Whole exome sequencing highlights variants in association with Keratoconus in Jordanian families

Tawfiq Froukh 1*, Ammar Hawwari 2 and Khalid Al Zubi 3

Abstract

Background: Keratoconus (KC) is usually bilateral, noninflammatory progressive corneal ectasia in which the cornea becomes progressively thin and conical, resulting in myopia, irregular astigmatism, and corneal scarring.

Methods: Eight families characterized by consanguineous marriages and/or multiple keratoconic individuals were examined genetically. Whole exome sequencing was done as trio or quadro per family. The output of the filtration procedure, based on minor allele frequency (MAF) less than 0.01 for homozygous variants and MAF equals 0 for heterozygous variants, is 22 missense variants.

Results: Based on the gene/protein function five candidate variants were highlighted in four families. Two variants were highlighted in one family within the genes MYOF and STX2, and one variant is highlighted in each of the other three families within the genes: COL6A5, ZNF676 and ZNF765.

Conclusion: This study is one of the very rare that highlights genetic variants in association with KC.

Keywords: NGS, Genome, Ocular, Epithelial, Dry-eye

Background

Keratoconus (KC) is sometimes bilateral, noninflammatory progressive corneal ectasia during which the cornea becomes progressively thin and conical, leading to myopia, irregular astigmatism, and corneal scarring. Patients with KC have cone shaped cornea (hence the name keratoconus, derived from the Greek word for cornea ('kerato') and cone shaped ('conus'). It usually arises in the teenage years and progresses, eventually it stabilizes in the 3rd/4th decades [1]. The clinical phenotypes of KC are highly variable; however, the common feature is corneal steepening which is normally detected at an early stage of the disease using computer-assisted corneal tomography. The current treatment of keratoconus is cross linking using riboflavin (vitamin B2) and ultraviolet light, which could prevent (stop) the progression of KC. In mild cases of KC, the refractive errors are usually treated by glasses or contact lenses but in more advanced cases, surgery is required in order to restore optimal visual acuity [2].

While the prevalence of KC in Jordan is unknown it is known that the severity of KC is related to the consanguineous marriages. For example the prevalence (per 100,000) of KC is high in societies with high consanguinity such as India (2300) [3], Iran (2500) [4], Lebanon (3330) [5], and Jerusalem (2340) [6] and low in societies with low consanguinity such as UK/Caucasian (57) [7], Denmark (86) [8], Finland (30) [1], Japan (17.3) [9], and USA (54.5) [10]. In Jordan consanguineous marriage ranges between 25.5% in the capital Amman to 52.1% in Irbid (www.consang.net), and therefore KC is estimated

* Correspondence: tfroukh@philadelphia.edu.jo
1Department of Biotechnology and Genetic Engineering, Philadelphia University, Jerash Road, Amman 11118, Jordan
2Full list of author information is available at the end of the article

© The Author(s). 2020 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
to have comparable prevalence rates to those in Lebanon and Jerusalem [11, 12].

Aetiology of KC involves environmental factors such as eye rubbing, atopy, and allergies as well as genetic factors. Because of its heterogeneity, KC is either associated with genetic systemic disorders or exhibited in isolated cases (i.e non-syndromic; Online Mendelian Inheritance In Man, OMIM#148300) [13]. Therefore, genetic factors are expected to cause non-syndromic KC especially in consanguineous families and/or families with multiple affected individuals.

The identification of genes causing non-syndromic KC has been the focus of many studies worldwide using genome wide linkage studies and genome wide association studies. Several variants have been implicated in genes coding for the primary composition of cornea including collagens and the components of the extracellular matrix. These genes include: LOX, CAST, DOCK9, IL1RN, SLCAA11, HGF, Rab3GAP1, TGFB1, ZNF469, ZEB1, VSX1, COL5A1, COL4A3, COL4A4, FNDC3B, FOXO1, MPDZ-NF1B, WNT10A, SOD1, IL1B, IL1A and MIR184. However not all analyses of these genes confirm their role in KC pathogenesis [14]. Mutations in VSX1 and SOD1 were assigned as causative to the isolated cases of KC [15, 16]. However, contradictory results were obtained regarding the pathogenic variants within these genes. While some studies revealed missense mutations associated with KC and thus suggesting an important role of these genes in KC [15, 17–19], others found that the mutations do not segregate with the disease or do not prove the pathogenicity of the disease [15, 20]. Moreover, in the era of Next Generation Sequencing (NGS) mutations in these genes were revealed in the genome aggregation database (gnomAD; https://gnomad.broadinstitute.org [21]), suggesting that the mutations in VSX1 and SOD1 may not contribute to KC. Other studies have shown that mutations in MIR184 (the most abundant expressed microRNA in the corneal and lens epithelia) are candidates (but not consistent) causing KC and cataract [22, 23]. In Jordan, NGS was recently used to identify variants causing neurodevelopmental diseases [24–31]. In one study homozygous frameshift variant in the gene GALNT14 is identified to be in association with KC [32].

In this study, Whole Exome Sequencing (WES) is used to identify rare variants associated with KC in eight Jordanian families.

**Methods**

**Families recruitment**

Patients with KC were recruited in July 2017 from Sight and Insight clinic/Amman in collaboration with Dr. Ammar Hiwari and from Al Karak hospital in collaboration with Dr. Khalid Al Zubi. KC was diagnosed during a routine eye test using slit lamp examination, snellen chart vision test, refraction Corneal topography by a pentacam machine “Oculus, Germany”. KC was confirmed using Belin / Ambrósio Enhanced Ectasia Display. Because the goal of this study is the identification of genetic factors being in association with KC, the targeted study group in this study is consanguineous families and/or multiple affected individuals in the same family.

Written informed consents were obtained from all participants in this study. Written informed consents were obtained from the parents of all participants under the age of 16. All participants have provided written informed consents to publish all identifying images, personal details and clinical details anonymously. Peripheral blood samples were collected from all participants.

**Molecular genetic analysis**

Total genomic DNA was extracted with Flexigene DNA kit. WES was conducted to 28 individuals (17 keratoconic and 11 healthy). Exome sequencing, bioinformatics analyses and variants validation by sanger sequencing were done as described previously [32].

**Variants filtration procedure**

We identified high-quality variants that are located in the protein coding genes including variants within the two base pair flanking the splicing sites (according to Ensembl-GRCh37.73). We maintained only the variants meeting the following quality criteria: at least 20X coverage and a mapping quality score ≥ 60. Variants meetings of high quality criteria were grouped into homozygous (recessive model) and heterozygous (dominant model) variants. In house controls of 200 Jordanian exomes were used in the filtration procedure.

Homozygous variants were filtered according to the following. Rare variants with minor allele frequency (MAF) less than 0.01 in the following databases: genome aggregation database (gnomAD), 1000 genome project and in the in-house sequenced controls. Shared homozygous variants between the keratoconic individuals per family but heterozygous in the parents (if not keratoconic) were maintained. Filtration was then done only for the loss of function variants (LOF; including stop gain, frameshift, splice site acceptor and splice site donor) and the missense variants predicted to be possibly- or probably-damaging by Polyphen2 Humvar and predicted to be deleterious by SIFT [33–35]. Loss of function (LOF) variants were also excluded if the genes they are located in carry other homozygous LOF variants in gnomAD, 1000 genome project or in the in-house sequenced controls. Variants located in genes reported with non-ocular OMIM disease were excluded because we aim to identify rare variants in association with non-syndromic KC (Fig. 1).
Heterozygous variants were filtered according to three criteria. Firstly, variants that are absent in the following databases: gnomAD, 1000 genome project and in the in-house sequenced controls. Secondly, shared heterozygous variants between the keratonic individuals per family but wild type in the parents (if not affected). Filtration was then done only for the loss of function variants (LOF; including stop gain, frameshift, splice site acceptor and splice site donor) and missense variants predicted to be possibly- or probably-damaging by Polyphen2 Humvar and predicted to be deleterious by SIFT. LOF variants were also excluded if the gene they are located in carry other LOF variants in gnomAD, 1000 genome project or in the in-house sequenced controls. Thirdly, variants were excluded if they are in genes reported with non-ocular OMIM disease because we aim to identify rare variants with non-syndromic KC (Fig. 1).

Results
Identified variants
Based on the filtration procedure, 22 variants were identified. All of the identified variants are missense. Four variants are homozygous in four families: KC003, KC004, KC010 and KC011 and 18 variants are heterozygous in five families: KC001(two variants), KC003 (three variants) KC005 (five variants), KC007 (one variants), KC008 (seven variants) (Table 1). None of the filtered variants is in a gene reported with OMIM disease.

Discussion
Family KC001
The family has 3 keratoconic siblings born to non-consanguineous parents (Fig. 2a). WES was performed to the 3 siblings (II-1, II-2 and II-3) and to their mother (I-1). Two heterozygous missense variants out of 118, 763 variants were identified in the genes CDC42BPA and LRRC16B. While no information about the gene LRRC16B, the gene CDC42BPA encodes a kinase protein that phosphorylates histone H1. This protein plays a role in the regulation of cytoskeleton and cell migration and abundant in the heart, brain, skeletal muscle, kidney and pancreas [36]. Despite the fact that the identified variants (c.4586G > A:p.Arg1529Gln in the gene CDC42BPA and c.2290G > A:p.Asp764Asn LRRC16B) in these two genes are predicted to be deleterious no relationship can be established between these two variants and KC because there are no reports link any of these two genes with ocular functions.

KC003
Distantly related parents in this family have two keratoconic progeny (Fig. 2b). Performing WES to II-1, II-2 and I-1 yields 113,885 variants. According to the filtration procedure four variants are identified. One variant is homozygous in the gene POLR2M and three are heterozygous in the genes GHITM, MYOF and STX2. The variants in the genes POLR2M and GHITM are excluded to be in association with keratoconus because the first encodes a subunit of a specific form of RNA polymerase II that act as negative regulator of transcriptional activation by the mediator complex [37], and the second encodes a growth hormone inducible transmembrane protein that is required for the mitochondrial tubular network and cristae organization [38]. Usually abnormalities in such functions are in association with developmental disorders.

The variants in the genes MYOF and STX2 are more likely to be in association with keratoconus because MYOF is calcium/phospholipid-binding protein that plays a role in the plasmalemma repair mechanism of endothelial cells that permits rapid rescaling of membranes disrupted by mechanical stress [39]. While the gene STX2 encodes a protein that regulates epithelial mesenchymal interactions and epithelial cell morphogenesis and activation [40].

KC004
The parents in this family I-1 and I-2 belong to two different families but from the same town nonetheless they have two keratoconic progeny II-1 and II-2 (Fig. 2c).
WES revealed 130,125 variants and the filtration identified one variant (c.5381G > A:p.Gly1794Glu) in the gene MLLT4 which is not likely contributing to keratoconus. This is because this gene encodes a protein involved in signalling and organization of cell junctions during embryogenesis [41]. Dysfunction here is most likely in association with central nervous system diseases and therefore, no variant can be assigned to this family in association with keratoconus.

**Table 1** Identified variants after the filtration procedure

| Family | Genotype | gene   | OMIM No. | Variant                                      |
|--------|----------|--------|----------|----------------------------------------------|
| KC001  | het      | CDC42BPA | 603,412  | ENST00000334218:c.4586G > A:p.Arg1529Gln    |
|        |          | LRR16B  | 614,716  | ENST00000342740:c.2290G > A:p.Asp646Asn     |
| KC003  | hom      | POLR2M  | 606,485  | ENST00000299638:c.42G > C:p.Glu14Asp        |
|        | het      | GHITM   | –        | ENST00000372134:c.770T > G:p.Val257Gly       |
|        |          | MYOF    | 604,603  | ENST00000359263:c.2906G > C:p.Cys969Ser     |
|        |          | STX2    | 132,350  | ENST00000392373:c.636T > G:p.His212Gln      |
| KC004  | hom      | MLLT4   | 159,559  | ENST00000400822:c.5381G > A:p.Gly1794Glu    |
| KC005  | het      | GOLGA4  | 602,509  | ENST00000356847:c.779G > A:p.Gly260Asp      |
|        |          | RPAP1   | 611,475  | ENST00000343030:c.190G > A:p.Asp646Asn      |
|        |          | NTN3    | 602,349  | ENST00000239373:c.142G > T:p.Ala48Ser       |
|        |          | TDRD12  | –        | ENST00000444215:c.1849T > G:p.Trp617Arg     |
|        |          | SPATA25 | –        | ENST00000372519:c.674C > G:p.Ser225Cys      |
| KC007  | het      | ZNF676  | –        | ENST00000397121:c.1240A > G:p.Ile414Val     |
| KC008  | hom      | XIR1P   | 609,777  | ENST00000340369:c.4G > A:p.Ala2Thr          |
|        |          | COL6A5  | 611,916  | ENST00000265379:c.5014T > G:p.Phe1672Val    |
|        |          | HLT5    | 603,257  | ENST00000310053:c.2269G > T:p.Asp757Tyr     |
|        |          | UTRN    | 128,240  | ENST00000367545:c.9911C > A:p.Ser3304Tyr    |
|        |          | TBATA   | 612,640  | ENST00000299290:c.130C > T:p.Pro44Ser      |
|        |          | KRTAP17–1| –       | ENST00000343202:c.13C > G:p.Pro45Ala       |
|        |          | SYCP2   | 604,105  | ENST00000357552:c.1278A > G:p.Gln426Hls    |
| KC010  | hom      | NPPB    | –        | ENST00000243400:c.1353G > T:p.Lys451Asn     |
| KC011  | hom      | ZNF765  | –        | ENST00000396408:c.1133C > T:p.Thr378Ile    |

WES revealed 130,125 variants and the filtration identified one variant (c.5381G > A:p.Gly1794Glu) in the gene MLLT4 which is not likely contributing to keratoconus. This is because this gene encodes a protein involved in signalling and organization of cell junctions during embryogenesis [41]. Dysfunction here is most likely in association with central nervous system diseases and therefore, no variant can be assigned to this family in association with keratoconus.

**KC005**

Distantly related parents with four keratoconic individuals including the mother were recruited in this family (Fig. 2d). The whole exome sequences in this family revealed 124,431 variants and the filtration procedure identified five variants in five different genes (Table 1). Screening the functions of the gene products of these five genes excludes any of them to be in association with keratoconus and therefore no variant can be assigned to this family.

**KC007 and KC011**

While the parents in these two families are first cousin with two keratoconic progeny (Fig. 2e and h), no homozygous variants were identified but rather a heterozygous variant (c.1240A > G:p.Ile414Val) in the gene ZNF676 for family KC007 and the variant c.1133C > T:p.Thr378Ile in the gene ZNF765. No detailed information is available about the proteins of these two genes except being zinc finger proteins that might be involved in transcriptional regulations. However, the identified variants in these two zinc finger proteins might be in association with KC because variants in another zinc finger protein (ZNF469) were previously reported in association with keratoconus [42–45].

**KC008**

In this family the mother and her daughter are keratoconic (Fig. 2f). The whole exome sequence revealed 103,750 variants and the filtered are seven variants in seven genes. By searching the functions of the proteins encoded by these seven genes, only one is prioritized COL6A5. This gene encodes a member of the collagen VI that acts as a cell-binding protein [46].

**KC010**

This is the only family in this study with one keratoconic individual who belong to nonconsanguineous parents (Fig. 2g). According to the filtration procedure only one variant was identified in the gene NPPB. This gene
encodes a nuclear pore complex interacting protein which does not seem to be involved with keratoconus and therefore no variants are assigned to this family.

Conclusion
In summary, WES and the validation by sanger sequencing allowed us to highlight variants in association with KC in four families (two variants in KC003 and one variant in each of KC007, KC008, and KC011) and no variants to highlight in three families (KC001, KC004, KC005 and KC010). The families were chosen in this study with multiple keratoconic individual as a sort of evidence for the genetic contribution on KC. Nonetheless, none of the prioritized variants in the families (KC003, KC007, KC008, and KC011) is strong candidate to cause KC. This is due to the lack of concrete information connecting between the functions of the gene products (proteins) and the cornea. Therefore, further experimental assays are needed to strengthen the possible causality of the identified variants and the candidate genes on KC.

Acknowledgments
The authors would like to thank all participating patients in this study, the lab technicians Aya Baraghithi and Talha Issa, and the informatician Anis Nazir.

Authors’ contributions
AH and KZ have designed the methodology in terms of recruiting the patients with keratoconus. TF analysed the whole exome data and wrote the manuscript. All authors have read the manuscript and approved it.

Funding
The deanship of scientific research and graduate studies/Philadelphia University/Jordan has completely funded this study.

Availability of data and materials
The highlighted variants in association with keratoconus in this article are available in Leiden Open Variation Database LOVD 3.0 (https://databases.lovd.nl/) according to reference genome GRCh37.73 used in this study.

KC003 (MYOF): https://databases.lovd.nl/shared/individuals/00307910
KC003 (STX2): https://databases.lovd.nl/shared/individuals/00307911
KC007 (ZNF676): https://databases.lovd.nl/shared/individuals/00307912
KC008 (COL6A5): https://databases.lovd.nl/shared/individuals/00307914
KC011 (ZNF765): https://databases.lovd.nl/shared/individuals/00307913

Ethics approval and consent to participate
This study was approved by the ethics committee of Philadelphia University in Jordan and conducted according to the principles outlined in the declaration of Helsinki in 1964. Written informed consents were obtained from all participants in this study. Written informed consents were obtained from the parents of all participants under the age of 16.

Consent for publication
All participants have provided written informed consents to publish all identifying images, personal details and clinical details anonymously. Moreover, written informed consents for publication of identifying images or other personal or clinical details were obtained from the parents or legal guardians of any participant under the age of 18.
