Gene 769 (2021) 145237

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Complete genome sequence and bioinformatics analysis of nine Egyptian females with clinical information from different geographic regions in Egypt

Mahmoud ElHefnawi [a,b,c,1,*], Elsayed Hegazy [a,1], Asmaa Elfiky [d,1], Yeonsu Jeon [e,f], Sungwon Jeon [e,f], Jong Bhak [e,f,g], Fateheya Mohamed Metwally [d], Sumio Sugano [h], Terumi Horiuchi [i], Abe Kazumi [h], Asta Blazyte [e,f,1]

[a] School of Information Technology and Computer Science, Nile University, Giza 12588, Egypt
[b] Informatics & Systems Department, the National Research Centre, Cairo, Egypt
[c] Biomedical Informatics and Chemoinformatics Group, Center of Excellence for Medical Research, National Research Centre, Cairo, Egypt
[d] Environmental and Occupational Medicine Department, Environmental Research Division, National Research Centre, Cairo, Egypt
[e] Korean Genomics Center (KOGIC), UNIST, Republic of Korea
[f] Department of Biomedical Engineering, School of Life Sciences, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea
[g] Personal Genomics Institute, Genome Research Foundation, Osong, Republic of Korea
[h] The Institute of Medical Science, University of Tokyo, Japan
[i] Graduate School of Frontier Sciences, University of Tokyo, Chiba, Japan

ARTICLE INFO

Original content: https://www.ncbi.nlm.nih.gov/sra/?term=SRP136979

Keywords:
Personal genome
Variant analysis

ABSTRACT

Egyptians are at a crossroad between Africa and Eurasia, providing useful genomic resources for analyzing both genetic and environmental factors for future personalized medicine. Two personal Egyptian whole genomes have been published previously by us and here nine female whole genome sequences with clinical information have been added to expand the genomic resource of Egyptian personal genomes. Here we report the analysis of whole genomes of nine Egyptian females from different regions using Illumina short-read sequencers. At 30x sequencing coverage, we identified 12 SNPs that were shared in most of the subjects associated with obesity which are concordant with their clinical diagnosis. Also, we found mtDNA mutation A4282G is common in all the samples and this is associated with chronic progressive external ophthalmoplegia (CPEO). Haplogroup and Admixture analyses revealed that most Egyptian samples are close to the other north Mediterranean, Middle Eastern, and European, respectively, possibly reflecting the into-Africa influx of human migration. In conclusion, we present whole-genome sequences of nine Egyptian females with personal clinical information that cover the diverse regions of Egypt. Although limited in sample size, the whole genomes data provides possible geno-phenotype candidate markers that are relevant to the region's diseases.

1. Introduction

Next-generation sequencing (NGS) technology is a powerful approach enabling an efficient study of a large scale genetic variants (Bentley et al., 2008; Wheeler et al., 2008; Metzker, 2010; Azim et al., 2013). Whole genome sequencing (WGS) by NGS can cover intronic areas that may contain rare and common deleterious mutations, that cannot be captured by whole exome sequencing (WES) that usually facilitates a deeper coverage of coding regions important for protein function analyses (Bamshad et al., 2011). One of the emerging applications of NGS is investigating complex diseases such as obesity along with its manifestations: insulin resistance, impaired glucose tolerance, and dyslipidemia. Even though Genome-Wide Association Studies (GWAS) may rely on genotyping data like in Frayling et al. (2007) and

Abbreviations: WGS, Whole-genome sequencing; CPEO, chronic progressive external ophthalmoplegia; NGS, next-generation sequencing; GWAS, Genome-Wide Association Studies; EGY, Egyptian; VEST, Variant Effect Scoring Tool; rCRS, Revised Cambridge Reference Sequence.

* Corresponding author at: Informatics & Systems Department, the National Research Centre, Cairo, Egypt.
E-mail address: mahef@aucegypt.edu (M. ElHefnawi).

[1] These authors contributed equally to the work.

https://doi.org/10.1016/j.gene.2020.145237
Received 25 November 2019; Received in revised form 3 August 2020; Accepted 11 October 2020
Available online 27 October 2020

Loos et al. (2008), wider coverage increases the possibility of finding novel variants or genes that were not associated with the obesity before. The Fat mass and obesity-associated gene (*FTO*) is an iconic target-gene in such studies as its variants were the first ones associated with body weight increment, obesity, and one variant specifically with insulin resistance (Sällman Almén et al., 2013). For example, by targeted resequencing (NGS) Volckmar et al. (2016) detected a novel *FTO* variant, Val83Leu, which leads to a loss of *FTO* function in lean individuals, apart from six already known rare and non-synonymous variants in the same gene. Nordang et al. (2017) applied NGS technology to target exons and exon–intron junctions in the five obesity-linked genes including, leptin (LEP) which is responsible for serum leptin levels (Nordang et al., 2017),leptin receptor *(LEPR)*, melanocortin-4-receptor (*MC4R*) involved in satiety regulation (Nordang et al., 2017); proprotein convertase subtilisin/kexin-type-*1* (*PCSK1*) associated with abnormal plasma glucose and insulin levels (Muhsin et al., 2020), and pro-opiomelanocortin (*POMC*) suspected to be involved in dyslipidemia (Sudo et al., 2005) for a Norwegian-based cohort, and found five novel variants. Moreover, they found four variants in *MC4R* gene that were predicted to be pathogenic or likely pathogenic, and three *MC4R* variants previously reported to be associated with monogenic obesity. Furthermore, Kleinendorst et al. (2018) developed an NGS based gene panel able to identify mutations in 52 obesity-related genes, including *MC4R,* which they concluded to be the most frequent genetic cause of obesity. Therefore, the NGS-based gene panel analysis in patients with obesity is a useful and constantly improving tool for estimating genetic predisposition to obesity and obesity-related metabolic issues. Moreover, environmental factors such as socioeconomic status and lifestyle, can complement the effect of the obesogenic factors (Trevino et al., 1999; Dehghan et al., 2005; Harper, 2006; Biro and Wien, 2010). However, the response to the environmental factors can also be ethnic-group specific, which makes GWAS difficult to interpret. For example, Caucasians with different genetic backgrounds, but living in a similar obesogenic environment is less susceptible to developing obesity (~32%) or T2D (~8%) when compared to Pima Indians who live in Arizona where obesity and T2D in those populations were (~64%) and (~30%), respectively (Vimaleswaran and Loos, 2010; Choquet and Meyre, 2011).

As for Egyptians, obesity studies have exceptional relevance as it is one of the nations with the steepest obesity rate in a 33 yearlong study (1980–2013) (Ng et al., 2014). A study published in 2016 reported >30 percent of the Egyptian population to be obese with nearly half of all females being obese (46%) which is nearly two times the magnitude of this issue in Egyptian males (22%) (Alebshehy et al., 2016). The majority (66%) of Egyptian women is overweight, in fact, Egyptian women of reproductive age in just 10 years showed an average increase in BMI by 2.21 kg/m^2 (Austin et al., 2013). A recent PCR-based study came to a conclusion that *LEPR* Gln223Arg, *UCP2* G 866 A and *INSR* exon 17 polymorphisms are linked to obesity in Egyptians and some of the risk variants are indeed sex biased (Hassan et al., 2017). In our previous work the two Egyptian male whole genomes designated as (EGP1 and EGP2) from Delta and Said regions, respectively, shared the SNPs Asn985Tyr in *RP1* and Thr55Ala in *FABP2* which are associated with an increased level of cholesterol and triglycerides, possibly related to obesity (ElHefnawi et al., 2018).

2. Materials and methods

2.1. Study group and ethical approval

A total of nine Egyptian females were recruited for this study, the mean age of study group was 41 ± 14.35 years. The subjects (EGY) were enrolled by National Research Center's outpatient clinic and approved by the Medical Research Ethics Committee of the National Research Centre –Egypt (MREC, NRC) with approval number (13177) and with 1964 Helsinki Declaration and its later amendments. Signed informed consents were obtained from the participants in this study. A clinical examination was conducted for each of them; weight (weight), height (Ht) and waist circumference (WC) were measured. Body mass index (BMI) was calculated according to WHO (Organization, 2000). Blood pressure was measured for each individual. Five ml of whole blood was centrifuged at 12,000 rpm for 10 min using (Heraeus Labofuge 400R) centrifuge to separate serum and stored at −80 °C until analysis.

2.2. Biochemical analyses

Total cholesterol, the levels of triglyceride and High-Density lipoprotein cholesterol (HDL-c) were assessed using an enzymatic colorimetric method by the commercial kits supplied by Centronic GmbH (Wartenberg, Germany). Serum low density lipoprotein cholesterol (LDL-C) was estimated using Friedwald's formula.

Fasting serum glucose was assessed using commercial kit (Spinreact, Spain), Fasting serum insulin was estimated by human insulin enzyme immunoassay test kit (Immunospec, USA). Insulin resistance was expressed as homeostasis model assessment of insulin resistance (HOMA-IR) higher than 2.5 (Reinehr et al., 2004).

2.3. Sample collection, DNA isolation and whole genome sequencing

Genomic DNA was extracted from blood with the GeneJET (Whole Blood) genomic DNA purification Kit (Thermo Scientific, USA), according to manufacturer's protocol. DNA Library preparation was carried out according to the manufacturer's instructions for sequencing on HiSeq2000 (Illumina, San Diego, CA, USA). For demultiplexing and conversion to FASTQ format, CASAVA 1.8.2 (Illumina) was used.

2.4. Bioinformatics analysis

2.4.1. Alignment of reads to reference

The sequences were mapped to the human reference genome (GRCh38/hg 38) BWA-MEM v.0.7.15 with default options (Li and Durbin, 2009). The SAM files were converted to BAM files using samtools 0.1.19 (Li et al., 2009). GATK v.3.5.0 (McKenna et al., 2010) was used for variant calling. Specifically, GATK v.3.5.0 UnifiedGenotyper with options -mbq 30 -stand_emit_conf 50 -stand_call_conf 50.

2.4.2. Annotation and functional analysis of variants

The variants were annotated using Annovar v.2018Apr16 (Wang et al., 2010) and SnpEff tool v4.3i (Cingolani et al., 2012). CRAVAT tool v.4 (Douville et al., 2013) was used for predicting mutation impact. It employs Variant Effect Scoring Tool (VEST) which classified mutations as pathogenic or benign. Furthermore, ClinVar database was used (Landrum et al., 2015) to predict the pathogenic mutations.

2.4.3. Identification of mitochondrial DNA haplogroups

The paired-end reads aligned to hg19 mitochondrial sequence were realigned to rCRS (Revised Cambridge Reference Sequence (Andrews et al., 1999) and then the variants were used to call haplogroups using HaploGrep software (Kloss-Brandstätter et al., 2011).

2.4.4. Admixture analysis

To visualize the heterogeneity of the Egyptian lineages we employed ADMIXTURE (v1.3.0) program (Alexander et al., 2009). We used a total of 80 human origin SNP panel (HOSP) (Lazaridis et al., 2014) contemporary representative genomes from Europe, West Asia, and various parts of Africa (four per population). The panel was merged with the Egyptian samples and pruned using PLINK (v.1.90) (Purcell et al., 2007) with the option '–indep-pairwise 200 25 0.4′. We explored the composition of the artificial ancestral populations (*K*) from two to five to provide the glimpse into the increasing Egyptian genomic complexity as *K* increases (supplementary figure S1), and chose $K = 4$ for the final interpretation because it showed a gradual transition of ancestral

component proportions from European and West Asian populations to North African as predicted by geographic location (supplementary figure S2). Out of the 9 Egyptian samples, sample EGY3 was excluded due to blood relations with another sample (EGY4) in the dataset to avoid bias.

3. Results

3.1. Clinical characteristics of the study group

The clinical characteristics of each subject are represented in supplementary Table S1. The mean value of BMI was 32.1 ± 11.28; ranged from (18–45.9), where five cases (aged 44–58 years old) were obese with BMI >30 kg/m^2. Four subjects (aged 20–35 years old) were classified as lean with BMI less than 25 kg/m^2. The mean level of total cholesterol among the research subjects was 192.7 ± 51.5. Three cases of hyperlipidemia (aged 29–58 years old) were observed; in one of the cases coinciding with hypertension. Regarding HOMA-IR, the mean value was 2.3 ± 1.59 and ranged from 0.9 to 5.08. Two out of the nine subjects (age 50 and 56 years old) were diagnosed with diabetes, and one had HOMA-IR 4.5 but without any confirmed diagnosis. Three out of the nine subjects (EGY3 age 35, EGY4 age 20, and EGY5 age 25 years old) weren't clinically diagnosed with any metabolic syndrome.

3.2. Genome sequencing and candidate gene identification

The nine Egyptian females represent different geographic regions in Egypt (supplementary Figure S2). The workflow of the sequence preprocessing, variant calling, annotation and analysis is summarized in eight steps (supplementary Figure S3.). We used an Illumina HiSeq 2500 short-read sequencing and generated approximately 3,099,0381,190 reads of 100 bps length that were aligned to the human reference genome GRCh38/hg 38, resulting in the average depth of coverage of 30x. The basic sequence data are summarized in supplementary Table S2.

From our data, we identified genes associated with obesity, T2D, and MS using the CRAVAT v.4 server and listed the ones with significant association (supplementary Table S3). Among them, we identified five candidate genes for obesity in the Egyptians; *CRHR1*, *VCP*, *NOTCH3*, *RHBG*, and *TMEM63A*. Variations in these genes were common within the subjects with BMI > 30 kg/m^2. *CRHR1* and *TMEM63A* were associated with T2D-diagnosed subjects. Also, *NUP88*, *CRHR1*, and *WISP1* were associated with high cholesterol trait. The gene *CRHR1* variations were found in both high cholesterol and T2D patients and therefore this gene is the most feasible obesity and obesity-related disease risk factor we can suggest based on these nine Egyptian females studied.

3.3. Variants analysis

We used CRAVAT v.4 server to identify the most significant SNPs in genes using ClinVar and variant effect scoring tool (VEST) score, (Carter et al., 2013) as well as phenotypes reported in the GWAS catalogue (MacArthur et al., 2016) which resulted in a total of 12 SNPs related to metabolic issues and obesity shared among the subjects. Some of them are associated with the pathological conditions confirmed in the Egyptian (EGY) subjects (supplementary Table S4). Among them, we found rs3733402, Ser143Asn in *KLKB1* that are associated with obesity according to GWAS catalog, and rs351855, Gly388Arg in *FGFR4* that are associated with Waist-to-hip ratio adjusted for body mass index (Rask-Andersen et al., 2019).

In the prehypertensive and hypertensive subjects (systolic and diastolic 130/90 &150/100) with high cholesterol, we found Asp131Glu in C1GALT1C1 (rs17261572, T > A), rs10065172, Leu105 in *IRGM* and rs1805010, Ile75Val in *IL4R* while in the prediabetic with HOMA-IR > 2.5 and diabetic participants, we found the following variants: rs1804495, Leu303Phe in *SERPINA7*, rs1800450, Gly54Asp in *MBL2*, rs1566734, Gln276Pro in *PTPRJ*, and rs61752717, Met694Val in *MEFV*.

3.4. Mitochondrial DNA analysis

The mitochondrial DNA (mtDNA) haplogroups were identified using and MitoSuite 1.0.9 and validated using Haplogrep 2.0 tools (Table 1) and showed diverse maternal ancestral backgrounds associated with Africa, Middle East, Europe or broadly - Mediterranean. The sample EGY1 and EGY2 revealed haplogroups associated with Mediterranean regions, M1a1b found in the north Mediterranean (Stevanovitch et al., 2004; Saunier et al., 2009; Winters, 2010; González et al., 2007) and T2f found in the European countries, respectively. The eastern Mediterranean region shares severe incidence of obesity with Egypt (Musaiger, 2004)showing a possible genetic linkage within these countries regarding metabolic profile. Whereas, EGY3 and EGY4 are related samples sharing haplogroup L2a1 and EGY5 had haplogroup M, in both cases, these haplogroups point to an African origin (29, 30) (Quintana-Murci et al., 1999). The strongest mtDNA ancestral connection to the Middle East was shown by EGY8 haplogroup H5 (Ennafaa et al., 2009), however, such origin could be attributed to EGY6 (H27e) and EGY7 (U8b1a2b) as well EGY8 (Achilli et al., 2004; Brotherton et al., 2013; Roostalu et al., 2006). EGY7 haplogroup showed a very specific distribution within Jordan and Italy (Turchi et al., 2008; Grignani et al., 2009; González et al., 2006); possibility of European or Middle Eastern admixture in this sample seems to be reflected in the ADMIXTURE analysis (supplementary Figure S1). man Haplogroup T1a7 found in the EGY9 is associated with very broad distribution within all of the geographic regions described regarding other EGY samples (Metspalu et al., 1999; Bonfiglio et al., 2012; Rowold et al., 2007). Moreover, mitochondrial genome of diabetic subjects and one of three cases of hyperlipidemia, EGY7, EGY9, and EGY2 contained T16189C mutation (Table 2) associated with insulin resistance and type 2 diabetes mellitus (Poulton et al., 1998, 2002; Park et al., 2008; Kim et al., 2002). Among the variants that were not related to obesity, we found that A4282G (Table 2) associated with chronic progressive external ophthalmoplegia (CPEO) was present in all samples (Jackson et al., 2014).

Table 1
List of Haplogroups and locations where they are common regarding each sample.

Sample	Haplogroup	Locations with high frequesncy	Lower resolution haplogroups and locations with high frequency
EGY1	M1a1b2	North Mediterranean (Pennarun et al., 2012)	M1 - eastern Africa/Ethiopia M1a1 – Mediterranean region (east Libya, Greece)
EGY2	T2f	Europe, Middle East (Pala et al., 2012)	T2 - Mediterranean and central and western Europe (high freq. in northern Italy) (Pala et al., 2012)
EGY3, EGY4	L2a1d1	Africans (Allard et al., 2005; Johnson et al., 2015)	L2a1 - Africans (Allard et al., 2005; Johnson et al., 2015)
EGY5	M3a1b	NA	M - northeast Africa (Quintana-Murci et al., 1999)
EGY6	H27e	NA	H - Europe, Middle East, Caucasus region (Achilli et al., 2004; Brotherton et al., 2013; Roostalu et al., 2006)
EGY7	U8b1a2b	NA	U8b – Italy, Jordan (Turchi et al., 2008; Grignani et al., 2009; González et al., 2006)
EGY8	H5	Middle East and Western Caucasus (Ennafaa et al., 2009)	H5 - Middle East and Western Caucasus (Ennafaa et al., 2009)
EGY9	T1a7	NA	T1 - Caucasus, Middle East, Europe, and North and East Africa (Metspalu et al., 1999; Bonfiglio et al., 2012; Rowold et al., 2007)

Table 2
List of mitochondrial disease associated variants in each sample.

	Mutations (disease associated)
EGY1	T195C, A4282G, A8108G, A10398G, G15043A, G16129A, A16183C and T16189C
EGY2	T4216C, A4282G, A4917G, A10398A and G15928A
EGY3	T195C, A4282G, A10398G T15784C, G16390A and T16519T
EGY4	T195C, A4282G, A10398G, T15784C, G16390A and 16,519
EGY5	A4282G, A10398G and G15043A
EGY6	A4282G, A10398A, T16093C and T16519T
EGY7	T195C, G3316A, A4282G, G9055A, A10398A, A11467G, A12308G, G12372A, G14831A, A16183C and T16189C
EGY8	A4282G, A10398A, C16192T and T16519T
EGY9	T4216C, A4282G, A4917G, A10398A, G15928A and T16189C

3.5. Admixture analysis

When $K = 4$ is chosen (Fig. 1), the major ancestral genomic components in Egyptians are dark green and dark blue. While the dark green varied between 81.10% and 92.43% among the Egyptian samples, the dark blue composed from 7.56% to 18.48% of the artificial Egyptian ancestry. Similar genomic trend was shown when the Tunisians, Algerians, Saharawi and the Bedouin samples used, which points to a strong uniform North African genomic signature, despite of nomadic lifestyle practiced by some of these populations (Losleben, 2003; Catassi et al., 1999) which could have resulted in a high admixture rate. However, the components themselves seem to have originated from two main sources: the dark green from the Southern Europe as well as the Middle East and the dark blue mainly from the West Africa as it was found in the highest proportions in the Gambian, Essan, and Yoruba populations. Besides the European/Middle Eastern and West African components small amount and highly variable (from 0.00% to 1.05%) East-African-associated component (soft orange) (supplementary Table S5) was found in highest proportions in EGY9, EGY6 and EGY5 among the Egyptian samples. This component had the strongest association with the Hadza indigenous people, however, in smaller proportions was also found among North-Eastern (Ethiopia) and East-Central (Dinka) African populations potentially reflecting genetic drift across the African continent. Most of the Egyptian samples used, especially, EGY5 and EGY9 are closely resembled with Tunisian ancestral proportions while EGY7, EGY8 and EGY2 exhibited higher proportions of the green (Southern European/West Asian) component which were common among the Bedouin.

4. Discussion

There is a limited amount of studies reporting complete genomes using NGS technology from different ethnic groups in North Africa (ElHefnawi et al., 2018; Ilyas et al., 2015; Azim et al., 2013; Thareja et al., 2015), especially, together with clinical data. We detected the variants of the candidate genes, *GCAT*, *SLC6A13*, *TESK2*, *VCP*, *PCNX*, and *NUP88* using the CRAVAT v.4 server. Also, we found four already reported variants in *CRHR1* among the subjects who are suffering from severe obesity and metabolic syndrome. Rankinen et al. (2005) demonstrated that *CRHR1* caused human obesity by single-gene mutations. Licinio et al. (2004) and Liu et al. (2007). Moreover, the participants suffering from obesity and obesity related diseases, (EGY2, EGY6, EGY7, and EGY9) had a SNP rs3733402, 428G > A (Ser143Asn) in *KLKB1*. This amino acid change has been associated with a Prekallikrein deficiency (Katsuda et al., 2007). Also, we found Gly388Arg in *FGFR4* gene (rs351855, 1162G > A) in the participants EGY1, EGY6, and EGY8, which is associated with Waist-to-hip ratio adjusted for body mass index (Rask-Andersen et al., 2019). Moreover, four SNPs were found in prediabetic subject with HOMA-IR > 2.5 (EGY6), and two diabetic subjects (EGY7, EGY9). One of these SNPs was in *MEFV*, 2080A > G, Met694Val, rs61752717 that is associated with Familial Mediterranean fever (Shohat and Halpern, 2011). It is reported that *MEFV* gene variations affect mostly populations living in the Mediterranean region, especially North African Jews, Armenians, Turks, and Arabs and were also reported in a sequenced Bedouin genome (Mohammadnejad and Farajnia, 2013).

The mtDNA analysis in the presented study revealed haplogroup for each subject as well as pathogenic mutations that can be transferred from mother to children. A mtDNA T16189C mutation observed in EGY7, EGY9, and EGY2, is already linked with insulin resistance and type 2 diabetes mellitus from the pooled case-manipulate research comparing the linkage of T16189C polymorphism with the threat of most cancers and T2DM progression (Kumari et al., 2018). Regarding other pathogenic mutations, each subject of the nine samples shared mutation A4282G (Table 2) related to chronic progressive external ophthalmoplegia (CPEO Plus) which is a complex slowly progressing ocular disease that may begin at any age and progresses for fifteen years and often manifests in older age (Lv et al., 2017). CPEO can be caused by large-scale single mtDNA deletions (Jackson et al., 2014; Goto et al., 1990; López-Gallardo et al., 2009) and mt-tRNA point mutations (Seibel et al., 1994; Taylor et al., 1998) (one of them m.4282G > A), or more often, because of multiple mtDNA deletions, which may arise from already compromised mtDNA maintenance (Hanisch et al., 2015; Van Goethem et al., 2003). Even though this variant may seem outside the scope of our study, we suggest it as a possible direction for the future large-scale studies, which could confirm the A4282G variant pathogenicity, prevalence and co-occurrence with other CPEO risk variants in the Egyptian population.

Overall, both mtDNA and ADMIXTURE studies suggested a diverse and admixed background of Egyptians sharing genetic signature and metabolic phenotypes with other Mediterranean nations. Although largely debated, into-Africa migration hypothesis (Haber et al., 2016; Busby et al., 2016; Llorente et al., 2015) seems to be supported by our largely diverse mtDNA haplotype collection demonstrated by just nine random samples. Egypt standing at the crossroad of different continents could have served as a gateway that allowed circulating people from Eurasia towards Africa since at least Holoscene (Haber et al., 2016) and lead to admixtures in east and sub-Saharan Africa that were reported before (Busby et al., 2016) and in this case lead to shared genetic predisposition to complex diseases such as obesity.

Our data can be a useful additional genomics resource for the greater

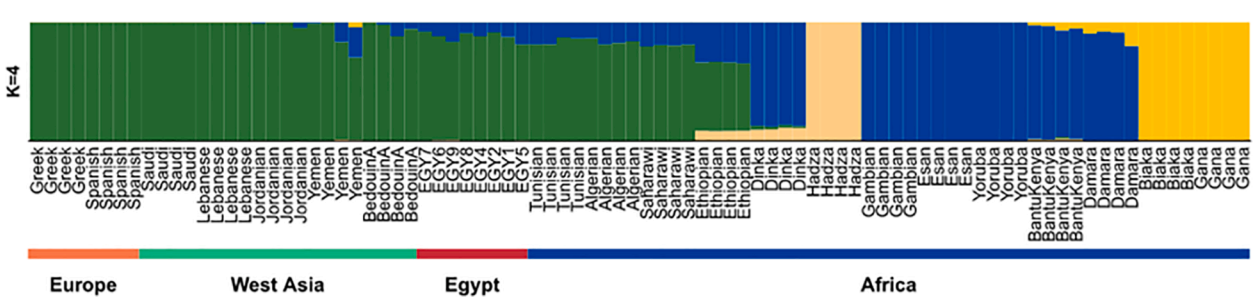

Fig. 1. The ancestral composition of Egyptians in comparison to other populations when K = 4.

genomic research of the nothern Africa and Middle East regions. The haplotype and admixture results point to the diversity of Egyptians, and to a non-African origin for most of them, pointing to the into Africa influx theory.

5. Conclusions

Here, we provide nine additional Egyptian genomic resources in terms of short-read based whole genome data and SNPs linked to metabolism-related phenotypes. These nine genomes are a valuable national resource as it contains matching clinical information for each of the nine participants. It can facilitate further larger-scale in depth research yielding more relevant variants and greater statistical power.

CRediT authorship contribution statement

Mahmoud ElHefnawi: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing - review & editing. **Elsayed Hegazy:** Data curation, Formal analysis, Investigation, Software. **Asmaa Elfiky:** Data curation, Methodology, Resources, Writing - original draft, Writing - review & editing. **Yeonsu Jeon:** Formal analysis, Software, Writing - review & editing. **Sungwon Jeon:** Formal analysis, Investigation, Software, Validation, Visualization, Writing - review & editing. **Jong Bhak:** Formal analysis, Funding acquisition, Project administration, Supervision, Writing - review & editing. **Fateheya Mohamed Metwally:** Resources, Writing - review & editing. **Sumio Sugano:** Methodology. **Terumi Horiuchi:** Methodology. **Abe Kazumi:** Methodology. **Asta Blazyte:** Data curation, Formal analysis, Investigation, Software, Validation, Visualization, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The datasets generated in this study are available in the NCBI Sequence Read Archive repository; under accession number SRP136979 “https://www.ncbi.nlm.nih.gov/sra/?term=SRP136979“.

Acknowledgements

We thank the National Research Centre clinic for providing for the samples. We thank the pan Asian population genomics initiative (PAPGI) for support with sequencing. This work was supported by U-K BRAND Research Fund (1.190007.01) of UNIST; Research Project Funded by Ulsan City Research Fund (1.190033.01) of UNIST.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2020.145237.

References

Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A., Scozzari, R., Cruciani, F., Zeviani, M., Briem, E., Carelli, V., 2004. The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. Am. J. Hum. Genet. 75, 910–918.
Alebshehy, R., Shuaib, N.M., Mbako, J.D., Barffo, D., KuuzagrNuotol, R., 2016. Determinant analysis of obesity among adult females in Egypt. Egypt. J. Hospital Med. 31, 1–8.
Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19, 1655–1664.
Allard, M.W., Polanskey, D., Miller, K., Wilson, M.R., Monson, K.L., Budowle, B., 2005. Characterization of human control region sequences of the African American SWGDAM forensic mtDNA data set. Forensic Sci. Int. 148, 169–179.
Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., Howell, N., 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat. Genet. 23, 147.
Austin, A.M., Hill, A.G., Fawzi, W.W., 2013. Maternal obesity trends in Egypt 1995–2005. Maternal Child Nutr. 9, 167–179.
Azim, M.K., Yang, C., Yan, Z., Choudhary, M.I., Khan, A., Sun, X., Li, R., Asif, H., Sharif, S., Zhang, Y., 2013. Complete genome sequencing and variant analysis of a Pakistani individual. J. Hum. Genet. 58, 622–626.
Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A., Shendure, J., 2011. Exome sequencing as a tool for Mendelian disease gene discovery. Nat. Rev. Genet. 12, 745.
Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R., 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature 53–59.
Biro, F.M., Wien, M., 2010. Childhood obesity and adult morbidities. Am. J. Clin. Nutr. 91, 1499S–1505S.
Bonfiglio, S., Ginja, C., De Gaetano, A., Achilli, A., Olivieri, A., Colli, L., Tesfaye, K., Agha, S.H., Gama, L.T., Cattonaro, F., 2012. Origin and spread of Bos taurus: new clues from mitochondrial genomes belonging to haplogroup T1. PLoS ONE 7, e38601.
Brotherton, P., Haak, W., Templeton, J., Brandt, G., Soubrier, J., Adler, C.J., Richards, S. M., Der Sarkissian, C., Ganslmeier, R., Friederich, S., 2013. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. Nat. Commun. 4, 1764.
Busby, G.B., Band, G., Le, Q.S., Jallow, M., Bougama, E., Mangano, V.D., Amenga-Etego, L.N., Enimil, A., Apinjoh, T., Ndila, C.M., 2016. Admixture into and within sub-Saharan Africa. Elife 5, e15266.
Carter, H., Douville, C., Stenson, P.D., Cooper, D.N., Karchin, R., 2013. Identifying Mendelian disease genes with the variant effect scoring tool. BMC Genomics 14, S3.
Catassi, C., Ratsch, I.-M., Gandolfi, L., Pratesi, R., Fabiani, E., El Asmar, R., Frijia, M., Bearzi, I., Vizzoni, L., 1999. Why is coeliac disease endemic in the people of the Sahara? The Lancet 354, 647–648.
Choquet, H., Meyre, D., 2011. Genetics of obesity: what have we learned? Curr. Genomics 12, 169–179.
Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly 6, 80–92.
Dehghan, M., Akhtar-Danesh, N., Merchant, A.T., 2005. Childhood obesity, prevalence and prevention. Nutr. J. 4, 24.
Douville, C., Carter, H., Kim, R., Niknafs, N., Diekhans, M., Stenson, P.D., Cooper, D.N., Ryan, M., Karchin, R., 2013. CRAVAT: cancer-related analysis of variants toolkit. Bioinformatics 29, 647–648.
ElHefnawi, M., Jeon, S., Bhak, Y., ElFiky, A., Horaiz, A., Jun, J., Kim, H., Bhak, J., 2018. Whole genome sequencing and bioinformatics analysis of two Egyptian genomes. Gene 668, 129–134.
Ennafaa, H., Cabrera, V.M., Abu-Amero, K.K., González, A.M., Amor, M.B., Bouhaha, R., Dzimiri, N., Elgaaïed, A.B., Larruga, J.M., 2009. Mitochondrial DNA haplogroup H structure in North Africa. BMC Genet. 10, 8.
Frayling, T.M., Timpson, N.J., Weedon, M.N., Zeggini, E., Freathy, R.M., Lindgren, C.M., Perry, J.R., Elliott, K.S., Lango, H., Rayner, N.W., 2007. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 316, 889–894.
González, A.M., García, O., Larruga, J.M., Cabrera, V.M., 2006. The mitochondrial lineage U8a reveals a Paleolithic settlement in the Basque country. BMC Genomics 7, 124.
González, A.M., Larruga, J.M., Abu-Amero, K.K., Shi, Y., Pestano, J., Cabrera, V.M., 2007. Mitochondrial lineage M1 traces an early human backflow to Africa. BMC Genomics 8, 223.
Goto, Y.-I., Koga, Y., Horai, S., Nonaka, I., 1990. Chronic progressive external ophthalmoplegia: a correlative study of mitochondrial DNA deletions and their phenotypic expression in muscle biopsies. J. Neurol. Sci. 100, 63–69.
Grignani, P., Turchi, C., Achilli, A., Peloso, G., Alù, M., Ricci, U., Robino, C., Pelotti, S., Carnevali, E., Boschi, I., 2009. Multiplex mtDNA coding region SNP assays for molecular dissection of haplogroups U/K and J/T. Forensic Sci. Int. Genet. 4, 21–25.
Haber, M., Mezzavilla, M., Bergström, A., Prado-Martinez, J., Hallast, P., Saif-Ali, R., Al-Habori, M., Dedoussis, G., Zeggini, E., Blue-Smith, J., 2016. Chad genetic diversity reveals an African history marked by multiple Holocene Eurasian migrations. Am. J. Hum. Genetics 99, 1316–1324.
Hanisch, F., Kornhuber, M., Alston, C.L., Taylor, R.W., Deschauer, M., Zierz, S., 2015. SANDO syndrome in a cohort of 107 patients with CPEO and mitochondrial DNA deletions. J. Neurol. Neurosurg. Psychiatry 86, 630–634.
Harper, M.G., 2006. Childhood obesity: strategies for prevention. Family Commun. Health 29, 288–298.
Hassan, N.E., El-Masry, S.A., Zarouk, W., El Banna, R.A., Mosaad, R.M., Al-Tohamy, M., Salamah, A.R., 2017. Obesity phenotype in relation to gene polymorphism among samples of Egyptian children and their mothers. Genes Dis. 5, 150–157.
Ilyas, M., Kim, J.-S., Cooper, J., Shin, Y.-A., Kim, H.-M., Cho, Y.S., Hwang, S., Kim, H., Moon, J., Chung, O., 2015. Whole genome sequencing of an ethnic Pathan (Pakhtun) from the north-west of Pakistan. BMC Genomics 16, 172.
Jackson, C.B., Neuwirth, C., Hahn, D., Nuoffer, J.M., Frank, S., Gallati, S., Schaller, A., 2014. Novel mitochondrial tRNA(Ile) m.4282A>G gene mutation leads to chronic progressive external ophthalmoplegia plus phenotype. Br. J. Ophthalmol. 98, 1453–1459.

Jackson, C.B., Neuwirth, C., Hahn, D., Nuoffer, J.-M., Frank, S., Gallati, S., Schaller, A., 2014. Novel mitochondrial tRNAIle m. 4282A> G gene mutation leads to chronic progressive external ophthalmoplegia plus phenotype. Br. J. Ophthalmol. 98, 1453–1459.
Johnson, D.C., Shrestha, S., Wiener, H.W., Makowsky, R., Kurundkar, A., Wilson, C.M., Aissani, B., 2015. Mitochondrial DNA diversity in the African American population. Mitochondrial DNA 26, 445–451.
Katsuda, I., Maruyama, F., Ezaki, K., Sawamura, T., Ichihara, Y., 2007. A new type of plasma prekallikrein deficiency associated with homozygosity for Gly104Arg and Asn124Ser in apple domain 2 of the heavy-chain region. Eur. J. Haematol. 79, 59–68.
Kim, J.-H., Park, K.S., Cho, Y.M., Kang, B., Kim, S., Jeon, H., Kim, S., Lee, H., 2002. The prevalence of the mitochondrial DNA 16189 variant in non-diabetic Korean adults and its association with higher fasting glucose and body mass index. Diabet. Med. 19, 681–684.
Kleinendorst, L., Massink, M.P., Cooiman, M.I., Savas, M., van der Baan-Slootweg, O.H., Roelants, R.J., Janssen, I.C., Meijers-Heijboer, H.J., Knoers, N.V., van Amstel, H.K.P., 2018. Genetic obesity: next-generation sequencing results of 1230 patients with obesity. J. Med. Genet. 55, 578–586.
Kloss-Brandstätter, A., Pacher, D., Schönherr, S., Weissensteiner, H., Binna, R., Specht, G., Kronenberg, F., 2011. HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. Hum. Mutat. 32, 25–32.
Kumari, T., Vachher, M., Bansal, S., Bamezai, R.N., Kumar, B., 2018. Meta-analysis of mitochondrial T16189C polymorphism for cancer and Type 2 diabetes risk. Clin. Chim. Acta 482, 136–143.
Landrum, M.J., Lee, J.M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Hoover, J., 2015. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 44, D862–D868.
Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P. H., Schraiber, J.G., Castellano, S., Lipson, M., 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature 513, 409.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079.
Licinio, J., O'kirwan, F., Irizarry, K., Merriman, B., Thakur, S., Jepson, R., Lake, S., Tantisira, K., Weiss, S., Wong, M., 2004. Association of a corticotropin-releasing hormone receptor 1 haplotype and antidepressant treatment response in Mexican-Americans. Mol. Psychiatry 9, 1075.
Liu, Z., Zhu, F., Wang, G., Xiao, Z., Tang, J., Liu, W., Wang, H., Liu, H., Wang, X., Wu, Y., 2007. Association study of corticotropin-releasing hormone receptor1 gene polymorphisms and antidepressant response in major depressive disorders. Neurosci. Lett. 414, 155–158.
Llorente, M.G., Jones, E.R., Eriksson, A., Siska, V., Arthur, K., Arthur, J., Curtis, M., Stock, J.T., Coltorti, M., Pieruccini, P., 2015. Ancient Ethiopian genome reveals extensive Eurasian admixture in Eastern Africa. Science 350, 820–822.
Loos, R.J., Lindgren, C.M., Li, S., Wheeler, E., Zhao, J.H., Prokopenko, I., Inouye, M., Freathy, R.M., Attwood, A.P., Beckmann, J.S., 2008. Common variants near MC4R are associated with fat mass, weight and risk of obesity. Nat. Genet. 40, 768.
López-Gallardo, E., López-Pérez, M.J., Montoya, J., Ruiz-Pesini, E., 2009. CPEO and KSS differ in the percentage and location of the mtDNA deletion. Mitochondrion 9, 314–317.
Losleben, E., 2003. The Bedouin of the Middle East. Lerner Publications.
Lv, Z.-Y., Xu, X.-M., Cao, X.-F., Wang, Q., Sun, D.-F., Tian, W.-J., Yang, Y., Wang, Y.-Z., Hao, Y.-L., 2017. Mitochondrial mutations in 12S rRNA and 16S rRNA presenting as chronic progressive external ophthalmoplegia (CPEO) plus: A case report. Medicine 96.
MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., Junkins, H., McMahon, A., Milano, A., Morales, J., 2016. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. 45, D896–D901.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303.
Metspalu, E., Kivisild, T., Kaldma, K., Parik, J., Reidla, M., Tambets, K., Villems, R., 1999. In: The Trans-Caucasus and the Expansion of the Caucasoid-specific Human Mitochondrial DNA. Springer, Genomic Diversity, pp. 121–133.
Metzker, M.L., 2010. Sequencing technologies—the next generation. Nat. Rev. Genet. 11, 31.
Mohammadnejad, L., Farajnia, S., 2013. Mediterranean Fever gene analysis in the azeri turk population with familial mediterranean Fever: evidence for new mutations associated with disease. Cell journal 15, 152–159.
Muhsin, N.I.A., Bentley, L., Bai, Y., Goldsworthy, M., Cox, R.D., 2020. A novel mutation in the mouse Pcsk1 gene showing obesity and diabetes. Mamm. Genome 31, 17–29.
Musaiger, A.O., 2004. Overweight and obesity in the Eastern Mediterranean Region: can we control it?.
Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., Mullany, E.C., Biryukov, S., Abbafati, C., Abera, S.F., Abraham, J.P., Abu-Rmeileh, N.M.E., Achoki, T., AlBuhairan, F.S., Alemu, Z.A., Alfonso, R., Ali, M.K., Ali, R., Guzman, N. A., Ammar, W., Anwari, P., Banerjee, A., Barquera, S., Basu, S., Bennett, D.A., Bhutta, Z., Blore, J., Cabral, N., Nonato, I.C., Chang, J.-C., Chowdhury, R., Courville, K.J., Criqui, M.H., Cundiff, D.K., Dabhadkar, K.C., Dandona, L., Davis, A., Dayama, A., Dharmaratne, S.D., Ding, E.L., Durrani, A.M., Esteghamati, A., Farzadfar, F., Fay, D.F.J., Feigin, V.L., Flaxman, A., Forouzanfar, M.H., Goto, A., Green, M.A., Gupta, R., Hafezi-Nejad, N., Hankey, G.J., Harewood, H.C., Havmoeller, R., Hay, S., Hernandez, L., Husseini, A., Idrisov, B.T., Ikeda, N., Islami, F., Jahangir, E., Jassal, S.K., Jee, S.H., Jeffreys, M., Jonas, J.B., Kabagambe, E.K., Khalifa, S.E.A.H., Kengne, A.P., Khader, Y.S., Khang, Y.-H., Kim, D., Kimokoti, R.W., Kinge, J.M., Kokubo, Y., Kosen, S., Kwan, G., Lai, T., Leinsalu, M., Li, Y., Liang, X., Liu, S., Logroscino, G., Lotufo, P.A., Lu, Y., Ma, J., Mainoo, N.K., Mensah, G.A., Merriman, T.R., Mokdad, A.H., Moschandreas, J., Naghavi, M., Naheed, A., Nand, D., Narayan, K.M.V., Nelson, E.L., Neuhouser, M.L., Nisar, M.I., Ohkubo, T., Oti, S.O., Pedroza, A., et al., 2014. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet 384, 766–781.
Nordang, G.B., Busk, Ø.L., Tveten, K., Hanevik, H.I., Fell, A.K.M., Hjelmesæth, J., Holla, Ø.L., Hertel, J.K., 2017. Next-generation sequencing of the monogenic obesity genes LEP, LEPR, MC4R, PCSK1 and POMC in a Norwegian cohort of patients with morbid obesity and normal weight controls. Mol. Genet. Metab. 121, 51–56.
Organization, W.H., 2000. Obesity: Preventing and Managing the Global Epidemic. World Health Organization.
Pala, M., Olivieri, A., Achilli, A., Accetturo, M., Metspalu, E., Reidla, M., Tamm, E., Karmin, M., Reisberg, T., Kashani, B.H., 2012. Mitochondrial DNA signals of late glacial recolonization of Europe from near eastern refugia. Am. J. Hum. Genet. 90, 915–924.
Park, K.S., Chan, J., Chuang, L.-M., Suzuki, S., Araki, E., Nanjo, K., Ji, L., Ng, M., Nishi, M., Furuta, H., 2008. A mitochondrial DNA variant at position 16189 is associated with type 2 diabetes mellitus in Asians. Diabetologia 51, 602–608.
Pennarun, E., Kivisild, T., Metspalu, E., Metspalu, M., Reisberg, T., Moisan, J.-P., Behar, D.M., Jones, S.C., Villems, R., 2012. Divorcing the Late Upper Palaeolithic demographic histories of mtDNA haplogroups M1 and U6 in Africa. BMC Evol. Biol. 12, 234.
Poulton, J., Brown, M.S., Cooper, A., Marchington, D., Phillips, D., 1998. A common mitochondrial DNA variant is associated with insulin resistance in adult life. Diabetologia 41, 54–58.
Poulton, J., Bednarz, A., Scott-Brown, M., Thompson, C., Macaulay, V., Simmons, D., 2002. The presence of a common mitochondrial DNA variant is associated with fasting insulin levels in Europeans in Auckland. Diabet. Med. 19, 969–971.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.
Quintana-Murci, L., Semino, O., Bandelt, H.-J., Passarino, G., McElreavey, K., Santachiara-Benerecetti, A.S., 1999. Genetic evidence of an early exit of Homo sapiens sapiens from Africa through eastern Africa. Nat. Genet. 23, 437.
Rankinen, T., Zuberi, A., Chagnon, Y.C., Weisnagel, S.J., Argyropoulos, G., Walts, B., Pérusse, L., Bouchard, C., 2005. The human obesity gene map: the update. Obesity 14 (2006), 529–644.
Rask-Andersen, M., Karlsson, T., Ek, W.E., Johansson, Å., 2019. Genome-wide association study of body fat distribution identifies adiposity loci and sex-specific genetic effects. Nat. Commun. 10, 339.
Reinehr, T., Kiess, W., Kapellen, T., Andler, W., 2004. Insulin sensitivity among obese children and adolescents, according to degree of weight loss. Pediatrics 114, 1569–1573.
Roostalu, U., Kutuev, I., Loogväli, E., Metspalu, E., Tambets, K., Reidla, M., Khusnutdinova, E., Usanga, E., Kivisild, T., Villems, R., 2006. Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. Mol. Biol. Evol. 24, 436–448.
Rowold, D., Luis, J., Terreros, M., Herrera, R.J., 2007. Mitochondrial DNA geneflow indicates preferred usage of the Levant Corridor over the Horn of Africa passageway. J. Hum. Genet. 52, 436.
Sällman Almén, M., Rask-Andersen, M., Jacobsson, J.A., Ameur, A., Kalnina, I., Moschonis, G., Juhlin, S., Bringeland, N., Hedberg, L.A., Ignatovica, V., Chrousos, G. P., Manios, Y., Klovins, J., Marcus, C., Gyllensten, U., Fredriksson, R., Schiöth, H.B., 2013. Determination of the obesity-associated gene variants within the entire FTO gene by ultra-deep targeted sequencing in obese and lean children. Int. J. Obesity 37, 424–431.
Saunier, J.L., Irwin, J.A., Strouss, K.M., Ragab, H., Sturk, K.A., Parsons, T.J., 2009. Mitochondrial control region sequences from an Egyptian population sample. Forensic Sci. Int. Genet. 3, e97–e103.
Seibel, P., Lauber, J., Klopstock, T., Marsac, C., Kadenbach, B., Reichmann, H., 1994. Chronic progressive external ophthalmoplegia is associated with a novel mutation in the mitochondrial tRNAAsn gene. Biochem. Biophys. Res. Commun. 204, 482–489.
Shohat, M., Halpern, G.J., 2011. Familial Mediterranean fever–a review. Genet. Med. 13, 487–498.
Stevanovitch, A., Gilles, A., Bouzaid, E., Kefi, R., Paris, F., Gayraud, R., Spadoni, J., El-Chenawi, F., Béraud-Colomb, E., 2004. Mitochondrial DNA sequence diversity in a sedentary population from Egypt. Ann. Hum. Genet. 68, 23–39.
Sudo, Y., Ezura, Y., Kajita, M., Yoshida, H., Suzuki, T., Hosoi, T., Inoue, S., Shiraki, M., Ito, H., Emi, M., 2005. Association of single nucleotide polymorphisms in the promoter region of the pro-opiomelanocortin gene (POMC) with low bone mineral density in adult women. J. Hum. Genet. 50, 235–240.
Taylor, R.W., Chinnery, P.F., Bates, M.J., Jackson, M.J., Johnson, M.A., Andrews, R.M., Turnbull, D.M., 1998. A Novel Mitochondrial DNA Point Mutation in the tRNAIleGene: Studies in a Patient Presenting with Chronic Progressive External Ophthalmoplegia and Multiple Sclerosis. Biochem. Biophys. Res. Commun. 243, 47–51.
Thareja, G., John, S.E., Hebbar, P., Behbehani, K., Thanaraj, T.A., Alsmadi, O., 2015. Sequence and analysis of a whole genome from Kuwaiti population subgroup of Persian ancestry. BMC Genomics 16, 92.

Trevino, R.P., Marshall, R.M., Hale, D.E., Rodriguez, R., Baker, G., Gomez, J., 1999. Diabetes risk factors in low-income Mexican-American children. Diabetes Care 22, 202–207.
Turchi, C., Buscemi, L., Previderè, C., Grignani, P., Brandstätter, A., Achilli, A., Parson, W., Tagliabracci, A., 2008. Italian mitochondrial DNA database: results of a collaborative exercise and proficiency testing. Int. J. Legal Med. 122, 199–204.
Van Goethem, G., Martin, J.-J., Van Broeckhoven, C., 2003. Progressive external ophthalmoplegia characterized by multiple deletions of mitochondrial DNA. NeuroMol. Med. 3, 129–146.
Vimaleswaran, K.S., Loos, R.J., 2010. Progress in the genetics of common obesity and type 2 diabetes. Expert Rev. Mol. Med. 12.
Volckmar, A.-L., Han, C.T., Pütter, C., Haas, S., Vogel, C.I., Knoll, N., Struve, C., Göbel, M., Haas, K., Herrfurth, N., 2016. Analysis of genes involved in body weight regulation by targeted re-sequencing. PLoS ONE 11.
Wang, K., Li, M., Hakonarson, H., 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38, e164.
Wheeler, D.A., Srinivasan, M., Egholm, M., Shen, Y., Chen, L., McGuire, A., He, W., Chen, Y.-J., Makhijani, V., Roth, G.T., 2008. The complete genome of an individual by massively parallel DNA sequencing. Nature 452, 872–876.
Winters, C., 2010. The African origin of mtDNA haplogroup M1. Europe 17, 14.15.