to remove potentially false positive calls. First, we visually inspected all aligned sequencing reads for *de novo* variants using the Integrative Genomics Viewer [12]. Second, we removed any calls based on low sequencing coverage (fewer than 10 sequencing total reads in the patient and in both parents). Finally, we Sanger sequenced and/or genotyped remaining *de novo* variants in the patients as well as their parents. Sequencing primers and genotyping probes are available upon request.

## Chromosomal Microarray

DNA samples from patients and unaffected family members were analyzed using Agilent SurePrint G3 custom 4×180K comparative genomic hybridization (CGH) and SNP microarray (Agilent Technologies) as previously described [13]. Candidate CNVs were filtered using standard clinical practices in the DNA diagnostic laboratory at Boston Children's Hospital.

## Variant Filtering

We hypothesized that these patients had rare highly penetrant genetic variants causing their short stature. We therefore only considered variants that were either not present or had minor allele frequency <1% in the 1000 Genomes Project (November 2012 release) [14], the National Institutes of Health Heart, Lung, and Blood Institute exome variant server (ESP 6500 release) [15], and the 50 control exomes whose variants were called in conjunction with our subject's variants. For *de novo* heterozygous variants and for consideration of dominant inheritance, we further restricted our analyses to variants not present in any of the above databases or in dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/). Analysis was limited to variants that were predicted to have an effect on the protein coding sequence, including nonsynonymous codon changes, frameshift variants, splice site variants, and coding indels.

We applied additional filters to consider several potential patterns of inheritance: 1) *de novo* – a novel variant present in the patient but neither parent, 2) homozygous recessive – a single candidate variant present in the homozygous state in the patient and heterozygous in each parent, 3) compound heterozygous – two or more candidate variants in a single gene

with at least a single candidate variant present in each parent, 4) and X-linked recessive in male patients – a maternally inherited candidate variant on the X-chromosome. For two families, we additionally considered a dominant mode of inheritance as described below.

## Results

### Description of cohort

We selected 14 patients with severe short stature (> 3 SD below the mean adjusted for age and gender at time of recruitment). The patients' heights and additional clinical features are listed in Table 1 (detailed clinical descriptions are provided in Supplementary File 1). For 12 families, DNA was available from the patient, both parents, and one sibling. For two families (P10 and P14), DNA was available from the patient and unaffected parents only. There were six male and eight female patients in the cohort. Seven of the patients in the cohort were treated with growth hormone therapy, but with variable clinical efficacy and for various lengths of time (Table 1).

The heights of the affected individuals ranged from −2.82 SD to −4.33 SD. These heights represent the most current height measurement, or if the patient has received GH therapy, the measurement immediately prior to GH administration (heights at enrollment are listed in Supplementary File 1). Immediate family members had heights above −2 SD and were considered to be unaffected except for families P04 and P05 where other family members had heights below −2 SD. All patients had undergone detailed evaluations in the pediatric endocrinology clinic and some had undergone additional evaluation by a clinical geneticist. All patients also had other genetic tests performed clinically (karyotypes, chromosomal microarrays, or diagnostic sequencing of candidate genes). Thus, prior to whole exome sequencing, for this cohort of patients, no cause of the short stature was identified despite extensive prior clinical evaluation (details are provided in Supplementary File 1).

### Variants identified by whole exome sequencing

Exome sequencing was performed to identify monogenic causes of short stature. Following the sequencing and filtering as described in Methods, we identified candidate variants for each patient (Table 2). On average per patient, we identified 1.1 (range 0 to 3) *de novo* variants, 0.5 (range 0 to 4) autosomal homozygous variants, 2.2 (range 0 to 4) X-linked hemizygous variants (for male patients), and 3.3 (range 0 to 8) genes with variants fitting a compound heterozygous model. We note that the rate of *de novo* variants we observed is higher than the expected number of nonsynonymous *de novo* variants expected (~0.65) [16]. However, the rate of *de novo* mutations in patients was similar to the rate found in paired unaffected siblings (average 0.75 *de novo* variants per sibling, p-value of 0.43, two-tailed paired t-test).

In addition to whole exome sequence analysis, copy number variants were assessed using a CGH+SNP microarray. After standard CNV filtering, we detected rare CNVs in four patients. Patient P02 had a deletion of *TFP1* (transferrin pseudogene 1) that was inherited from the unaffected father. Patient P04 had a deletion of *JAKMIP1* (encodes Janus kinase and microtubule-interacting protein 1) that was inherited from the unaffected mother. Patient

P11 had a duplication of *ARPC1A* (encoding actin-related protein 2/3 complex, subunit 1A) of unknown clinical significance and parental microarray data was unavailable. Patient P14 had a 493 kb duplication on chromosome 14q24.3, but this duplication was also found in the mother. Thus, no likely pathogenic copy number variants were identified. No runs of homozygosity (>10Mb) were detected in any patient to suggest uniparental disomy or parental consanguinity.

For each patient, we then manually searched in the Online Mendelian Inheritance of Man (OMIM) database and PubMed for known disease associations with the candidate variants. *In silico* predictions of functional effect of missense variants was assessed with the PolyPhen2 prediction tool [17]. All genetic diagnoses were made in the context of the patient's clinical features. We identified the genetic cause of short stature in five of the cases, including two cases of 3-M syndrome, one case of Floating Harbor syndrome, one case of the Kenny-Caffey syndrome, and one case of Ehlers-Danlos syndrome progeroid variant. In the remaining 9 cases, we generated a list of candidate genes based on our variant filtering criteria (Supplementary Table 2). We did not have enough evidence to assign causality for these candidate variants, but a few of these variants are interesting candidates (see Discussion). Details of all candidate variants are provided in Supplementary Table 2.

## Patients with identified genetic cause

As a result of our exome sequencing and analysis, we were able to identify a genetic cause of the short stature for five patients (Table 3), two of whom we have previously described (P02 and P11). Patient P02 received a genetic diagnosis of the progeroid form of Ehlers-Danlos syndrome due to compound heterozygous mutations in *B4GALT7* [18]. The patient presented with various skeletal abnormalities (most notably congenital dislocations of the elbow), skin received a genetic diagnosis of the 3-M syndrome due to recessive mutations in *CUL7* [11]. This patient was born small for gestational age (SGA) and had minor dysmorphic facial features in addition to severe short stature. Hypergonadotropic hypogonadism was discovered in this patient subsequent to the genetic diagnosis.

We report for the first time three other patients who we were able to diagnose using whole exome sequencing. Patient P09 is a South Asian male born SGA (5 lbs, 0 oz). Skeletal survey revealed proportionate short stature and gracile bones but no distinct skeletal dysplasia was identified. Growth hormone and thyroid hormone axes were normal. The patient reached appropriate developmental milestones. Russell-Silver syndrome was suspected, but molecular testing was negative. We identified a homozygous mutation in *OBSL1* (c.2134+1C>T) at a conserved splice site. *OBSL1* is one of the three genes (along with *CUL7* and *CCDC8*) known to cause the 3-M syndrome [19]. Notably, there is considerable overlap in clinical features between the 3-M syndrome and Russell-Silver syndrome [20]. Although hypergonadotropic hypogonadism has been previously associated with 3-M syndrome [11], the patient did not have this feature upon subsequent clinical evaluation (though the subject is currently only 14 years old).

Patient P07 is a Caucasian female born SGA (5 lbs, 0 ounces) as twin A of a term uncomplicated pregnancy. The patient experienced developmental delay most notable for significant language delay. In addition to short stature, she has a history of hypothyroidism.

The patient also has asymptomatic bilateral renal cysts, tethered cord, and severe scoliosis. Bone age was delayed as evaluated by the Greulich and Pyle method. Physical exam revealed a mildly dysmorphic facial appearance with a v-shaped nares, short philtrum, slightly beaked nose, and moderate scoliosis. We identified a *de novo* heterozygous point mutation (c.7330C>T) causing a p.Arg2444* nonsense mutation in the last exon of the *SRCAP* gene. Mutations in *SRCAP* are known to cause the Floating-Harbor syndrome [21], and the patient's mutation is the most frequent mutation identified in a cohort of 52 patients with Floating Harbor syndrome [10].

Patient P14 is a female who was born of a normal pregnancy and reached appropriate developmental milestones. The patient experienced failure to thrive in infancy requiring hospitalization. Testing at the time revealed elevated liver enzymes which self-resolved. Review of systems is notable for severe myopia but no other ophthalmological abnormalities, impaired dentition, and delayed primary eruption of teeth. Facial appearance demonstrates a prominent forehead, a flat nasal bridge, and a curved and upturned nose. Skeletal survey was notable for bilateral coxa valga and mild medullary stenosis and cortical thickening of the long bones, particularly the femur. Growth hormone levels were normal. The patient was variably treated with GH and IGF-1 therapy with an overall moderate response (Supplementary File 1). The patient also demonstrated early normal pubertal onset and was treated with leuprolide. Importantly, calcium, phosphate, and parathyroid hormone levels were all normal. We identified a *de novo* heterozygous point mutation (c.1706G>A) in *FAM111A* causing a missense change (p.Arg569His). This is a known recurrent *de novo* variant causing the Kenny-Caffey syndrome, although she is the first patient with this diagnosis with documented normocalcemia [22, 23].

## Additional Consideration of Dominant Inheritance

For families P04 and P05, we additionally considered dominant modes of inheritance, as for these two families, it was less clear if other family members were affected. For family P04, both the father (−2.65 SD) and sister (−2.32 SD) had milder short stature, while in family P05, the father had significant short stature (−3.46 SD) and the brother had mild short stature (−2.42 SD). Therefore, we considered dominant modes of inheritance with transmission from the father to the patient and sibling, or to the patient alone. When considering the patients' siblings as affected, we identified 30 rare heterozygous nonsynonymous variants in family P04 and 29 such variants in family P05 (Supplementary Table 3, 4). When considering the patients' siblings as unaffected, we identified 17 rare nonsynonymous variants in family P04 and 28 such variants in family P05 (Supplementary Table 3, 4). We did not identify any variants that were clearly causal, but we do note one interesting missense variant in *NPR2* (p.E389D) in family P04 that was seen in the patient, her father, and her sister. Although heterozygous variants in *NPR2* are associated with short stature [24, 25], PolyPhen2 predicted the variant to be benign and this is not a previously reported pathogenic variant. We therefore did not have enough evidence to assign causality for this variant. In family P05, there was a frameshift mutation NM_033360.2:c.556_557insT in the last exon of an alternative transcript of *KRAS* that was present in the three affected members of the family. Activating heterozygous mutations in *KRAS* are known to cause various RASopathies associated with short stature. However, as this mutation is a frameshift

(making it more likely to cause decreased function) and it occurs in an alternative transcript, it is unlikely that this mutation is the cause of the short stature. Furthermore, the affected individuals do not have any of the clinical features consistent with a RASopathy; thus, we do not believe that the frameshift mutation in *KRAS* is the cause of the patient's short stature.

## Discussion

Short stature is a common reason for visits to pediatric clinics but for most patients, a cause of the short stature cannot be identified. This inability to identify a cause for these patients hampers clinicians' ability to effectively manage and treat many of these patients. However, given the strong role that genetics plays in regulating human height, it is likely that many patients will have a genetic cause of their short stature. Whole exome sequencing provides the opportunity to rapidly assess for coding variants throughout the genome and can be highly advantageous for the diagnosis of rare monogenic syndromes that are difficult to clinically diagnose otherwise. In our cohort, we were able to provide a genetic diagnosis in 5 out of 14 patients. We also generated and report candidate variants in the remaining 9 patients, which will allow other researchers studying short stature or other genetic syndromes to cross-reference with our list for potential overlap.

Our study demonstrates how whole exome sequencing can be effectively used to identify genetic causes of disease. By obtaining samples from unaffected members of the same family, we can effectively identify *de novo* and inherited variants. Considering only rare variants (minor allele frequency <1% in reference databases), we are able to generate a tractable list of candidate variants/genes for each patient that is amenable to interpretation. Through this approach, we were able to identify the causal gene in 36% of families analyzed. Our sample size is quite small, limiting our ability to generalize our findings. However, our diagnostic yield is comparable to that found in a larger study of clinical exome sequencing applied to a wide range of phenotypes [26]. As an additional caveat, we intentionally selected patients with significant short stature thus representing the more severe end of the spectrum of patients seen by pediatric endocrinologists, which likely increased our yield. Nonetheless, applied in the pediatric endocrinology clinic setting, our results indicate that a significant proportion of patients with severe short stature could receive a genetic diagnosis. Larger studies of more varied patient populations are needed to better understand the efficacy of diagnostic whole exome sequencing for patients with short stature. As sequencing and interpretation technologies improve, the diagnostic yield is likely to increase. We note that sequencing of the unaffected sibling is not always necessary, but can aid in filtering out variants also present in the unaffected sibling.

In current clinical practice, a traditional candidate gene approach is taken where clinicians order specific tests based on clinical features resembling a known syndrome. This approach can impede diagnosis of atypical cases, in which the relevant syndrome is unlikely to be considered by the clinician. In contrast, whole exome sequencing considers the vast majority of the coding regions of the genome without a pre-existing clinical suspicion, thus allowing for the diagnosis of atypical presentations of rare syndromes. This is typified by patient P02, whom we diagnosed with the progeroid form of Ehlers-Danlos syndrome. Although the

syndrome resulting from *B4GALT7* deficiency had been previously characterized by a progeroid facial appearance, our patient did not have progeroid features despite having causal mutations in *B4GALT7* [18], and his description has helped refine the phenotypic spectrum of this disorder. Patients with Kenny-Caffey syndrome (including patients with the same causal variant as our patient P14) have been described as having hypoparathyroidism and concomitant hypocalcemia [22, 23]. By contrast, patient P14 demonstrated normal blood calcium, phosphate, and parathyroid hormone levels both in infancy and as an adolescent. As we learn more about the variation in clinical presentation of various gene disorders, “atypical” presentations may actually become more prevalent than the “classic” textbook syndromic presentations.

It is interesting to note that we identified the 3-M syndrome in two patients (P09 and P11) in our cohort. The 3-M syndrome is believed to be extremely rare, with roughly 200 patients reported worldwide, but is likely significantly underdiagnosed [27]. Our observation of two patients in our small cohort is consistent with substantial underdiagnosis in pediatric clinics. The syndrome bears strong phenotypic resemblance to the Russell-Silver syndrome [20] and one of our patients (P09) was evaluated for Russell-Silver syndrome. Thus, we believe that pediatricians and pediatric endocrinologists should be more aware of the 3-M syndrome, and consider this diagnosis in all patients with negative testing for Russell-Silver syndrome. The diagnosis of 3-M syndrome can have important clinical ramifications as some patients with this syndrome develop hypergonadotropic hypogonadism [11].

Our entire approach to exome analysis is predicated on the assumption that these patients’ short stature is due to monogenic highly penetrant rare genetic variants following a classic pattern of Mendelian inheritance. This scenario is clearly an oversimplification, but it is currently the most tractable interpretive approach for exome data analysis and is also the approach that is taken in clinical exome interpretation. This standard approach has several limitations. First, this approach largely ignores the possibility of genetic variants in multiple genes contributing to the short stature (i.e. oligogenicity or polygenicity). Second, we did not analyze variants that are not rare (>1% minor allele frequency in reference databases) although these variants are unlikely to have a large effect on stature. Additionally, we also only analyzed nonsynonymous variants, although it is well known that synonymous changes can perturb gene expression and function. We did not consider noncoding variants, epigenetic changes, and somatic changes (i.e., non-germline), as these are not captured by whole exome sequencing and are not currently amenable to interpretation. Finally, we note that it can be difficult to definitively assign other family members as affected or unaffected (see discussion above for patients P04 and P05). This caveat is important as our strategy is predicated on the accurate assignment of phenotypic status and the assumption that causal variants are highly penetrant and manifest as consistent phenotypes. Despite these limitations, we were able to identify probable major genetic contributors to short stature in five of fourteen patients.

While we were successful in identifying patients with rare syndromes for which the underlying genetic cause is already known, we were not successful in identifying novel genetic etiologies of short stature in this cohort. There are a number of factors that limit our ability to definitively identify novel pathogenic variants. We can identify numerous variants

that segregate with the short stature in the family, are damaging *in silico*, and meet our strict allele frequency criteria; however, many of these variants are in genes with little or no known biological evidence to support or preclude a causal role in short stature. Our approach also suffers from not having multiplex families or other individuals outside of the family with the same syndrome, preventing us from cross-referencing variants or performing linkage analysis. However, this scenario is not atypical in clinical medicine, where the clinician is usually evaluating an isolated patient with short stature. This highlights how large-scale international collaborations are needed in which detailed phenotypic and genotypic data are freely shared to facilitate the discovery of novel genetic causes of short stature.

We did identify several variants that are compelling candidates, but for which we do not have enough evidence at this time to make a convincing case for causality. Here we highlight two examples. Patient P01 has a missense *de novo* hemizygous change in *BTK*, which is known to cause a syndrome of X-linked agammaglobulinemia with isolated growth hormone deficiency. Although the patient has growth hormone deficiency, he does not have low immunoglobulin levels and it is not known if *BTK* mutations can cause GH deficiency in isolation. Patient P08 has compound heterozygous variants in *COL2A1*, which is known to cause a variety of skeletal dysplasia and short stature syndromes. However, patient P08 does not have any known skeletal anomalies, and it is unknown if this patient's variants in *COL2A1* could cause idiopathic short stature. We have provided lists of all variants that meet our filtering criteria with the hope that other researchers will encounter other patients with similar clinical features and pathogenic variants in the same gene and help to nominate causal genes across a cohort of patients.

In summary, our study outlines and demonstrates an approach to whole exome sequencing that can help identify genetic causes of short stature in patients with severe short stature without a known cause. Our approach is effective in the context of defined genetic syndromes, but in isolated families may be less effective in identifying novel genetic causes of short stature. Applied in the clinic, whole exome sequencing has the potential to reduce the diagnostic odyssey of patients with severe short stature especially in those with other clinical and/or biochemical abnormalities suggestive of a genetic syndrome. This process can be particularly challenging for patients with relatively nonspecific features who might present to a pediatric specialty clinic for evaluation of short stature. However, as new genetic causes of short stature and growth disorders are being identified, we expect that whole exome sequencing will become more widely used in pediatric endocrinology clinics and will facilitate obtaining a genetic diagnosis in a subset of children with idiopathic short stature.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

Disclosure summary: A.D. has previously consulted for Ipsen Pharma and Pfizer Inc. J.N.H. received grant support from Pfizer Inc (2011-present).

Grant Support: This work was conducted with support from Harvard Catalyst | The Harvard Clinical and Translational Science Center (National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health Award 8UL1TR000170-05 and financial contributions from Harvard University and its affiliated academic health care centers). The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University and its affiliated academic health care centers, or the National Institutes of Health. Sequencing experiments were performed by the Sequencing Core Facility of the Molecular Genetics Core Facility at Boston Children's Hospital supported by NIH P30-HD18655. This work was also supported by NIH Grant 1K23HD073351 (to A.D.), the Pediatric Endocrine Society Clinical Scholar Award (to A.D.), the Translational Research Program at Boston Children's Hospital, March of Dimes Grant 6-FY09-507 (to J.N.H.), Shanghai Science and Technology Commission for Major Issue Grant 11DZ1950300 (to Y.S.), Chinese National Natural Science Foundation Grant 81371903 (to Y.S), and a Harvard Medical School Genetics and Genomics PhD Training Grant 5T32GM096911-03 to M.H.G).

## References

1. Seaver LH, Irons M. American College of Medical Genetics (ACMG) Professional Practice and Guidelines Committee. ACMG practice guideline: genetic evaluation of short stature. Genet Med. 2009; 11:465–470. [PubMed: 19451827]
2. Cohen P, Rogol AD, Deal CL, Saenger P, Reiter EO, Ross JL, Chernausek SD, Savage MO, Wit JM. 2007 ISS Consensus Workshop participants. Consensus statement on the diagnosis and treatment of children with idiopathic short stature: a summary of the Growth Hormone Research Society, the Lawson Wilkins Pediatric Endocrine Society, and the European Society for Paediatric Endocrinology Workshop. J Clin Endocrinol Metab. 2008; 93:4210–4217. [PubMed: 18782877]
3. Lindsay R, Feldkamp M, Harris D, Robertson J, Rallison M. Utah Growth Study: growth standards and the prevalence of growth hormone deficiency. J Pediatr. 1994; 125:29–35. [PubMed: 8021781]
4. David A, Hwa V, Metherell LA, Netchine I, Camacho-Hubner C, Clark AJ, Rosenfeld RG, Savage MO. Evidence for a continuum of genetic, phenotypic, and biochemical abnormalities in children with growth hormone insensitivity. Endocr Rev. 2011; 32:472–497. [PubMed: 21525302]
5. Zahnleiter D, Uebe S, Ekici AB, Hoyer J, Wiesener A, Wieczorek D, Kunstmann E, Reis A, Doerr HG, Rauch A, Thiel CT. Rare copy number variants are a common cause of short stature. PLoS Genet. 2013; 9:e1003365. [PubMed: 23516380]
6. Dauber A, Yu Y, Turchin MC, Chiang CW, Meng YA, Demerath EW, Patel SR, Rich SS, Rotter JI, Schreiner PJ, Wilson JG, Shen Y, Wu BL, Hirschhorn JN. Genome-wide association of copy-number variation reveals an association between short stature and the presence of low- frequency genomic deletions. Am J Hum Genet. 2011; 89:751–759. [PubMed: 22118881]
7. van Duyvenvoorde HA, Lui JC, Kant SG, Oostdijk W, Gijsbers AC, Hoffer MJ, Karperien M, Walenkamp MJ, Noordam C, Voorhoeve PG, Mericq V, Pereira AM, Claahsen-van de Grinten HL, van Gool SA, Breuning MH, Losekoot M, Baron J, Ruivenkamp CA, Wit JM. Copy number variants in patients with short stature. Eur J Hum Genet. 2013
8. Wang SR, Carmichael H, Andrew SF, Miller TC, Moon JE, Derr MA, Hwa V, Hirschhorn JN, Dauber A. Large Scale Pooled Next-Generation Sequencing of 1077 Genes To Identify Genetic Causes of Short Stature. J Clin Endocrinol Metab. 2013
9. Dauber A, Golzio C, Guenot C, Jodelka FM, Kibaek M, Kjaergaard S, Leheup B, Martinet D, Nowaczyk MJ, Rosenfeld JA, Zeesman S, Zunich J, Beckmann JS, Hirschhorn JN, Hastings ML, Jacquemont S, Katsanis N. SCRIB and PUF60 Are Primary Drivers of the Multisystemic Phenotypes of the 8q23.4 Copy-Number Variant. Am J Hum Genet. 2013
10. Nikkel SM, Dauber A, de Munnik S, Connolly M, Hood RL, Caluseriu O, Hurst J, Kini U, Nowaczyk MJ, Afenjar A, Albrecht B, Allanson JE, Balestri P, Ben-Omran T, Brancati F, Cordeiro I, da Cunha BS, Delaney LA, Destree A, Fitzpatrick D, Forzano F, Ghali N, Gillies G, Harwood K, Hendriks YM, Heron D, Hoischen A, Honey EM, Hoefsloot LH, Ibrahim J, Jacob CM, Kant SG, Kim CA, Kirk EP, Knoers NV, Lacombe D, Lee C, Lo IF, Lucas LS, Mari F, Mericq V, Moilanen JS, Moller ST, Moortgat S, Pilz DT, Pope K, Price S, Renieri A, Sa J, Schoots J, Silveira EL, Simon ME, Slavotinek A, Temple IK, van der Burgt I, de Vries BB, Weisfeld-Adams JD, Whiteford ML, Wierczorek D, Wit JM, Yee CF, Beaulieu CL, White SM, Bulman DE, Bongers E, Brunner H, Feingold M, Boycott KM. The phenotype of floating-harbor

syndrome: clinical characterization of 52 individuals with mutations in exon 34 of SRCAP. Orphanet J Rare Dis. 2013; 8:63. [PubMed: 23621943]

11. Dauber A, Stoler J, Hechter E, Safer J, Hirschhorn JN. Whole exome sequencing reveals a novel mutation in CUL7 in a patient with an undiagnosed growth disorder. J Pediatr. 2013; 162:202–204. e1. [PubMed: 22974575]
12. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol. 2011; 29:24–26. [PubMed: 21221095]
13. Carmichael H, Shen Y, Nguyen T, Hirschhorn J, Dauber A. Whole exome sequencing in a patient with uniparental disomy of chromosome 2 and a complex phenotype. Clin Genet. 2013; 84:213–222. [PubMed: 23167750]
14. 1000 Genomes Project Consortium. A map of human genome variation from population- scale sequencing. Nature. 2010; 467:1061–1073. [PubMed: 20981092]
15. NHLBI GO Exome Sequencing Project (ESP). Seattle, WA: Exome Variant Server; Available at: http://evs.gs.washington.edu/EVS/. [Accessed February 16, 2012]
16. Keightley PD. Rates and fitness consequences of new mutations in humans. Genetics. 2012; 190:295–304. [PubMed: 22345605]
17. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods. 2010; 7:248–249. [PubMed: 20354512]
18. Guo MH, Stoler J, Lui J, Nilsson O, Bianchi DW, Hirschhorn JN, Dauber A. Redefining the progeroid form of ehlers-danlos syndrome: Report of the fourth patient with B4GALT7 deficiency and review of the literature. Am J Med Genet A. 2013
19. Hanson D, Murray PG, Sud A, Temtamy SA, Aglan M, Superti-Furga A, Holder SE, Urquhart J, Hilton E, Manson FD, Scambler P, Black GC, Clayton PE. The primordial growth disorder 3-M syndrome connects ubiquitination to the cytoskeletal adaptor OBSL1. Am J Hum Genet. 2009; 84:801–806. [PubMed: 19481195]
20. Akawi NA, Ali BR, Hamamy H, Al-Hadidy A, Al-Gazali L. Is autosomal recessive Silver- Russel syndrome a separate entity or is it part of the 3-M syndrome spectrum? Am J Med Genet A. 2011; 155A:1236–1245. [PubMed: 21548126]
21. Hood RL, Lines MA, Nikkel SM, Schwartzentruber J, Beaulieu C, Nowaczyk MJ, Allanson J, Kim CA, Wieczorek D, Moilanen JS, Lacombe D, Gillessen-Kaesbach G, Whiteford ML, Quaio CR, Gomy I, Bertola DR, Albrecht B, Platzer K, McGillivray G, Zou R, McLeod DR, Chudley AE, Chodirker BN, Marcadier J, FORGE Canada Consortium, Majewski J, Bulman DE, White SM, Boycott KM. Mutations in SRCAP, encoding SNF2-related CREBBP activator protein, cause Floating-Harbor syndrome. Am J Hum Genet. 2012; 90:308–313. [PubMed: 22265015]
22. Unger S, Gorna MW, Le Bechec A, Do Vale-Pereira S, Bedeschi MF, Geiberger S, Grigelioniene G, Horemuzova E, Lalatta F, Lausch E, Magnani C, Nampoothiri S, Nishimura G, Petrella D, Rojas-Ringeling F, Utsunomiya A, Zabel B, Pradervand S, Harshman K, Campos-Xavier B, Bonafe L, Superti-Furga G, Stevenson B, Superti-Furga A. FAM111A Mutations Result in Hypoparathyroidism and Impaired Skeletal Development. Am J Hum Genet. 2013
23. Isojima T, Doi K, Mitsui J, Oda Y, Tokuhiro E, Yasoda A, Yorifuji T, Horikawa R, Yoshimura J, Ishiura H, Morishita S, Tsuji S, Kitanaka S. A recurrent de novo FAM111A mutation causes Kenny-Caffey syndrome type 2. J Bone Miner Res. 2013
24. Olney RC, Bukulmez H, Bartels CF, Prickett TC, Espiner EA, Potter LR, Warman ML. Heterozygous mutations in natriuretic peptide receptor-B (NPR2) are associated with short stature. J Clin Endocrinol Metab. 2006; 91:1229–1232. [PubMed: 16384845]
25. Vasques GA, Amano N, Docko AJ, Funari MF, Quedas EP, Nishi MY, Arnhold IJ, Hasegawa T, Jorge AA. Heterozygous Mutations in Natriuretic Peptide Receptor-B (NPR2) Gene as a Cause of Short Stature in Patients Initially Classified as Idiopathic Short Stature. J Clin Endocrinol Metab. 2013; 98:E1636–E1644. [PubMed: 24001744]
26. Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA, Braxton A, Beuten J, Xia F, Niu Z, Hardison M, Person R, Bekheirnia MR, Leduc MS, Kirby A, Pham P, Scull J, Wang M, Ding Y, Plon SE, Lupski JR, Beaudet AL, Gibbs RA, Eng CM. Clinical Whole-Exome Sequencing for the Diagnosis of Mendelian Disorders. N Engl J Med. 2013

27. Clayton PE, Hanson D, Magee L, Murray PG, Saunders E, Abu-Amero SN, Moore GE, Black GC. Exploring the spectrum of 3-M syndrome, a primordial short stature disorder of disrupted ubiquitination. Clin Endocrinol (Oxf). 2012; 77:335–342. [PubMed: 22624670]

**Table 1**

Patient Characteristics: Heights for all participants are shown as the height z-scores adjusted for age and gender of the participant. NA indicates sibling height not available. Patient birth weights are percentiles adjusted for gestational age and gender. More detailed clinical descriptions can be found in the Supplementary File 1.

| Subject Number | Gender | Patient's Height | Mother's Height | Father's Height | Sibling's Height | Patient Birth Weight (Percentile) | Additional Clinical Features | GH Treatment |
|---|---|---|---|---|---|---|---|---|
| P01 | M | −4.05 | −0.44 | −1.04 | −0.76 | 5 | GH deficiency, osteopenia, delayed bone age, febrile seizures | Yes, with good response |
| P02 | M | −4.33 | −0.12 | 1.07 | NA | 8 | Hyperflexible joints, hyperextensible skin, bilateral radioulnar synostosis | Yes, with no response |
| P03 | M | −3.54 | −1.21 | −1.81 | −1.87 | <1 | IUGR, low IGF-1, vitiligo | Yes, with mild transient response |
| P04 | F | −3.84 | −0.34 | −2.65 | −2.32 | <1 | implant | Yes, with mild transient response |
| P05 | F | −3.15 | −0.50 | −3.46 | −2.42 | <1 | SGA, thoracic hemivertebrae | Not treated |
| P06 | F | −2.89 | −1.28 | −1.45 | −1.95 | 3 | Borderline SGA, ISS, anxiety disorder | Not treated |
| P07 | F | −2.98 | 0.42 | 0.16 | −0.26 | <1 | SGA, Dysmorphic facial appearance, scoliosis, tethered cord, bilateral polycystic kidneys, hypothyroidism, and developmental delay | Not treated |
| P08 | F | −3.25 | −0.89 | −0.64 | −0.17 | 8 | ISS | Not treated |
| P09 | M | −3.28 | −0.17 | −1.74 | NA | <1 | SGA | Yes, but immediately discontinued due to side effects |
| P10 | F | −3.30 | 1.50 | 0.45 | NA (Trio) | 8 | ISS, delayed bone age | Not treated |
| P11 | M | −4.17 | 1.54 | 2.98 | −0.41 | <1 | SGA, hypergonadotropic hypogonadism | Yes, with moderate response |
| P12 | F | −3.11 | −1.03 | −1.58 | −0.21 | 3 | Borderline SGA, delayed bone age | Not treated |
| P13 | M | −2.82 | −0.18 | −1.21 | −0.35 | 21 | ISS, delayed bone age | Not treated |
| P14 | F | −5.99 | 0.52 | −1.85 | NA (Trio) | 20 | dental abnormalities, myopia | Yes, with moderate response |

**Table 2**

Number of variants per patient. Shown is the number of variants (or number of genes for compound heterozygous) meeting our filtering criteria (as described in Methods) for each patient. Listed in parentheses are the candidate genes meeting filtering criteria for each patient (details of these candidate variants are provided in Supplementary Table 2). The causal genes are highlighted in bold.

| Subject Number | Compound Heterozygous | *de novo Heterozygous* | Autosomal Recessive | X-linked Recessive |
|---|---|---|---|---|
| P01 | 4 (*TLDC1*, *LTBP4*, *TTN*, *PELI3)* | 1 (ZBED4) | 1 (*KCNT1)* | 3 (*BTK*, *GPKOW*, *MAGEE2*) |
| P02 | 3 (***B4GALT7***, *GPR98*, *PYGB*) | 1 (*PPFIA1*) | 0 | 4 (*ACE2*, *BCORL1*, *IL13RA2*, *VSIG1*) |
| P03 | 1 (PCDH15) | 2 (IKZF4, *PDXP*) | 0 | 1 (PHF16) |
| P04 | 3 (*RNF123*, *BAZ2B*, *SH3TC1*) | 2 (*LRRC6*, *HSP90AB1*) | 0 | Female |
| P05 | 3 (*SV2C*, *MMS19*, *TAS2R9*) | 0 | 0 | Female |
| P06 | 3 (*ATM*, *AHNAK2*, *LRP1B*) | 0 | 1 (*FAM129A*) | Female |
| P07 | 4 (*MYCBP2*, *LRP1B*, *PMFBP1*, *C17ORF66*) | 3 (*KLHL26*, *NUCB1*, ***SRCAP***) | 0 | Female |
| P08 | 4 (*TTN*, *COL2A1*, *ANO5*, *ZNF507*) | 0 | 0 | Female |
| P09 | 6 (*AFF1*, *DKFZ*, *FAM129C*, *LRRC14B*, *MUC5B*, *NOL6*) | 0 | 4 (*ABCA13*, ***OBSL1***(2), *PKD1L1*) | 4 (*ACOT9*, *ARHGEF6*, *P2RY4*, *WAS*) |
| P10 | 3 (*ATF6B*, *COL6A6*, *MAP1A*) | 2 (*FCGR1B*, *BTK*) | 0 | Female |
| P11 | 0 | 1 (FAM134*A*) | 1 (***CUL7***) | 0 |
| P12 | 8 (*ABCA12*, *CACNA1G*, *HERC2*, *MSLNL*, *NBEAL2*, *OGFOD2*, *PKD1L1*, *PLCG2*) | 1 (*NEFM*) | 0 | Female |
| P13 | 3 (CP, *RALGAPA1*, *DTX2*) | 1 (PIK3AP*1*) | 0 | 1 (*SSX7*) |
| P14 | 1 (*AL359075.1*) | 2 (***FAM111A***, *LRRC30*) | 0 | Female |
| **Average** | **3.29** | **1.14** | **0.50** | **2.17** |

**Table 3**

Pathogenic variants identified. The table lists the causal variants we have identified for the five patients for which we identified a genetic cause of the short stature. Variant location, reference and variant alleles are reported from hg19 coordinates. Minor allele frequency is based on the overall allele frequency in the NHLBI Exome Variant Server (ESP 6500 release) (NHLBI GO exome variant server). Functional effect was assessed with PolyPhen2 prediction tool, which estimates the effect of missense variants (Adzhubei, 2010). A score of 0.00 is least likely to perturb protein function, while a score of 1.00 indicates a missense variant that is mostly likely to perturb protein function. Functional annotation of SNPs was assessed using SnpEff 2.0.5 (Cingolani, 2012) and validated manually. Protein function was summarized using information found in the UniProt database. Online Mendelian Inheritance of Man (OMIM) database was used to identify any known disease associations. Additional variants that passed our filtering criteria for these patients are listed in Supplementary Table 2.

| Patient | Gene | Inheritance Pattern | Position (hg19) | Frequency (Exome Variant Server) | Functional Annotation (SnpEff) | AA change | Function (UniProt) | Associated Diseases (OMIM) |
|---|---|---|---|---|---|---|---|---|
| P02 | *B4GALT7* | Compound Heterozygous | Chr 5: 177035995<br>Chr 5: 177031251 | NA<br>0.000077 | missense<br>missense | L41P<br>R270C | Proteoglycan Synthesis | Progeroid variant of Ehlers Danlos Syndrome |
| P07 | *SRCAP* | *de novo* heterozygous | Chr 16: 30748691 | NA | nonsense | R2444* | chromatin remodeling and transcription coactivator | Floating Harbor Syndrome |
| P09 | *OBSL1* | Autosomal Recessive | Chr 2: 220431551 | NA | splice site donor | NA | cell scaffolding protein | 3M Syndrome |
| P11 | *CUL7* | Autosomal Recessive | Chr 6: 43013346 | NA | frame shift | c.2837_2840dupAGAT | cell scaffolding protein | 3M Syndrome |
| P14 | *FAM111A* | *de novo heterozygous* | Chr11: 58920847 | NA | missense | R569H | unknown | Kenny-Caffey Syndrome |