Namaqualand hip dysplasia in South Africa: The molecular determinant elucidated

G Agenbag, PhD; **A Vorster,** MSc; **S Julius,** BSc (Med) Hons; **R Ramesar,** PhD; **P Beighton,** MD, PhD, FRCP, FRSSA

Division of Human Genetics, Department of Pathology, Faculty of Health Sciences, University of Cape Town, South Africa

Corresponding author: G Agenbag (gloudi.agenbag@gmail.com)

Background. Namaqualand hip dysplasia (NHD) is a mild form of spondyloepiphyseal dysplasia in which progressive arthropathy of the hip joint is a major manifestation. The disorder was documented in a multigenerational South African (SA) family with antecedents from Namaqualand, a region in the north-west of the country. Linkage analysis revealed a locus that includes the collagen type II gene, *COL2A1*.
Objectives. To identify the pathogenic *COL2A1* variant causing NHD in an SA family.
Methods. One affected male with a clear diagnosis of NHD was selected for whole-exome sequencing (WES) on the Ion Torrent Proton platform. A probe-based assay and direct cycle sequencing were used to confirm that the prioritised variant segregated with the phenotype in the NHD family and was not present in unrelated controls from the same population.
Results. WES identified one heterozygous variant, c.2014G>T; p.(Gly672Cys), in the coding sequence of the *COL2A1* gene. The variant segregated with NHD in 23 affected family members and was previously reported in a Caucasian male with Perthes disease-like presentation.
Conclusions. It is now possible to provide a molecular diagnosis of NHD before hip problems present. The large, clinically well-characterised NHD family is a valuable resource that could provide more insight into the mechanisms responsible for the variable expression observed in individuals with this variant.

S Afr Med J 2021;111(1):57-60. https://doi.org/10.7196/SAMJ.2021.v111i1.14561

Namaqualand hip dysplasia (NHD), also referred to as Namaqualand spondyloepiphyseal dysplasia, is a mild form of spondylo-epiphyseal dysplasia (SED) in which progressive arthropathy of the hip joint is a major manifestation. The disorder was initially documented in 45 members in five generations of a South African (SA) family with antecedents from Namaqualand, a region in the north-west of the country.[1] NHD differs from the classic forms of SED in terms of relatively normal stature and inconsistency of mild to moderate vertebral changes. The disorder mostly affects the femoral capital epiphyses of the hip joints and the vertebral bodies, and does not involve other systems (Fig. 1). The anatomical distribution of NHD is localised, with normal epiphyses and metaphyses throughout the remainder of the skeleton. Transmission of NHD is consistently autosomal dominant with variable expression. Onset is in childhood (average age of onset 10 years) with progressive disability. Both sexes are similarly affected. In childhood, the clinical and radiological manifestations are similar to those of Perthes disease and initial misdiagnosis is not unusual. As the orthopaedic management of these two disorders is very different, accurate diagnosis is essential. Linkage analysis in NHD revealed a locus that includes the collagen type II gene, *COL2A1*,[2] but the disease-causing variant has not been identified.

Objectives

The study aimed to address this deficiency by means of whole-exome sequencing (WES) of an affected person from the Namaqualand family.

Methods

Cohort selection

Approval to undertake this study was obtained from the University of Cape Town's Human Research Ethics Committee (ref. no. HREC 675/2015). Informed consent was obtained from a large multigenerational SA family presenting with NHD and unrelated individuals from the mixed-ancestry population who served as healthy controls (*N*=76). One affected male with a clear diagnosis of NHD was selected for WES. Pathogenicity of identified variants was confirmed with segregation analysis in the NHD family by selecting individuals for whom biological material was available; this group included 23 affected and 32 unaffected persons. The family, the NHD phenotype and their clinical diagnosis were previously described and documented for 45 affected individuals.[1-3] The cohort was extended to 55 individuals for the present study.

DNA extraction and quantification

Genomic DNA was extracted from peripheral blood using an adapted salting out method[4] and quantified with a Nanodrop ND-1000 (Thermo Fisher Scientific, USA). The sample destined for WES was further quantified with a CFX96 real-time polymerase chain reaction (PCR) system (Bio-Rad, USA) and the TaqMan RNase P detection kit (MAN0007732 Rev. 2.0; Thermo Fisher Scientific), according to the manufacturer's instructions. DNA was diluted accordingly.

Next-generation sequencing: Exomes

An Ion Torrent adapter-ligated library was generated with 1 μg genomic DNA using the Ion AmpliSeq Exome RDY kit with a SimpliAmp thermal cycler following the manufacturer's protocol (MAN0010084 Rev. A.0; Thermo Fisher Scientific). Following AMPure bead purification (Beckman Coulter, USA), the concentrations of the adapter-ligated libraries were determined using the Bio-Rad CFX96 real-time PCR system and the Ion Library TaqMan quantification kit (MAN0015802 Rev. B.0; Thermo Fisher Scientific) and diluted to 100 pM. The emulsion PCR, emulsion breaking and enrichment were performed using the Ion PI Hi-Q Chef 200 kit
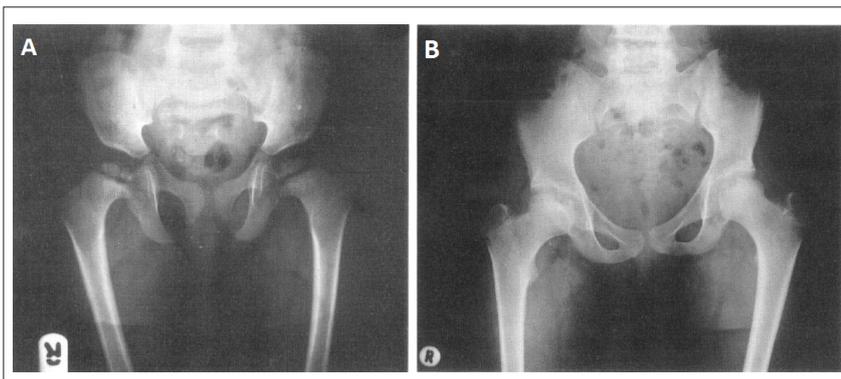
Fig. 1. Anteroposterior radiograph of the hips in (A) an NHD-affected child aged 10 years showing fragmentation of the femoral capital epiphyses, and (B) an NHD-affected adult aged 36 years showing severe arthropathy. (NHD = Namaqualand hip dysplasia.)

on the Ion Chef liquid handler according to the manufacturer's instructions (Thermo Fisher Scientific). An Ion PI chip kit v3 was templated with the Ion Chef and sequencing was performed on the Ion Proton. The Torrent Suite v4.0 (Thermo Fisher Scientific) was used to de-multiplex bar-coded reads and to generate run matrices, including chip-loading efficiency, coverage analysis and read quality. Raw signal data from the Ion Proton runs were processed using the Ion Torrent platform-specific pipeline to generate sequence reads, trim adapter sequences, filter, and remove poor signal-profile reads and align sequences to the reference (UCSC hg19). Variant calling was performed on the Ion Reporter v5.2 cloud-based software (Thermo Fisher Scientific) using the default workflow and parameters for this application.

The collagen genes, and specifically *COL2A1*, were prioritised for analysis based on linkage analysis results.[2] Variants were filtered according to their position, selecting only exonic and splice variants and excluding synonymous variants. Variants identified with WES were confirmed with direct cycle sequencing. Primers were designed using Primer3Plus[5] and standard conditions were used for PCR amplification of the relevant regions. Sequencing was performed with the BigDye Terminator v3.1 cycle sequencing kit according to the manufacturer's instructions (pub 4337035 Rev. D; Thermo Fisher Scientific) and capillary electrophoresis on a 3130*xl* Genetic Analyzer (Applied Biosystems, USA). Sequencing results were compared with the reference sequence for *COL2A1* (Genbank accession NM_001844.4).

Genotyping
Genotyping was carried out in NHD family members and population controls using a KASP (kompetitive allele-specific PCR) genotyping assay (LGC Genomics, UK) on a Bio-Rad CFX96 real-time PCR system according to the instructions provided by the assay manufacturer. A KASP-on-demand assay was designed based on reference sequences. Alleles were assigned based on the clustering profiles obtained from controls for each genotype. Any ambiguities were resolved with direct cycle sequencing.

Results
WES identified one heterozygous non-synonymous variant in the coding sequence of the *COL2A1* gene, and its presence was confirmed by direct cycle sequencing. The variant, c.2014G>T in exon 31, results in an amino acid substitution, p.(Gly672Cys). Analysis with a KASP assay showed the heterozygous genotype to be present in 23 NHD-affected individuals and absent in 32 unaffected individuals in this multigenerational SA family. The variant could not be detected in 76 unrelated controls from the same population group. The variant is also not present in the larger public genome databases (gnomAD, 1000 Genomes, ExAC and NCBI dbSNP), which confirms it to be an extremely rare variant. The above supports a pathogenic interpretation based on the guidelines published by the American College of Medical Genetics and Genomics (ACMG) (https://varsome.com/).[6]

Discussion
Various reports were published in the late 1980s (1987 - 1990) that attributed SED and related disorders affecting the epiphyses to variants in the *COL2A1* gene (MIM#108300). At least 460 *COL2A1* pathogenic variants have been identified in up to 21 distinct disorders in this category.[7,8] Based on these reports, and also the localisation of type II collagen fibrils in the cartilaginous epiphyses,

the locus was proposed as a likely candidate for NHD. Linkage studies confirmed that the link to the collagen type II locus was associated with the pathogenesis of NHD.[2] The current study, concerning 23 affected members of a large family with antecedents residing in Namaqualand, reports that the *COL2A1* c.2014G>T; p.(Gly672Cys) variant is responsible for the NHD phenotype.

Namaqualand is a remote semi-desert region in the north-west of SA, which was relatively isolated a few decades ago. The population has genetic admixture from Khoisan, Cape Malay (Javanese and Sumatran), indigenous African and Western European sources. Genealogical studies have not determined the source of the mutation, be it from one of the ancestral groups or whether it has arisen more recently by spontaneous mutation. The functional significance of this variant was not in question, as both the position and the nucleotide change have been reported previously.[9,10]

The *COL2A1* gene encodes the alpha 1 chain of procollagen type II. Three pro-alpha 1 chains twist together to form the procollagen molecule, which is processed into fibrils. The fibrils are cross-linked to form mature type II collagen fibres. Variants have been reported that are associated with a spectrum of rare, autosomal dominant conditions mainly characterised by skeletal dysplasia and short stature, collectively termed type II collagenopathies (Leiden Open Variation Database, http://databases. lovd.nl/shared/genes/COL2A1; Human Gene Mutation Database, http://www.hgmd.cf.ac. uk/ac/index.php). A common structural feature of collagen proteins is the presence of a triple helix. In COL2A1 the helical region is encoded by codons 201 - 1 214 and is composed of a core repeat of three residues that all begin with a glycine (Gly-X-Y). Most variants associated with type II collagenopathies are located in the region coding for the triple helix, with about a third replacing the glycine residue,[11] causing a dominant-negative effect that impairs homotrimer assembly and stability. Haplo-insufficiency, on the other hand, results from nonsense substitutions and premature termination due to out-of-frame deletions or insertions. Nonsense-mediated RNA decay is initiated that results in reduced but normal synthesis of collagen, leading to a weakened matrix of articular cartilage. The effect is associated with milder phenotypes compared with those mediated by amino acid substitutions. This model is well established for type I collagenopathies, but

less so for type II collagenopathies.[9] A previous study has shown that the position of amino acid substitutions affects the structure, thermostability and transport of mutant procollagen molecules and also the self-assembly of collagen fibrils.[12]

The *COL2A1* c.2014G>T variant (codon 672) is located in the helical region of the protein, resulting in the substitution of a glycine residue for cysteine. The variant is classified as pathogenic based on the most recent ACMG guidelines. Three cases involving codon 672, causing glycine substitutions, have been reported. The variant was previously detected in a boy with Perthes disease-like presentation.[10] The affected child presented with bilateral dysplastic capital femoral epiphyses and normal height, but various additional skeletal abnormalities were observed. The initial presentation of NHD can mimic Perthes disease. NHD differs by being typically bilateral, and the involvement of the spine, femoral necks and acetabula is also distinctive. The second account is for a missense variant at the same position, but resulting in a different amino acid substitution, p.(Gly672Ser), reported in a Japanese female presenting as SED congenita with severe coxa vara (SEDC-S).[9] The category 'SED' represents a range of disorders, which may include NHD. The SEDC-S phenotype, however, has additional features, including short stature, which is a trademark of classic SED. It is relevant that the case reported by Nishimura *et al.*[9] had bone abnormalities that were evident at birth. Codon 672 has also been reported to be associated with Stickler syndrome type I (c.2015G>C, p.(Gly672Ala), rs794727546) (NCBI ClinVar). Unfortunately no clinical information was provided for this report.

It has been found that glycine to serine substitutions in *COL2A1* result in alternating zones that produce severer and milder phenotypes, whereas glycine to non-serine residues cause more severe phenotypes but never a severe and perinatal lethal condition.[13] Although more severe, cysteine introductions are still producing a broad spectrum of phenotypes, with either normal or short stature. In the current study, a cysteine substitution at codon 672 caused NHD, a milder phenotype than reported for a serine substation at the same position.[9] A remarkably different, more severe phenotype was also reported for the cysteine substitution at codon 672.[10] There is still no clear phenotype-genotype correlation for *COL2A1* variants, and the impact of variants is unpredictable. Different substitutions at the same position may produce phenotypes varying from mild to moderate to lethal, and similar or identical variants are related to conditions with different phenotypes.[7] The variable phenotypes observed for p.(Gly672X) support the notion of genetic or environmental modifiers that family members may share.

Clinical variability and phenotypic overlap in *COL2A1*-related disorders are commonly observed.[9,14] NHD is behaving as a typical autosomal dominant skeletal disorder, and severity is variable and age related. In general terms the condition often presents in early childhood with unilateral or bilateral pain in the hips and associated difficulty and discomfort in walking. The disorder is bilateral but not necessarily symmetrical, and no other joints are involved. By adulthood the severity of disability necessitates hip joint replacement in a proportion of affected individuals. Some degree of platyspondyly was observed (60%), and variable broadness of the femoral necks and flatness of the humeral heads were observed in 20% of the individuals studied radiologically.[1] Inconsistent mild to moderate spinal malalignment and myopia in a small proportion of affected persons were the only other significant clinical problems.[2] Additional, mild abnormalities were observed that were considered insignificant at the time; however, the variable clinical presentation supports the notion of genetic and/or environmental modifiers and warrants further investigation. The large, clinically well-characterised NHD family is a valuable resource that could provide more insight into the mechanisms responsible for the variable expression observed in individuals with the variant.

Before NHD was clearly delineated, many affected family members had received a spurious diagnosis of Perthes disease. Currently, both Perthes disease and NHD have been accurately delineated, and they differ significantly in that Perthes disease is not familial, whereas in NHD all affected persons so far are members of the same extended family. NHD usually presents in early childhood, whereas Perthes disease is unusual before adolescence. There are no concomitants other than hip joint involvement in either of these conditions. The management of NHD is mainly orthopaedic and includes the provision of walking aids. Prosthetic joint replacement has proved to be effective in more severe cases.[3] Identification of the causative variant will permit presymptomatic diagnosis, which has practical significance for the clinical management of the asymptomatic potentially affected child. Although the *COL2A1* gene has a large number of exons (54), it is small enough to consider screening with direct cycle sequencing. However, a next-generation sequencing (NGS) approach is justified as it allows for the simultaneous interrogation of a larger set of coding sequences utilising a targeted gene panel or WES. NGS is becoming a mainstream modality in the diagnosis of genetic collagen disorders. The predominant femoral head involvement in NHD is similar to that in other unique SA disorders of uncertain aetiology, such as Upington bone disease and Beukes arthropathy. The observation that the *COL2A1* gene is involved in NHD has clear implications for the investigation of these conditions.

Conclusions

We identified a missense pathogenic variant in the *COL2A1* gene in 23 affected persons with NHD in a five-generation SA family. In view of the large size of this kindred, NHD is likely to be encountered in orthopaedic practice in this country. It is now possible to provide a molecular diagnosis before hip problems present in infancy and early childhood. *COL2A1* variants are known to cause variable clinical expression. This large family is a valuable resource to further study modifiers of the clinical expression of NHD in individuals with a shared genetic background.

1. Beighton P, Christy G, Learmonth ID. Namaqualand hip dysplasia: An autosomal dominant entity. Am J Med Genet 1984;19(1):161-169. https://doi.org/10.1002/ajmg.1320190116
2. Sher C, Ramesar R, Martell R, Learmonth I, Tsipouras P, Beighton P. Mild spondyloepiphyseal dysplasia (Namaqualand type): Genetic linkage to the type II collagen gene *COL2A1*. Am J Hum Genet 1991;48(3):518-524.
3. Learmonth ID, Christy G, Beighton P. Namaqualand hip dysplasia: Orthopedic implications. Clin Orthop Relat Res 1987;218:142-147. https://doi.org/10.1097/00003086-198705000-00020
4. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16(3):1215. https://doi.org/10.1093/nar/16.3.1215

5. Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res 2007;35(Suppl 2):71-74. https://doi.org/10.1093/nar/gkm306

6. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17(5):405-424. https://doi.org/10.1038/gim.2015.30

7. Deng H, Huang X, Yuan L. Molecular genetics of the *COL2A1*-related disorders. Mutat Res Rev Mutat Res 2016;768:1-13. https://doi.org/10.1016/j.mrrev.2016.02.003

8. Zhang B, Zhang Y, Wu N, Li J, Liu H, Wang J. Integrated analysis of *COL2A1* variant data and classification of type II collagenopathies. Clin Genet 2020;97(3):383-395. https://doi.org/10.1111/cge.13680

9. Nishimura G, Haga N, Kitoh H, et al. The phenotypic spectrum of *COL2A1* mutations. Hum Mutat 2005;26(1):36-43. https://doi.org/10.1002/humu.20179

10. Kannu P, Irving M, Aftimos S, Savarirayan R. Two novel *COL2A1* mutations associated with a Legg-Calvé-Perthes disease-like presentation. Clin Orthop Relat Res 2011;469(6):1785-1790. https://doi.org/10.1007/s11999-011-1850-x

11. Barat-Houari M, Sarrabay G, Gatinois V, et al. Mutation update for *COL2A1* gene variants associated with type II collagenopathies. Hum Mutat 2016;37(1):7-15. https://doi.org/10.1002/humu.22915

12. Steplewski A, Majsterek I, McAdams E, et al. Thermostability gradient in the collagen triple helix reveals its multi-domain structure. J Mol Biol 2004;338(5):989-998. https://doi.org/10.1016/j.jmb.2004.03.037

13. Hoornaert KP, Dewinter C, Vereecke I, et al. The phenotypic spectrum in patients with arginine to cysteine mutations in the *COL2A1* gene. J Med Genet 2006;43(5):406-413. https://doi.org/10.1136/jmg.2005.035717

14. Nagendran S, Richards AJ, McNinch A, Sandford RN, Snead MP. Somatic mosaicism and the phenotypic expression of *COL2A1* mutations. Am J Med Genet A 2012;158(5):1204-1207. https://doi.org/10.1002/ajmg.a.35303