

## Precision medicine of monogenic disorders: Lessons learned from the Saudi human genome

Malak Abedalthagafi<sup>1,2</sup>

<sup>1</sup>Saudi Human Genome Project, King Fahad Medical City and King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia, <sup>2</sup>Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

### TABLE OF CONTENTS

1. Abstract
2. Introduction
  - 2.1. National human genome projects
  - 2.2. SHGP sequencing approaches
  - 2.3. Intellectual disability
  - 2.4. Other CNS diseases
  - 2.5. Autoimmune disease
  - 2.6. Kidney disease
  - 2.7. Hereditary cancer syndromes
  - 2.8. Congenital anomalies
  - 2.9. Muscle, joint and developmental limb disorders
  - 2.10. Other diseases
3. Summary
4. Acknowledgements
5. References

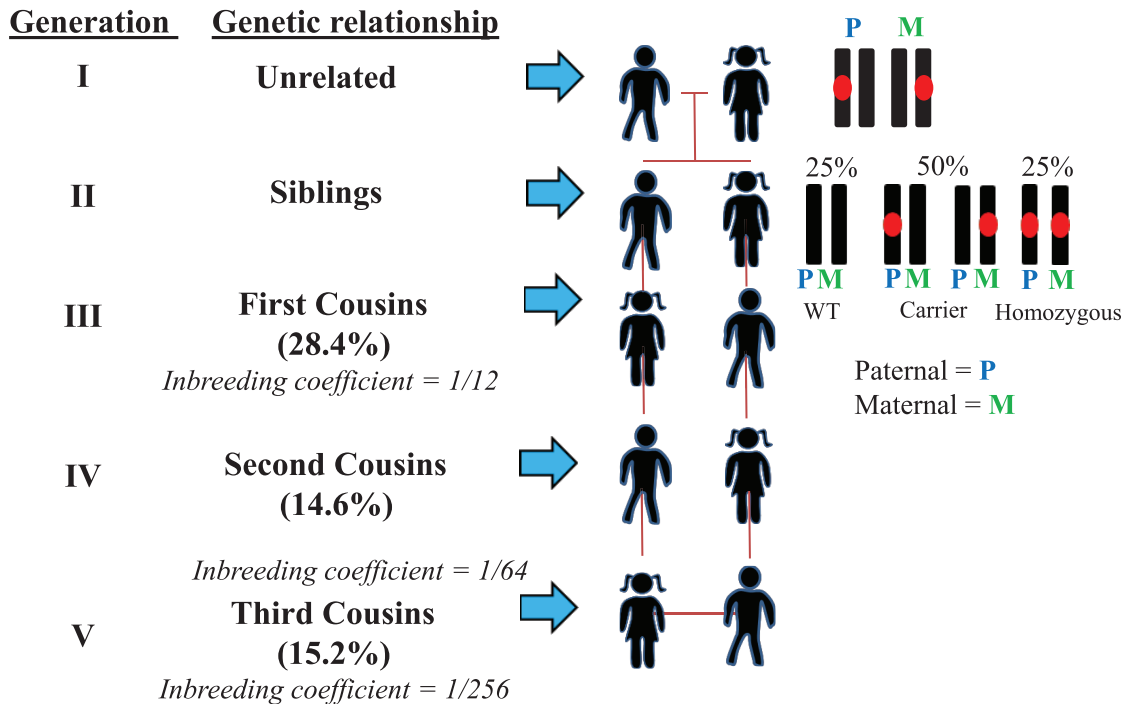
### 1. ABSTRACT

Congenital and genetic disorders cause many diseases in Arab countries due to large family sizes and high levels of inbreeding. Saudi Arabia (SA) has the highest consanguinity rates among Middle Eastern countries (~60% of all marriages) and is burdened by the highest number of genetic diseases. Genetic diseases can be life-threatening, often manifesting early in life. Approximately 8% of births in SA are affected, and more common genetic diseases, such as metabolic disease and cancer, manifest later in life in up to 20% of the population. This represents a massive healthcare burden to SA hospitals. The number of genetic disorders in the human population ranges from 7000 to 8000, over 3000 of which are caused by unknown mutations. In 2013, SA initiated the Saudi Human Genome Program (SHGP), which aims to sequence over 100,000 human genomes, with the goal of identifying strategies to discover, prevent, diagnose and treat genetic disorders through precision therapy. High-technology

genomics and informatic-based centers that exploit next-generation sequencing (NGS) have now identified mutations underlying many unexplained diseases.

### 2. INTRODUCTION

The Saudi Human Genome Program (SHGP) was designed to identify mutations in genes that predispose individuals to disease (1-3). The abundance of genetic diseases in SA combined with large family sizes makes it easier to identify the genes/mutations that underlie a particular disease (Figure 1) (4). With this knowledge, preventative counseling can be planned, or a regimen of rational precision-based therapies can be formulated (5, 6). In addition, the growing catalogue of disease-causing mutations allows premarital DNA tests for prospective partners, reducing the rates of afflicted offspring (Figure 1). In many cases, the disease-causing genes and

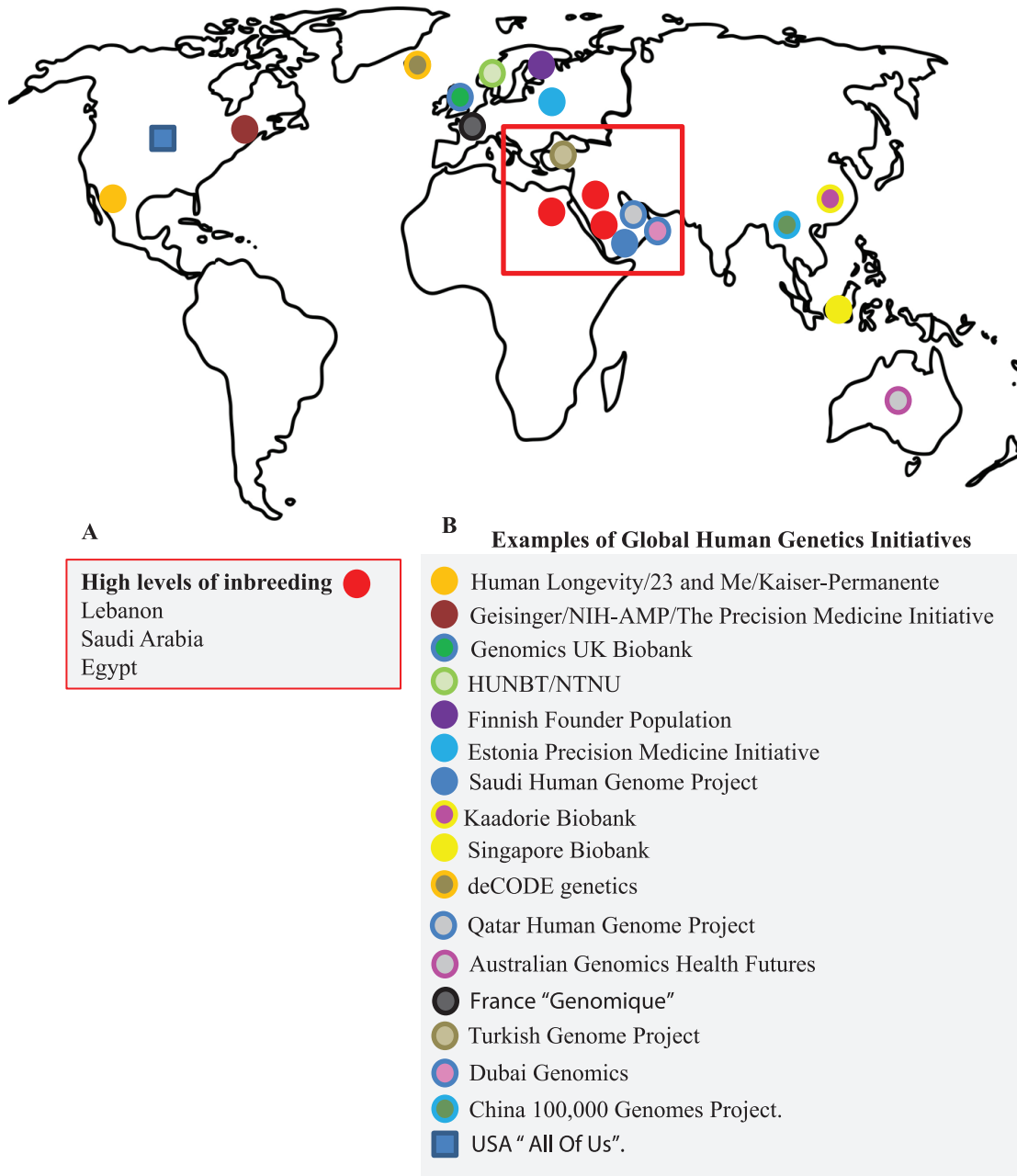


**Figure 1.** Genetic relationships and the rates of consanguinity in SA. Studies suggest that 57.7% of the families screened in SA are consanguineous (9, 55). The most frequent consanguineous marriages are between first cousins (28.4%), followed by distant relatives (15.2%) and second cousins (14.6%). These rates can differ among provinces. The offspring of consanguineous unions will be at an increased risk of genetic disorders (red dots) because of the expression of autosomal recessive gene mutations inherited from a common ancestor. The inbreeding coefficient denotes the probability that an individual is homozygous for an ancestral allele by inheritance.

gene variants identified in the SA population can be used to verify results obtained by research groups outside of SA that draw on conclusions from fewer case studies (7). Thus, despite its national nature, the project has the potential to benefit the global fight against inherited genetic diseases, but further improvements are still required (Figure 2).

The SHGP was launched in December 2013 in the city of Riyadh, at the location of the King Abdulaziz City for Science and Technology (KACST). The SHGP was funded and organized by the KACST, and it involves a national network of seven genome centers that were designed to recruit subjects and to carry out sequencing. The core technology in genome centers is next-generation sequencing (NGS), which has enabled efficient and cost-effective readings of entire human genomes. Over the

past decade, NGS has exponentially increased our understanding of genetic Mendelian diseases by permitting the analysis of multiple genomic regions in a single reaction (8). NGS has been utilized with success in the clinic, with a reported diagnostic yield as high as ~25% (9). The use of NGS enables the generation of large volumes of sequencing data in hours, representing a large advance over the human genome project that required years to be completed at a cost of 3 billion US dollars. NGS has proven to be a powerful application for the detection of disease-causing mutations, particularly in patients with monogenic disorders. While NGS has benefited the field of human genetics as a whole, the study of rare genetic diseases has perhaps witnessed the greatest advances (8). This permitted the genetic diagnosis of disorders for which the molecular mechanisms leading to disease

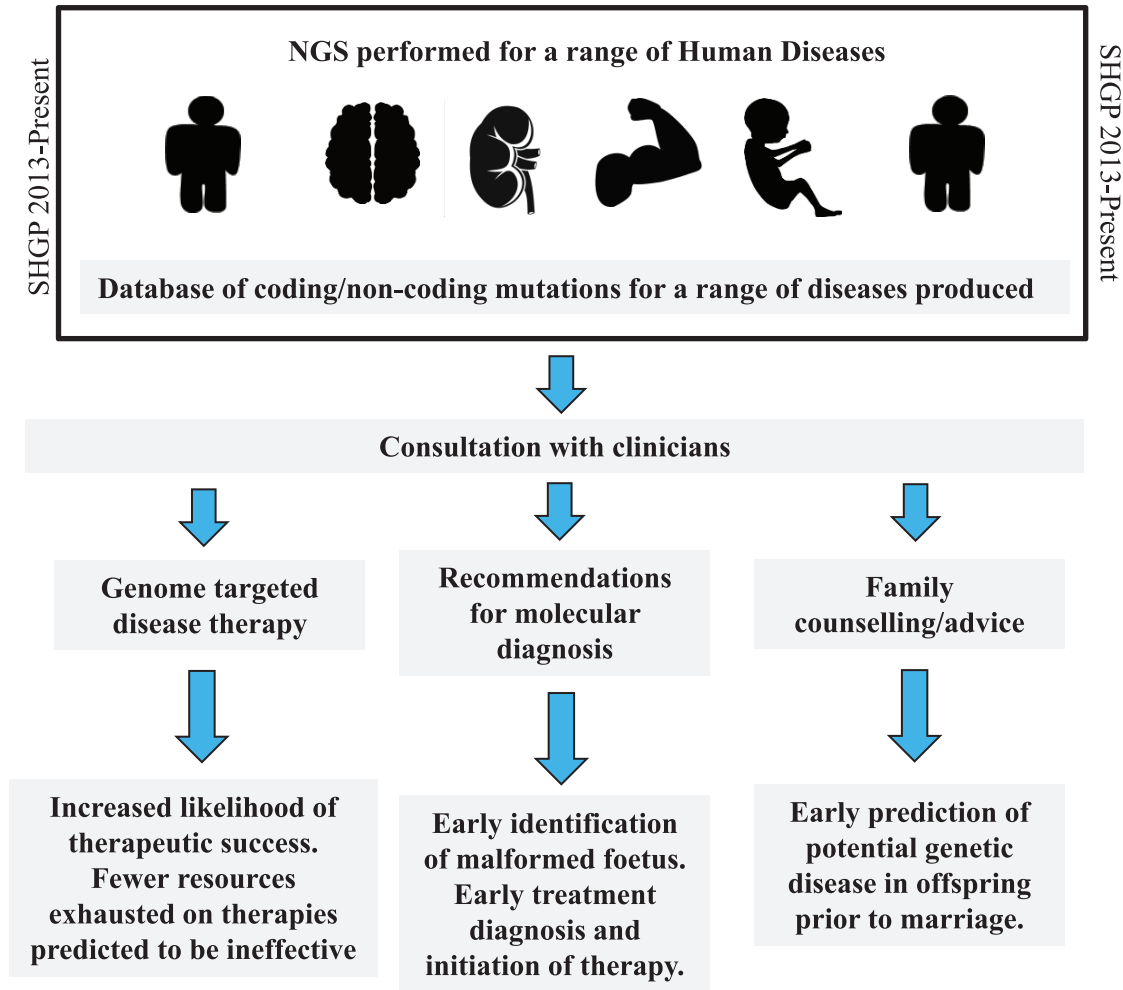


**Figure 2.** Human genomics in global health. (A) Red dots indicate the Arab countries with the highest rates of inbred marriages and thus genetic diseases. (B) Colored dots highlight several global projects in addition to the Saudi Human Genome Project now underway to sequence inherited disease. These projects aim to share information and to develop innovative approaches in the field of human genetics and genome-based therapy.

were unknown. As an example, the focus on mitochondrial proteins helped identify ACAD9 in cases of complex I deficiency (10).

The datasets produced from NGS sequence reads are extensive; as such, advanced computing infrastructure to process

### SHGP-Targeted Therapy Workflow



### Output: Reduced burden of genetic disease in the Arab world

**Figure 3.** Workflow highlighting how data from the SHGP can benefit disease diagnosis and revolutionize gene- and pathway-targeted therapeutics. The SHGP will allow more rational, efficient, and effective drug development pipelines centered on disease-specific genetic abnormalities. This progress provides opportunities for identifying new and effective treatment approaches and rapid disease diagnoses in SA and other Arab countries.

the genomic data has been established. A centralized knowledge base at the KACST was also developed to store information on population variations, including those causing disease, and to make this information available to clinicians in SA to enable future diagnostic and screening efforts (Figure 3) (7). This success

of the SHGP requires the ability to collect an array of subjects that fully represent normal and disease states in the Saudi population. The genomes must be sequenced, the data correctly interpreted and validated, and the results stored in an accessible knowledge-base for future use.

Over the last five years, the project has helped identify the genetic basis of different inherited diseases. To date, over 99 publications charter its success. The project has advanced treatment plans using stem cell and gene editing, in which a defective gene is manipulated by silencing or reintroduced using gene therapy. The KACST plans to support initiatives from genetic data to develop further technologies and treatment regimens for afflicted SA patients. To date, the SHGP is the largest study on the mutational spectrum of genetic diseases in the SA and Arab populations. The non-selective nature of tested families and its representation of all regions of SA allows inference of important patterns of national genetic diseases (11). The highly consanguineous population of SA differs from outbred populations (such as those in Western countries) and thus impacts the landscape of disease-causing mutations (12). The consanguineous nature biases the occurrence of homozygous recessive mutations in diseases that were characterized as dominant, particularly for mutations leading to intellectual disabilities (13). These characteristics increase the sensitivity of NGS assays, allowing the genetic basis of many diseases to be defined.

Using new genomic data, the SHGP has identified a number of diseases that are common or endemic in SA and have a strong genetic component, including intellectual disorders, CNS disorders, metabolic diseases, developmental diseases, hereditary cancers, autoimmune diseases and cardiovascular diseases (11). Some common limitations are shared by the studies published from the SHGP, including a lack of functional follow-up studies, the absence of a replication cohort for gene discovery studies and the differing analysis pipelines. Herein, we will discuss the new genetic knowledge of these diseases and the development of personalized medicine that may arise from this new knowledge.

### **2.1. National human genome projects**

The SHGP is one of an array of national genome projects that have aimed to describe

the genetic background of specific populations (Figure 2). The projects share common aims, namely, (1) the development of infrastructure to create a network of state-of-the-art human genome sequencing centers; (2) training to create the skills, capabilities and capacity for human genome sequencing; and (3) the ability to solve disease genetics and identify traits or disorders within a specific population that can facilitate the genetic distribution of disease within a country. The projects also share common drawbacks, including the incorrect interpretation of a pathogenic variant, resulting from an incomplete human reference genome; an incomplete exon being represented in public databases; and an incorrect annotation of transcripts/exons owing to differing tissue-specific expression patterns (14). The use of genetics to guide therapy is also not without critique. Genomic medicine, through the ability to control the basic genes of a human, may lead to a loss of human diversity, the development of “designer” humans, and a genetic racism for inherited diseases.

An array of national genome projects based on NGS approaches is well established, including projects in Iceland (deCODE), the United Kingdom, the Netherlands, Finland and Sweden (reviewed in (15)). The scale of interest in national genomic screening is highlighted by the emergence of several new sequencing programs, including the initiative on rare and undiagnosed diseases in Japan that was launched in 2015 through collaborations with more than 400 hospitals, including 34 clinical centers; the 100,000 genomes project in China that launched a five-year project in December 2017 and will sequence the genomes of 100,000 individuals from different Chinese ethnic backgrounds; the genome program in Qatar that aims to establish a Qatari reference genome map through the sequencing of 3,000 genomes (~1% of the Qatari population); the genomic health futures mission in Australia that launched in May 2018; the “all of us” research program in the United States that aims to compile data from  $\geq 1$  million individuals beginning with its inception in May 2018; the

personalized medicine program in Estonia that launched in 2016 from an initial cohort of 51,535 gene donors; the genomic program (génomique) in France that was launched in 2016 that will encompass a network of 12 genomic sequencing centers; the Dubai genomics program in the UAE that is currently focusing on building genomic medicine infrastructure and initiating large-scale whole-genome sequencing; and the Turkish genome project in Turkey that plans to sequence 1 million genomes by 2023. When compared to the scale of these studies, the SA population of ~32 million may appear low, but the SHGP has developed  $\geq 13$  gene panels covering  $\geq 5,000$  inherited diseases, identifying  $\geq 2,000$  variants underlying disease (1-3). Over 500 of these mutations are common and have been represented in multiple patients. An underrepresentation of specific ethnic groups in open access international DNA databases occurs, which highlights the importance of having DNA databases from ethnically matching healthy individuals. This is a feature that nation-specific human genome sequencing efforts such as the SHGP have achieved. However, nation-specific studies such as the SHGP are independent and do not form part of a global consortium or open access multipopulation DNA database, limiting the availability and global applications of the study findings. This differs from projects such as the 1000 Genomes Project that allows access to sequencing data from  $\geq 26$  countries to be globally available.

## 2.2. SHGP sequencing approaches

The initial aim of the SHGP was to read the genomes of affected individuals to identify the casual disease genes and their mutations. If tens of thousands of datasets are correctly analyzed, the major disease risk factors in the Saudi population can be identified (1-5). Entire genomes are sequenced when genes that cause human disease are to be identified in a Mendelian fashion (9). Sequencing of the coding regions using whole-exome sequencing (WES) has been successful and has clinical utility, particularly in the case of neurological disorders. However, WES typically uncovers

a number of variants from which identifying causal variants can be challenging. Several tools have been developed to prioritize WES variants, but these require expert analysis. A leading concern of WES in clinical genomics is secondary findings, defined as variants that are potentially medical relevant but are missed as they are unrelated to the medical rationale for the test. This inability of WES to identify incidental/secondary findings is a commonly encountered problem (9, 16).

An alternative to WES is the use of gene panels, which are when a range of genes relevant to a particular phenotype are sequenced in patients presenting that phenotype. This approach provides a focused analysis of clinically relevant genes minimizing many WES challenges. Difficulties in this process arise from the design of an appropriate gene list for a given phenotype, as many diseases have variable clinical presentation. Indeed, gene panels developed across laboratories often differ for the same disease. This can be overcome by relaxing the clinical indications and the integration of supporting evidence to classify diseases into five categories as follows: (1) pathogenic, (2) likely to be pathogenic, (3) uncertain significance of variant, (4) likely benign, or (5) benign. These guidelines have improved disease classification; however, in practice, they remain challenging to standardize. In addition, a shortcoming of gene panels is their need to be updated with the discovery of new gene-disease associations. This is challenging as further disease mutants accumulate from large international datasets.

The Saudi Mendeliome Group defined 13 broad clinical themes in which ~3000 Mendelian genes were distributed (17). The power of this approach was highlighted by the success of the gene panel on the diagnosis of 2300 patients presenting with a range of medical and surgical diseases. Al-Mousa and coworkers also displayed the benefits of targeted NGS panels for primary immunodeficiency diseases, which were more sensitive and cost-effective



than WES approaches, particularly in patients with an atypical presentation of known primary immunodeficiency genes (18). Mustafa and coworkers advanced these studies, demonstrating the suitability of a targeted AmpliSeq Inherited Disease Panel (IDP) (consisting of 328 genes underlying more than 700 inherited diseases) for the first-line screening of genetic diseases following clinical validation (19). It is therefore clear that the sequencing approach is dependent on the type of disorder and degree of genetic or phenotypic heterogeneity. WES has utility for gene discovery studies when exploring the genetic landscape of the population and in diseases with genetic or phenotypic heterogeneity. For those diseases in which little genetic information is available, gene panels have utility, particularly for clinical conditions with distinct and homogenous phenotypes, for which there is sufficient knowledge of their genetic basis.

Despite the obvious progress, knowledge gaps and areas of concern remain. A leading challenge remains that NGS technology is underutilized. Clinical genomics are primarily employed to detect SNVs and indels, but its application in the screening of CNVs, RNA or methylation status remains limited (16). NGS also fails to detect repeat expansion mutations or to quantify alleles. Experts have also expressed concerns over the small numbers of clinical geneticists and/or genetic counselors, as well as the lack of genomic knowledge among active clinicians. This results in a loss of confidence in the application of NGS data to clinical decision making.

### 2.3. Intellectual disability

Intellectual disability (ID) is a common developmental disorder characterized by congenital limitations in intellectual function. Advances in high-throughput whole-genome sequencing (WGS) and single-cell sequencing have increased the number of causative genes identified for this human disease. As the clinical features of ID display heterogeneity regarding their genetic causative factors,

the characterization of ID has benefited from these advances. In SA, the most common cause of ID is genetic and due to the marriage of relatives (1). The SHGP has expanded the locus and allelic heterogeneity of ID in recent years, and the culmination of these studies demonstrates the power of positional mapping to reveal new and unusual mutational mechanisms (20-25). Anazi *et al.* (21) described the phenotypic and genetic findings of 105 patients originating from 68 families with ID, revealing several new disease variants. These genes included *TRAK1*, *GTF3C3*, *SPTBN4* and *NKX6-2* as well as novel variants in 14 other genes, including *ANKHD1*, *ASTN2*, *ATP13A1*, *FMO4*, *MADD*, *MFSD11*, *NCKAP1*, *NFASC*, *PCDHGA10*, *PPP1R21*, *SLC12A2*, *SLK*, *STK32C* and *ZFAT* (21). In particular, *MADD* and *PCDHGA10* were compelling candidates as biallelic deleterious variants in two independent ID families were discovered (21). In other studies, Anazi *et al.* (20) prospectively used molecular karyotyping, multigene panels and whole-exome sequencing (WES) in a cohort of over 300 ID subjects and compared the data with the clinical evaluations of the patients. WES identified independent mutations in three new candidate ID genes, *DENND5A*, *NEMF* and *DNHD1*, and all patients harboring these mutations displayed comparable phenotypes. In addition, *de novo* and recessive variants in 32 other genes were identified, namely, *MAMDC2*, *TUBAL3*, *CPNE6*, *KLHL24*, *USP2*, *PIP5K1A*, *UBE4A*, *TP53TG5*, *ATOH1*, *C16ORF90*, *SLC39A14*, *TRERF1*, *RGL1*, *CDH11*, *SYDE2*, *HIRA*, *FEZF2*, *PROCA1*, *PIANP*, *PLK2*, *QRFPR*, *AP3B2*, *NUDT2*, *UFC1*, *BTN3A2*, *TADA1*, *ARFGEF3*, *FAM160B1*, *ZMYM5*, *SLC45A1*, *ARHGAP33* and *CAPS2* (20). Causal variants from previously published ID genes, including *ASTN1*, *HELZ*, *THOC6*, *WDR45B*, *ADRA2B* and *CLIP1*, were also confirmed, strengthening the accuracy of the study. The very recent identification of *WDR45B* as an ID gene in two large cohorts of affected individuals was revealed in independent studies (25). Additionally, *KIF14* dysfunction was revealed as a cause of autosomal recessive primary microcephaly

(an extremely rare condition characterized by a reduced cerebral cortex accompanied by ID) (26). KIF14 is a mitotic motor protein that is required for spindle localization of the mitotic citron rho-interacting kinase, CIT, highlighting a cellular mechanism of disease progression.

Genetic sequencing has also expanded our knowledge of ID formation. Patel and colleagues demonstrated variable expressivity in three consanguineous families linked to novel 3' UTR mutations in SLC4A4, a gene known to be mutated in a syndromic form of ID (27). The 3' UTR motif mediates posttranscriptional control of many genes, and a marked reduction in the transcript level of SLC4A4 was observed in cells from ID patients. This novel mutational mechanism expanded our knowledge of the variables that underlie phenotypic expressivity in human disease (27).

#### 2.4. Other CNS diseases

For many years, knowledge of the mutations in genes that lead to neurological disorders in SA has been lacking. The gaps in this knowledge have been filled using WES in multiple consanguineous families presenting diseases associated with impaired brain function. Alazami *et al.* (28) identified a total of 33 genes (*SPDL1*, *TUBA3E*, *INO80*, *NID1*, *TSEN15*, *DMBX1*, *CLHC1*, *C12orf4*, *WDR93*, *ST7*, *MATN4*, *SEC24D*, *PCDHB4*, *PTPN23*, *TAF6*, *TBCK*, *FAM177A1*, *KIAA1109*, *MTSS1L*, *XIRP1*, *KCTD3*, *CHAF1B*, *ARV1*, *ISCA2*, *PTRH2*, *GEMIN4*, *MYOCD*, *PDPR*, *DPH1*, *NUP107*, *TMEM92*, *EPB41L4A*, and *FAM120AOS*) as mutations that may impair brain activity, but *in vitro* and *in vivo* data that support a direct link between the identified mutations and impaired brain activity are required. Sowada *et al.* (29) described new forms of developmental and epileptic encephalopathies (DEE) with deleterious biallelic variants in *PTPN23*, a tyrosine kinase strongly expressed in neuronal tissue. The phenotype was characterized by early onset drug-resistant epilepsy, developmental delay, microcephaly, and on occasion premature death. Ramadan *et al.* (30) identified that

recessive mutations in *SCN1B* can cause severe epilepsy in humans in addition to the known dominant *SCN1B* mutations, which were previously reported. This highlighted the need to consider recessive mutations in the interpretation of variants in typically dominant genes.

Shamseldin *et al.* (31) identified other candidate genes involved in a distinct neurodevelopmental phenotype in a multiplex consanguineous family. Mutations in *MICU2*, a major component of the mitochondrial calcium uniporter complex, lead to impaired mitochondrial  $Ca^{2+}$  homeostasis and severe cognitive impairment, spasticity, and white matter involvement (31). Al Mutairi and colleagues also identified mutations in the short-chain enoyl-CoA hydratase (SCEH), a mitochondrial enzyme involved in the oxidation of fatty acids and valine catabolism in early childhood Leigh syndrome (32). When these mutations were accompanied by a missense mutation in *ECHS1*, a lethal phenotype was observed.

The phenotypic and molecular spectrum of Aicardi-Goutières Syndrome (a rare genetic neurological disorder) was recently noted by Al-Mutairi and coworkers to be caused by homozygous mutations in *RNASEH2B*, *RNASEH2A*, *RNASEH2C*, *SAMHD1*, and *TREX1* as well as heterozygous mutations in *IFIH1* in Arab pediatric patients (13). Additionally, in cases of spastic ataxia and hypomyelination, homozygosity mapping and WES revealed missense mutations in *NKX6-2*, which are known to encode a transcriptional repressor with high CNS expression (33). Accordingly, when *NKX6-2* was silenced in mice, hypomyelination was observed (15, 33).

#### 2.5. Autoimmune disease

Studies assessing systemic lupus erythematosus (SLE) in Arab countries suggest its high prevalence in the Arab world (34-36). Using NGS, Carbonella *et al.* (37) identified a homozygous 2 bp deletion in *DNASE1L3* that leads to autosomal recessive autoimmune disease (AID), which mimics SLE. The same



mutations were reported in three siblings from consanguineous parents who presented with hypocomplementemic urticarial vasculitis syndrome (HUVS). As approximately 50% of individuals with HUVS develop SLE, whether SLE is a subphenotype or a separate condition was not clear. Studies to identify the cause of an autosomal-recessive form of systemic juvenile idiopathic arthritis (JIA), which displays both autoimmune- and auto-inflammatory-etologies, in SA patients identified a homoallelic missense mutation in *LACC1* (38). This gene encodes the enzyme laccase, which is a multicopper oxidoreductase. Missense mutations were identified using WES and confirmed by Sanger sequencing. The mutations were present in all investigated cases of JIA disease based on an autosomal-recessive pattern of inheritance. Given the known association of *LACC1* with Crohn's disease and leprosy, this further highlighted this enzyme as a drug target in genetic autoinflammatory disorders (38). This originates from data demonstrating that leprosy hijacks the immune system, manipulating defense mechanisms and causing them to attack neurons.

## 2.6. Kidney disease

Inherited cystic kidney disorders commonly cause end-stage renal disease. To define the phenotype and genotype of cystic kidney disease in fetuses and neonates, Al-Hamed *et al.* (39) correlated antenatal- and postnatal-renal-ultrasound examination with targeted exon sequencing using a defined renal gene panel. The disease phenotypes observed were severe, with 36 cases of stillbirth or perinatal death. Renal gene panel testing identified causative mutations in up to 60% of the families, and mutations were found in 12 genes, including an inferred novel variant in *NEK8* (39). Mutations in *CC2D2A* were the most common cause of antenatal cystic kidney disease and suspected ciliopathy. The renal gene panel from this study holds promise for a rapid molecular diagnosis for this disease in the future.

Meckel-Gruber syndrome (MKS) is characterized by occipital encephalocele,

polydactyly and polycystic kidneys and is genetically heterogeneous with mutations in up to twelve genes known to date. Using a gene panel and NGS in 25 MKS families, Shaheen *et al.* (40) identified a homozygous splice variant in *TMEM107*, which encodes for a protein that localizes to the transition zone at the proximal region of the ciliary axoneme. As *TMEM107*-deficient mice are known to display typical ciliopathy phenotypes, *TMEM107* mutations are now accepted as a factor contributing to human MKS and thus represent a novel drug target for this disorder. In a large population of patients with phenotypes that span the entire ciliopathy spectrum, loss of function mutations in *TXNDC15* encoding a thiol isomerase causing MKS were identified (41). Other novel candidate genes for ciliopathies were also discovered in these studies, including *TRAPPC3*, *EXOC3L2*, *FAM98C*, *C17orf61*, *LRRCC1*, *NEK4*, and *CELSR2*.

## 2.7. Hereditary cancer syndromes

Hereditary cancer syndromes account for approximately 5% of all malignancies. Hereditary cancer syndromes occur due to an inherited mutation increasing the risk of tumor development, which typically occurs at an early age (42). In the majority of hereditary malignant cancers, the elevated risk is due to the mutation of a single gene (monogenic hereditary disease). The ability to identify germline variants in hereditary cancer cases has been a challenge. A history of incomplete cataloguing of cancer-relevant genes combined with a lack of agreement on the patients who should be tested has posed significant barriers. In a recent study from SA, Siraj *et al.* (43) designed a hereditary oncogenesis predisposition evaluation (HOPE) that consisted of many of the genes with known associations with cancer and assessed the effectiveness of HOPE on  $\geq$  1300 tumor/blood samples from patients with ovarian, breast, colorectal and thyroid cancer. Most notably, pathogenic alleles in DNA repair/genomic instability genes other than *BRCA2*, *ATM* and *PALB2* accounted for at least 11.1, 16.8, 50 and 45.5% of mutation-positive ovarian, breast, thyroid and colorectal cancer

patients (CRC), respectively (43). A family history was absent in many of these mutation-positive cases, suggesting a high contribution of germline mutations to cancer predisposition and extending our knowledge beyond “classic” hereditary cancer genes.

## 2.8. Congenital anomalies

Embryonic lethality is recognized as a phenotypic result of individual gene mutations. However, identifying embryonically lethal genes in humans is challenging, particularly when the phenotype manifests at the preimplantation stage. To catalogue recessively acting embryonic lethal genes in SA, Alazami *et al.* (44) identified two families with a female-limited infertility phenotype. Using autozygosity mapping and WES, the phenotype was matched to a single mutation in TLE6, a maternal-effect gene that encodes a member of the subcortical maternal complex (44). Female patients homozygous for TLE6 mutations were found to not undergo early cleavage, leading to sterility. The human mutation abrogates TLE6 phosphorylation, which is reported to be critical for oocyte meiosis (44). This was the first report to identify a human defect in the subcortical maternal complex. Maddirevula *et al.* (45) performed positional mapping and WES and revealed two homozygous deleterious variants in PATL2, which is also necessary for female meiosis and fertility. PATL2 encodes a highly conserved oocyte-specific mRNP repressor of translation, again identifying a novel disease mechanism in humans. The diagnosis of these mutations in IVF clinics may thus hold great promise for the identification of fetal abnormalities.

Orofacial clefting is among the most prevalent type of birth defect. Mutations in the *HYAL2* gene encoding hyaluronidase 2 (which degrades extracellular hyaluronan) were identified as a critical component of the developing heart and palatal shelf matrix (46). In SA studies, Harms *et al.* (47) identified a homozygous truncating variant in *CDH11* associated with Elsahy–Waters syndrome (EWS), also known as branchial–skeletal–genital syndrome. Included in the

clinical features of *CDH11*, mutation-positive individuals had upper eyelid coloboma, which is a new phenotype of this disease that may aid in its diagnosis.

In other studies, Shaheen *et al.* (48) identified loss-of-function mutations in *SMG9* that lead to a loss of nonsense-mediated decay (NMD), which is an important process to degrade transcripts containing premature stop codons that removes their potentially harmful consequences. *SMG9* was shown to be required for normal human development, most likely through this transcriptional regulatory role.

Congenital hydrocephalus is an important birth defect that results in children born with an excessive accumulation of cerebrospinal fluid in the brain. WES combined with positional mapping identified causal mutations in 16 genes; interestingly, none of these mutations were X-linked. Ciliopathies and dystroglycanopathies were the most common etiologies of this congenital hydrocephalus cohort. In a single family with four affected members, a homozygous truncating variant in *EML1* was identified (49). Recessive mutations in *WDR81*, previously linked to cerebellar ataxia, mental retardation, and disequilibrium syndrome, were also found. Other previously identified candidates, including *MPDZ*, were identified, thus highlighting the importance of recessive mutations in this disease condition (49).

## 2.9. Muscle, joint and developmental limb disorders

Congenital disorders of glycosylation are frequently associated with muscle weakness in the Arab population. Using WES, Monies and coworkers reported clinical and pathological features resulting from a homozygous mutation of *ALG2* in an extended family (50). Mutations of *ALG2* manifested as limb and muscle wasting, with defects at both the neuromuscular junction and sarcomere evident. The same group developed a first-line diagnostic assay for limb-girdle muscular dystrophy and myopathies

by evaluating a panel of 759 OMIM genes associated with neurological disorders in Saudi patients presenting with muscle weakness (51).

Larsen syndrome (LS) is characterized by the dislocation of large joints with heterozygous *FLNB* mutations, accounting for the majority of cases. However, biallelic mutations in *CHST3* and *B4GALT7* were recently identified by Patel *et al.* (52), confirming recessive forms of the disease. In a multiplex consanguineous Saudi family affected by severe and recurrent large joint dislocation and severe myopia, a homozygous truncating variant was identified in *GZF1* through a combined autozygome and exome approach (52). The same approach identified a second homozygous truncating *GZF1* variant in a second multiplex consanguineous family affected by severe myopia and milder skeletal involvement (53). *GZF1* encodes the GDNF-inducible zinc finger protein 1, which is a transcription factor of unknown developmental function, and it is expressed in the eyes and limbs of developing mice. These studies reveal new functionality of this protein and its contribution to human disease when dysfunctional.

The spectrum of genetic muscle diseases also extends to diseases of the heart muscle. Dilated cardiomyopathy (DCM) is a common cardiomyopathy that leads to systolic dysfunction and heart failure due to the heart becoming enlarged, thick and rigid. Despite knowledge of over 30 known genes for DCM (primarily in sarcomere and cytoskeletal proteins), Al-Yacoub *et al.* (54) identified mutations in *FBXO32* as a novel DCM-causing locus. The protein encoded by *FBXO32* (also known as atrogin-1) is an E3 ligase that is expressed selectively in skeletal muscle and cardiomyocytes.

### 2.10. Other diseases

Molecular genetic studies are of increasing importance in the diagnosis and classification of several other diseases. Causative mutations in *SKIV2L* and *TTC37*

that cause congenital diarrheal disorders in SA patients were identified by Monies and coworkers (55). Congenital cafe-au-lait spots on the pelvis and lower limbs were a unique and consistent clinical feature of these patients and will aid the differential diagnosis of congenital diarrheal disorders in the future (55).

Ectodermal dysplasia is a highly heterogeneous group of disorders that affect skin, hair, nails and teeth. In an autosomal dominant form of ectodermal dysplasia, Shamseldin *et al.* (53) identified a novel variant in *KDF1* in a Saudi family. The recapitulation of the phenotype was observed in *KDF1* knockout mice, suggesting a causal role played by the *KDF1* variant (56).

Diabetic retinopathy (DR) is a common clinical expression of diabetes mellitus-induced vasculopathy and is a major cause of vision loss. In a cohort of SA diabetic individuals who did not develop DR 10 years after diagnosis compared to SA diabetic individuals with DR, WES identified three genes (*NME3*, *LOC728699*, and *FASTK*) whose rare variant burden protects against DR, highlighting them as attractive candidate drug targets (57). In syndromic RD cases, WES revealed *AGBL5* and *CDH16* as likely gene candidates for the disease (58). A homozygous truncating mutation in *DNAJC17* in a family with retinitis pigmentosa and hypogammaglobulinemia was also identified, thus expanding the allelic spectrum of known RD genes. Patel *et al.* (56), when investigating the genetic basis of pediatric cataracts, applied a multigene panel as well as WES and identified 15 novel genes, including *GEMIN4*, which was mutated in families with cataracts and global developmental delay. Mutations in *RIC1* were identified in patients with cataracts, brain atrophy and microcephaly. Among others, two further candidates that were biallelically inactivated in single cataract families were reported, including *TAF1A* (cataract with global developmental delay) and *WDR87* (nonsyndromic cataract) (56). This expanded the allelic and locus heterogeneity of pediatric cataracts for future diagnosis/treatment.

Advanced cholestatic liver disease is a leading referral to pediatric liver transplant centers in Arab countries. Using an NGS panel of pediatric patients with advanced cholestatic liver disease, causal mutations were identified in the majority of cases and included the novel identification of *TJP2* and *VIPAS39* (59). Due to the alarmingly high carrier frequency of founder mutations identified in this cohort, primary prevention through carrier screening was suggested as the most effective method of minimizing this disease incidence.

### 3. SUMMARY

It is clear that the SHGP has made strides to improve our knowledge of genetic diseases. From the findings, the “one-size-fits-all” approach to medicine, which is based on broad population averages, is not the solution for future therapies in SA. Our enhanced knowledge of the genetic basis for disease offers the prospect of precision medicine. Personalized treatments can allow the individualization of therapy tailored to the molecular profile of a patient’s disease. In the case of hereditary cancer, for example, the identification of tumor-specific alterations, such as point mutations and neoantigen production, can determine a patient’s prognosis and the aggressiveness of the tumor and can predict how well the patient will respond to specific chemotherapy treatments or anti-cancer vaccines (60-65). In the diagnosis, management and treatment of genetic errors of metabolism, disorders may be treatable with stem cell transplantation or with enzyme replacement therapy using CRISPR Cas-9, which was designed to compensate for the identified defective gene. Such approaches have proven effective in the successful correction of inherited liver diseases (66).

Armed with disease-related genetic knowledge, it is also important that genetic counseling services and disease prevention programs run in parallel. Genetic counselors can provide advice for a wide range of genetic disorders with a focus on their prevention through intensive family education and

preventive reproductive options. Prenatal diagnosis and preimplantation genetic diagnosis are of further benefit to recognize genetically inflicted siblings at an early stage. From this information, it is recommended that SA and other Arab governments establish programs that demonstrate the benefits of marriage outside of families (thus reducing inbred genetic diseases) and offer embryonic counseling services for those already afflicted. The combination of efficient diagnosis and precision-based therapy should reduce the occurrence of these genetic defects while improving patient care in the future.

From a sequencing perspective, the advent of NGS technologies has increased its availability. This is highlighted by platforms such as Illumina that have led to a surge in screening programs. NGS, however, has limitations. NGS relies on short  $\leq 1$  kb reads; thus, millions of fragments must be scaled in parallel. This is not conducive to *de novo* assembly as the data require reference genomes. Deep sequencing can compensate for this, but repetitive regions remain challenging. Sample preparation is also time-intensive. This area has been addressed with an array of amplification kits that simplify DNA purification, but further advances in this area are required. In this regard, third-generation sequencers can read kbs of sequence without clonal amplification, subverting the preparation steps required for NGS. This simplifies *de novo* assembly and the reading of repetitive regions. The drawback is that single-molecule sequencing can lack accuracy, but technology is improving, and error rates are decreasing. Regarding human genome projects, further harnessing of the power of NGS is required. For example, NGS now allows the quantification of RNA modifications at a genome-wide scale, revolutionizing our understanding of diverse RNA modifications and their association with disease. The advances in this field require simultaneous advances in chemical reagents, specific antibodies against RNA modifications, and the application of single-molecule RNA sequencing technologies. It is therefore clear that as NGS develops, a step change in both



ease-of-use and speed is predicted. Future technologies should simplify the workflow of technical preparation steps, making it more available for national genome sequencing programs that mirror the SHGP in new countries and territories. The SHGP should now look to form collaborative efforts with other national sequencing programs to increase data accessibility and our understanding of genetic diseases on a global level.

#### 4. ACKNOWLEDGEMENTS

The author thanks King Abdulaziz City for Science and Technology and the Saudi Human Genome Project for technical support.

#### 5. REFERENCES

1. S. G. Project Team: The Saudi human genome program: an oasis in the desert of Arab medicine is providing clues to genetic disease. *IEEE Pulse* 6, 22-26 (2015)  
DOI: 10.1109/MPUL.2015.2476541  
PMid:26583887
2. A. P. Shaik, A. S. Shaik, Y. A. Al-Sheikh: Colorectal cancer: a review of the genome-wide association studies in the kingdom of Saudi Arabia. *Saudi J Gastroenterol* 21, 123-128 (2015)  
DOI: 10.4103/1319-3767.157548  
PMid:26021770 PMCid:PMC4455141
3. S. M. Wakil, R. Ram, N. P. Muiya, M. Mehta, E. Andres, N. Mazhar, B. Baz, S. Hagos, M. Alshahid, B. F. Meyer, G. Morahan, N. Dzimiri: A genome-wide association study reveals susceptibility loci for myocardial infarction/coronary artery disease in Saudi Arabs. *Atherosclerosis* 245, 62-70 (2016)  
DOI: 10.1016/j.atherosclerosis.2015.11.019  
PMid:26708285
4. M. Abu-Elmagd, M. Assidi, H. J. Schulten, A. Dallol, P. Pushparaj, F. Ahmed, S. W. Scherer, M. Al-Qahtani: Individualized medicine enabled by genomics in Saudi Arabia. *BMC Med Genomics* 8, S3 (2015)
5. E. A. Ashley: Towards precision medicine. *Nat Rev Genet* 17, 507-522 (2016)  
DOI: 10.1038/nrg.2016.86  
PMid:27528417
6. M. W. Vegter: Towards precision medicine; a new biomedical cosmology. *Med Health Care Philos* (2018). DOI: 10.1007/s11019-018-9828-z  
DOI: 10.1007/s11019-018-9828-z
7. M. Abouelhoda, T. Faquih, M. El-Kalioby, F. S. Alkuraya: Revisiting the morbid genome of Mendelian disorders. *Genome Biol* 17, 235 (2016)
8. D. C. Koboldt, K. M. Steinberg, D. E. Larson, R. K. Wilson, E. R. Mardis: The next-generation sequencing revolution and its impact on genomics. *Cell* 155, 27-38 (2013)  
DOI: 10.1016/j.cell.2013.09.006  
PMid:24074859 PMCid:PMC3969849
9. S. S. Jamuar, E. C. Tan: Clinical application of next-generation sequencing for Mendelian diseases. *Hum Genomics* 9, 10 (2015)
10. T. B. Haack, K. Danhauser, B. Haberberger, J. Hoser, V. Strecker, D. Boehm, G. Uziel, E. Lamantea, F. Invernizzi, J. Poulton, B. Rolinski, A. Iuso, S. Biskup, T. Schmidt, H. W. Mewes, I. Wittig, T. Meitinger, M. Zeviani, H. Prokisch: Exome sequencing identifies ACAD9 mutations as a cause of complex I deficiency. *Nat Genet* 42, 1131-1134 (2010)  
DOI: 10.1038/ng.706  
PMid:21057504
11. D. Monies, M. Abouelhoda, M. AlSayed, Z. Alhassnan, M. Alotaibi, H. Kayyali, M. Al-Owain, A. Shah, Z. Rahbeeni, M. A. Al-Muhaizea, H. I. Alzaidan, E. Cupler, S. Bohlega, E. Faqeih, M. Faden, B. Alyounes, D. Jaroudi, E. Goljan, H. Elbardisy, A. Akilan, R. Albar, H. Aldhalaan, S. Gulab, A. Chedrawi, B. K. Al Saud, W. Kurdi, N. Makhseed, T. Alqasim, H. Y. El Khashab, H. Al-Mousa, A. Alhashem, I. Kanaan, T. Algoufi, K. Alsaleem, T. A. Basha, F. Al-Murshedi, S. Khan, A. Al-Kindy, M.



- Alnemer, S. Al-Hajjar, S. Alyamani, H. Aldhekri, A. Al-Mehaidib, R. Arnaut, O. Dabbagh, M. Shagrani, D. Broering, M. Tulbah, A. Alqassmi, M. Almugbel, M. AlQuaiz, A. Alsaman, K. Al-Thihli, R. A. Sulaiman, W. Al-Dekhail, A. Alsaegh, F. A. Bashiri, A. Qari, S. Alhomadi, H. Alkuraya, M. Alsebayel, M. H. Hamad, L. Szonyi, F. Abaalkhail, S. M. Al-Mayouf, H. Almojalli, K. S. Alqadi, H. Elsiesy, T. M. Shuaib, M. Z. Seidahmed, I. Abosoudah, H. Akleh, A. AlGhoniaum, T. M. Alkharfy, F. Al Mutairi, W. Eyaid, A. Alshanbary, F. R. Sheikh, F. I. Alsohaibani, A. Alsonbul, S. Al Tala, S. Balkhy, R. Bassiouni, A. S. Alenizi, M. H. Hussein, S. Hassan, M. Khalil, B. Tabarki, S. Alshahwan, A. Oshi, Y. Sabr, S. Alsaadoun, M. A. Salih, S. Mohamed, H. Sultana, A. Tamim, M. El-Haj, S. Alshahrani, D. K. Bubshait, M. Alfadhel, T. Faquih, M. El-Kalioby, S. Subhani, Z. Shah, N. Moghrabi, B. F. Meyer, F. S. Alkuraya: The landscape of genetic diseases in Saudi Arabia based on the first 1000 diagnostic panels and exomes. *Hum Genet* 136, 921-939 (2017)  
DOI: 10.1007/s00439-017-1821-8  
PMid:28600779 PMCid:PMC5502059
12. M. A. El-Hazmi, A. R. Al-Swailem, A. S. Warsy, A. M. Al-Swailem, R. Sulaimani, A. A. Al-Meshari: Consanguinity among the Saudi Arabian population. *J Med Genet* 32, 623-626 (1995)  
DOI: 10.1136/jmg.32.8.623  
PMid:7473654 PMCid:PMC1051637
13. F. Al Mutairi, M. Alfadhel, M. Nashabat, A. W. El-Hattab, T. Ben-Omran, J. Hertecant, W. Eyaid, R. Ali, A. Alasmari, M. Kara, W. Al-Twaijri, R. Filimban, A. Alshenqiti, M. Al-Owain, E. Faqeih, F. S. Alkuraya: Phenotypic and molecular spectrum of Aicardi-Goutieres Syndrome: a study of 24 patients. *Pediatr Neurol* 78, 35-40 (2018)  
DOI: 10.1016/j.pediatrneurol.2017.09.002  
PMid:29239743
14. C. A. Steward, A. P. J. Parker, B. A. Minassian, S. M. Sisodiya, A. Frankish, J. Harrow: Genome annotation for clinical genomic diagnostics: strengths and weaknesses. *Genome Med* 9, 49 (2017)
15. I. Dorboz, C. Aiello, C. Simons, R. T. Stone, M. Niceta, M. Elmaleh, M. Abuawad, D. Doummar, A. Bruselles, N. I. Wolf, L. Travaglini, O. Boespflug-Tanguy, M. Tartaglia, A. Vanderver, D. Rodriguez, E. Bertini: Biallelic mutations in the homeodomain of NKX6-2 underlie a severe hypomyelinating leukodystrophy. *Brain* 140, 2550-2556 (2017)  
DOI: 10.1093/brain/awx207  
PMid:28969374
16. J. B. Krier, S. S. Kalia, R. C. Green: Genomic sequencing in clinical practice: applications, challenges, and opportunities. *Dialogues Clin Neurosci* 18, 299-312 (2016)
17. G. S. Mendeliome: Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. *Genome Biol* 16, 134 (2015)
18. H. Al-Mousa, M. Abouelhoda, D. M. Monies, N. Al-Tassan, A. Al-Ghoniaum, B. Al-Saud, H. Al-Dhekri, R. Arnaut, S. Al-Muhsen, N. Ades, S. Elshorbagi, S. Al Gazlan, F. Sheikh, M. Dasouki, L. El-Baik, T. Elamin, A. Jaber, O. Kheir, M. El-Kalioby, S. Subhani, E. Al Idrissi, M. Al-Zahrani, M. Alhelale, N. Alnader, A. Al-Otaibi, R. Kattan, K. Al Abdelrahman, M. M. Al Breacan, F. S. Bin Humaid, S. M. Wakil, F. Alzayer, H. Al-Dusery, T. Faquih, S. Al-Hissi, B. F. Meyer, A. Hawwari: Unbiased targeted next-generation sequencing molecular approach for primary immunodeficiency diseases. *J Allergy Clin Immunol* 137, 1780-1787 (2016)  
DOI: 10.1016/j.jaci.2015.12.1310  
PMid:26915675
19. A. E. Mustafa, T. Faquih, B. Baz, R. Kattan, A. Al-Issa, A. I. Tahir, F. Imtiaz, K. Ramzan, M. Al-Sayed, M. Alowain, Z. Al-Hassnan, H. Al-Zaidan, M. Abouelhoda, B. R. Al-Mubarak, N. A. Al Tassan: Validation of

- Ion Torrent(TM) inherited disease panel with the PGM(TM) sequencing platform for rapid and comprehensive mutation detection. *Genes* 9 (2018). DOI: 10.3390/genes9050267  
DOI: 10.3390/genes9050267
20. S. Anazi, S. Maddirevula, E. Faqeih, H. Alsedairy, F. Alzahrani, H. E. Shamseldin, N. Patel, M. Hashem, N. Ibrahim, F. Abdulwahab, N. Ewida, H. S. Alsaif, H. Al Sharif, W. Alamoudi, A. Kentab, F. A. Bashiri, M. Alnaser, A. H. AlWadei, M. Alfadhel, W. Eyaid, A. Hashem, A. Al Asmari, M. M. Saleh, A. AlSaman, K. A. Alhasan, M. Alsughayir, M. Al Shammari, A. Mahmoud, Z. N. Al-Hassnan, M. Al-Husain, R. O. Khalil, N. Abd El Meguid, A. Masri, R. Ali, T. Ben-Omran, P. El Fishway, A. Hashish, A. E. Sencicek, M. State, A. M. Alazami, M. A. Salih, N. Altassan, S. T. Arold, M. Abouelhoda, S. M. Wakil, D. Monies, R. Shaheen, F. S. Alkuraya: Clinical genomics expands the morbid genome of intellectual disability and offers a high diagnostic yield. *Mol Psychiatry* 22, 615-624 (2017)  
DOI: 10.1038/mp.2016.113  
PMid:27431290
  21. S. Anazi, S. Maddirevula, V. Salpietro, Y. T. Asi, S. Alsahli, A. Alhashem, H. E. Shamseldin, F. AlZahrani, N. Patel, N. Ibrahim, F. M. Abdulwahab, M. Hashem, N. Alhashmi, F. Al Murshedi, A. Al Kindy, A. Alshaer, A. Rumayyan, S. Al Tala, W. Kurdi, A. Alsaman, A. Alasmari, S. Banu, T. Sultan, M. M. Saleh, H. Alkuraya, M. A. Salih, H. Aldhalaan, T. Ben-Omran, F. Al Musafri, R. Ali, J. Suleiman, B. Tabarki, A. W. El-Hattab, C. Bupp, M. Alfadhel, N. Al Tassan, D. Monies, S. T. Arold, M. Abouelhoda, T. Lashley, H. Houlden, E. Faqeih, F. S. Alkuraya: Expanding the genetic heterogeneity of intellectual disability. *Hum Genet* 136, 1419-1429 (2017)  
DOI: 10.1007/s00439-017-1843-2  
PMid:28940097
  22. S. Anazi, S. Maddirevula, V. Salpietro, Y. T. Asi, S. Alsahli, A. Alhashem, H. E. Shamseldin, F. AlZahrani, N. Patel, N. Ibrahim, F. M. Abdulwahab, M. Hashem, N. Alhashmi, F. Al Murshedi, A. Al Kindy, A. Alshaer, A. Rumayyan, S. Al Tala, W. Kurdi, A. Alsaman, A. Alasmari, S. Banu, T. Sultan, M. M. Saleh, H. Alkuraya, M. A. Salih, H. Aldhalaan, T. Ben-Omran, F. Al Musafri, R. Ali, J. Suleiman, B. Tabarki, A. W. El-Hattab, C. Bupp, M. Alfadhel, N. Al Tassan, D. Monies, S. T. Arold, M. Abouelhoda, T. Lashley, H. Houlden, E. Faqeih, F. S. Alkuraya: Correction to: expanding the genetic heterogeneity of intellectual disability. *Hum Genet* 137, 105-109 (2018)  
DOI: 10.1007/s00439-017-1859-7  
PMid:29288388
  23. S. Anazi, H. E. Shamseldin, D. AlNageb, M. Abouelhoda, D. Monies, M. A. Salih, K. Al-Rubeaan, F. S. Alkuraya: A null mutation in TNIK defines a novel locus for intellectual disability. *Hum Genet* 135, 773-778 (2016)  
DOI: 10.1007/s00439-016-1671-9  
PMid:27106596
  24. C. M. Loucks, J. S. Parboosingh, R. Shaheen, F. P. Bernier, D. R. McLeod, M. Z. Seidahmed, E. G. Puffenberger, C. Ober, R. A. Hegele, K. M. Boycott, F. S. Alkuraya, A. M. Innes: Matching two independent cohorts validates DPH1 as a gene responsible for autosomal recessive intellectual disability with short stature, craniofacial, and ectodermal anomalies. *Hum Mutat* 36, 1015-1019 (2015)  
DOI: 10.1002/humu.22843  
PMid:27106596
  25. J. Suleiman, D. Allingham-Hawkins, M. Hashem, H. E. Shamseldin, F. S. Alkuraya, A. W. El-Hattab: WDR45B-related intellectual disability, spastic quadriplegia, epilepsy, and cerebral hypoplasia: a consistent Neurodevelopmental Syndrome. *Clin Genet* 93, 360-364 (2018)  
DOI: 10.1111/cge.13054  
PMid:28503735
  26. A. Moawia, R. Shaheen, S. Rasool, S. S. Waseem, N. Ewida, B. Budde, A. Kawalia, S. Motameny, K. Khan, A.

- Fatima, M. Jameel, F. Ullah, T. Akram, Z. Ali, U. Abdullah, S. Irshad, W. Hohne, A. A. Noegel, M. Al-Owain, K. Hortnagel, P. Stobe, S. M. Baig, P. Nurnberg, F. S. Alkuraya, A. Hahn, M. S. Hussain: Mutations of KIF14 cause primary microcephaly by impairing cytokinesis. *Ann Neurol* 82, 562-577 (2017)  
DOI: 10.1002/ana.25044  
PMid:28892560
27. N. Patel, A. O. Khan, M. Al-Saif, W. N. Moghrabi, B. M. AlMaarik, N. Ibrahim, F. Abdulwahab, M. Hashem, T. Alshidi, E. Alobeid, R. A. Alomar, S. Al-Harbi, M. Abouelhoda, K. S. A. Khabar, F. S. Alkuraya: A novel mechanism for variable phenotypic expressivity in Mendelian diseases uncovered by an AU-rich element (ARE)-creating mutation. *Genome Biol* 18, 144 (2017)
28. A. M. Alazami, N. Patel, H. E. Shamseldin, S. Anazi, M. S. Al-Dosari, F. Alzahrani, H. Hijazi, M. Alshammari, M. A. Aldahmesh, M. A. Salih, E. Fageih, A. Alhashem, F. A. Bashiri, M. Al-Owain, A. Y. Kentab, S. Sogaty, S. Al Tala, M. H. Tamsah, M. Tulbah, R. F. Aljelaify, S. A. Alshahwan, M. Z. Seidahmed, A. A. Alhadid, H. Aldhalaan, F. AlQallaf, W. Kurdi, M. Alfadhel, Z. Babay, M. Alsogheer, N. Kaya, Z. N. Al-Hassnan, G. M. Abdel-Salam, N. Al-Sannaa, F. Al Mutairi, H. Y. El Khashab, S. Bohlega, X. Jia, H. C. Nguyen, R. Hammami, N. Adly, J. Y. Mohamed, F. Abdulwahab, N. Ibrahim, E. A. Naim, B. Al-Younes, B. F. Meyer, M. Hashem, R. Shaheen, Y. Xiong, M. Abouelhoda, A. A. Aldeeri, D. M. Monies, F. S. Alkuraya: Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. *Cell Rep* 10, 148-161 (2015)  
DOI: 10.1016/j.celrep.2014.12.015  
PMid:25558065
29. N. Sowada, M. O. Hashem, R. Yilmaz, M. Hamad, N. Kakar, H. Thiele, S. T. Arold, H. Bode, F. S. Alkuraya, G. Borck: Mutations of PTPN23 in developmental and epileptic encephalopathy. *Hum Genet* 136, 1455-1461 (2017)  
DOI: 10.1007/s00439-017-1850-3  
PMid:29090338
30. W. Ramadan, N. Patel, S. Anazi, A. Y. Kentab, F. A. Bashiri, M. H. Hamad, L. Jad, M. A. Salih, H. Alsaif, M. Hashem, E. Fageih, H. E. Shamseddin, F. S. Alkuraya: Confirming the recessive inheritance of SCN1B mutations in developmental epileptic encephalopathy. *Clin Genet* 92, 327-331 (2017)  
DOI: 10.1111/cge.12999  
PMid:28218389
31. H. E. Shamseldin, A. Alasmari, M. A. Salih, M. M. Samman, S. A. Mian, T. Alshidi, N. Ibrahim, M. Hashem, E. Fageih, F. Al-Mohanna, F. S. Alkuraya: A null mutation in MICU2 causes abnormal mitochondrial calcium homeostasis and a severe neurodevelopmental disorder. *Brain* 140, 2806-2813 (2017)  
DOI: 10.1093/brain/awx237  
PMid:29053821
32. F. Al Mutairi, H. E. Shamseldin, M. Alfadhel, R. J. Rodenburg, F. S. Alkuraya: A lethal neonatal phenotype of mitochondrial short-chain enoyl-CoA hydratase-1 deficiency. *Clin Genet* 91, 629-633 (2017)  
DOI: 10.1111/cge.12891  
PMid:27905109
33. V. Chelban, N. Patel, J. Vandrovцова, M. N. Zanetti, D. S. Lynch, M. Ryten, J. A. Botia, O. Bello, E. Tribollet, S. Efthymiou, I. Davagnanam, F. A. Bashiri, N. W. Wood, J. E. Rothman, F. S. Alkuraya, H. Houlden: Mutations in NKX6-2 cause progressive spastic ataxia and hypomyelination. *Am J Hum Genet* 100, 969-977 (2017)  
DOI: 10.1016/j.ajhg.2017.05.009  
PMid:28575651 PMCid:PMC5473715
34. A. M. AlDhanhani, M. Agarwal, Y. S. Othman, O. Bakoush: Incidence and prevalence of systemic lupus erythematosus among the native Arab population in UAE. *Lupus* 26, 664-669 (2017)  
DOI: 10.1177/0961203316678677  
PMid:27831539

35. S. AlFadhli, A. AlFailakawi, A. A. Ghanem: Th-17 related regulatory network in the pathogenesis of Arab patients with systemic lupus erythematosus and lupus nephritis. *Int J Rheum Dis* 19, 512-520 (2016)  
DOI: 10.1111/1756-185X.12393  
PMid:25496163
36. M. Housey, P. DeGuire, S. Lyon-Callo, L. Wang, W. Marder, W. J. McCune, C. G. Helmick, C. Gordon, J. P. Dhar, J. Leisen, E. C. Somers: Incidence and prevalence of systemic lupus erythematosus among Arab and Chaldean Americans in Southeastern Michigan: the Michigan lupus epidemiology and surveillance program. *Am J Public Health* 105, e74-e79 (2015)  
DOI: 10.2105/AJPH.2014.302423  
PMid:25790387 PMCid:PMC4386493
37. A. Carbonella, G. Mancano, E. Gremese, F. S. Alkuraya, N. Patel, F. Gurrieri, G. Ferraccioli: An autosomal recessive DNASE1L3-related autoimmune disease with unusual clinical presentation mimicking systemic lupus erythematosus. *Lupus* 26, 768-772 (2017)  
DOI: 10.1177/0961203316676382  
PMid:27821515
38. S. M. Wakil, D. M. Monies, M. Abouelhoda, N. Al-Tassan, H. Al-Dusery, E. A. Naim, B. Al-Younes, J. Shinwari, F. A. Al-Mohanna, B. F. Meyer, S. Al-Mayouf: Association of a mutation in LACC1 with a monogenic form of systemic juvenile idiopathic arthritis. *Arthritis Rheumatol* 67, 288-295 (2015)  
DOI: 10.1002/art.38877  
PMid:25220867
39. M. H. Al-Hamed, W. Kurdi, N. Alsahan, Z. Alabdullah, R. Abudraz, M. Tulbah, M. Alnemer, R. Khan, H. Al-Jurayb, A. Alahmed, A. I. Tahir, D. Khalil, N. Edwards, B. AlAbdulaziz, F. S. Binhumaid, S. Majid, T. Faquih, M. El-Kalioby, M. Abouelhoda, N. Altassan, D. Monies, B. Meyer, J. A. Sayer, M. Albaqumi: Genetic spectrum of Saudi Arabian patients with antenatal cystic kidney disease and ciliopathy phenotypes using a targeted renal gene panel. *J Med Genet* 53, 338-347 (2016)  
DOI: 10.1136/jmedgenet-2015-103469  
PMid:26862157 PMCid:PMC4853542
40. R. Shaheen, A. Almoisheer, E. Faqeih, Z. Babay, D. Monies, N. Tassan, M. Abouelhoda, W. Kurdi, E. Al Mardawi, M. M. Khalil, M. Z. Seidahmed, M. Alnemer, N. Alsahan, S. Sogaty, A. Alhashem, A. Singh, M. Goyal, S. Kapoor, R. Alomar, N. Ibrahim, F. S. Alkuraya: Identification of a novel MKS locus defined by TMEM107 mutation. *Hum Mol Genet* 24, 5211-5218 (2015)  
DOI: 10.1093/hmg/ddv242  
PMid:26123494
41. R. Shaheen, K. Szymanska, B. Basu, N. Patel, N. Ewida, E. Faqeih, A. Al Hashem, N. Derar, H. Alsharif, M. A. Aldahmesh, A. M. Alazami, M. Hashem, N. Ibrahim, F. M. Abdulwahab, R. Sonbul, H. Alkuraya, M. Alnemer, S. Al Tala, M. Al-Husain, H. Morsy, M. Z. Seidahmed, N. Meriki, M. Al-Owain, S. AlShahwan, B. Tabarki, M. A. Salih, T. Faquih, M. El-Kalioby, M. Ueffing, K. Boldt, C. V. Logan, D. A. Parry, N. Al Tassan, D. Monies, A. Megarbane, M. Abouelhoda, A. Halees, C. A. Johnson, F. S. Alkuraya: Characterizing the morbid genome of ciliopathies. *Genome Biol* 17, 242 (2016)
42. N. Rahner, V. Steinke: Hereditary cancer syndromes. *Dtsch Arztebl Int* 105, 706-714 (2008)  
DOI: 10.3238/arztebl.2008.0706
43. A. K. Siraj, T. Masoodi, R. Bu, S. K. Parvathareddy, I. A. Al-Badawi, N. Al-Sanea, L. H. Ashari, A. Abduljabbar, S. Alhomoud, S. S. Al-Sobhi, A. Tulbah, D. Ajarim, K. Alzoman, M. Aljuboury, H. B. Yousef, M. Al-Dawish, F. Al-Dayel, F. S. Alkuraya, K. S. Al-Kuraya: Expanding the spectrum of germline variants in cancer. *Hum Genet* 136, 1431-1444 (2017)  
DOI: 10.1007/s00439-017-1845-0  
PMid:28975465



44. A. M. Alazami, S. M. Awad, S. Coskun, S. Al-Hassan, H. Hijazi, F. M. Abdulwahab, C. Poizat, F. S. Alkuraya: TLE6 mutation causes the earliest known human embryonic lethality. *Genome Biol* 16, 240 (2015)
45. S. Maddirevula, S. Coskun, S. Alhassan, A. Elnour, H. S. Alsaif, N. Ibrahim, F. Abdulwahab, S. T. Arold, F. S. Alkuraya: Female infertility caused by mutations in the oocyte-specific translational repressor PATL2. *Am J Hum Genet* 101, 603-608 (2017)  
DOI: 10.1016/j.ajhg.2017.08.009  
PMid:28965844 PMCID:PMC5630161
46. M. M. Muggenthaler, B. Chowdhury, S. N. Hasan, H. E. Cross, B. Mark, G. V. Harlalka, M. A. Patton, M. Ishida, E. R. Behr, S. Sharma, K. Zahka, E. Faqeih, B. Blakley, M. Jackson, M. Lees, V. Dolinsky, L. Cross, P. Stanier, C. Salter, E. L. Baple, F. S. Alkuraya, A. H. Crosby, B. Triggs-Raine, B. A. Chioza: Mutations in HYAL2, encoding hyaluronidase 2, cause a syndrome of orofacial clefting and cor triatriatum sinister in humans and mice. *PLoS Genet* 13, e1006470 (2017)
47. F. L. Harms, S. Nampoothiri, S. Anazi, D. Yesodharan, M. Alawi, K. Kutsche, F. S. Alkuraya: Elsayh-Waters Syndrome is caused by biallelic mutations in CDH11. *Am J Med Genet A* 176, 477-482 (2018)  
DOI: 10.1002/ajmg.a.38568  
PMid:29271567
48. R. Shaheen, S. Anazi, T. Ben-Omran, M. Z. Seidahmed, L. B. Caddle, K. Palmer, R. Ali, T. Alshidi, S. Hagos, L. Goodwin, M. Hashem, S. M. Wakil, M. Abouelhoda, D. Colak, S. A. Murray, F. S. Alkuraya: Mutations in SMG9, encoding an essential component of nonsense-mediated decay machinery, cause a Multiple Congenital Anomaly Syndrome in humans and mice. *Am J Hum Genet* 98, 643-652 (2016)  
DOI: 10.1016/j.ajhg.2016.02.010  
PMid:27018474 PMCID:PMC4833216
49. R. Shaheen, M. A. Sebai, N. Patel, N. Ewida, W. Kurdi, I. Altweijri, S. Sogaty, E. Almardawi, M. Z. Seidahmed, A. Alnemri, S. Maddirevula, N. Ibrahim, F. Abdulwahab, M. Hashem, T. Al-Sheddi, R. Alomar, E. Alobeid, B. Sallout, B. AlBaqawi, W. AlAali, N. Ajaji, H. Lesmana, R. J. Hopkin, L. Dupuis, R. Mendoza-Londono, H. Al Rukban, G. Yoon, E. Faqeih, F. S. Alkuraya: The genetic landscape of familial congenital hydrocephalus. *Ann Neurol* 81, 890-897 (2017)  
DOI: 10.1002/ana.24964  
PMid:28556411
50. D. M. Monies, H. N. Al-Hindi, M. A. Al-Muhaizea, D. J. Jaroudi, B. Al-Younes, E. A. Naim, S. M. Wakil, B. F. Meyer, S. Bohlega: Clinical and pathological heterogeneity of a congenital disorder of glycosylation manifesting as a Myasthenic/Myopathic Syndrome. *Neuromuscul Disord* 24, 353-359 (2014)  
DOI: 10.1016/j.nmd.2013.12.010  
PMid:24461433
51. D. Monies, H. N. Alhindi, M. A. Almuhaizea, M. Abouelhoda, A. M. Alazami, E. Goljan, B. Alyounes, D. Jaroudi, A. Allssa, K. Alabdulrahman, S. Subhani, M. El-Kalioby, T. Faquih, S. M. Wakil, N. A. Altassan, B. F. Meyer, S. Bohlega: A first-line diagnostic assay for limb-girdle muscular dystrophy and other myopathies. *Hum Genomics* 10, 32 (2016)
52. N. Patel, H. E. Shamseldin, N. Sakati, A. O. Khan, A. Softa, F. M. Al-Fadhli, M. Hashem, F. M. Abdulwahab, T. Alshidi, R. Alomar, E. Alobeid, S. M. Wakil, D. Colak, F. S. Alkuraya: GZF1 mutations expand the genetic heterogeneity of Larsen Syndrome. *Am J Hum Genet* 100, 831-836 (2017)  
DOI: 10.1016/j.ajhg.2017.04.008  
PMid:28475863 PMCID:PMC5420357
53. H. E. Shamseldin, O. Khalifa, Y. M. Binamer, A. Almutawa, S. T. Arold, H. Zaidan, F. S. Alkuraya: KDF1, encoding keratinocyte differentiation factor 1, is mutated in a multigenerational family with ectodermal dysplasia. *Hum Genet* 136,



- 99-105 (2017)  
DOI: 10.1007/s00439-016-1741-z  
PMid:27838789
54. N. Al-Yacoub, R. Shaheen, S. M. Awad, M. Kunhi, N. Dzimiri, H. C. Nguyen, Y. Xiong, J. Al-Buraiki, W. Al-Habeeb, F. S. Alkuraya, C. Poizat: FBXO32, encoding a member of the SCF complex, is mutated in dilated cardiomyopathy. *Genome Biol* 17, 2 (2016)
  55. D. M. Monies, Z. Rahbeeni, M. Abouelhoda, E. A. Naim, B. Al-Younes, B. F. Meyer, A. Al-Mehaidib: Expanding phenotypic and allelic heterogeneity of Tricho-Hepato-Enteric Syndrome. *J Pediatr Gastroenterol Nutr* 60, 352-356 (2015)  
DOI: 10.1097/MPG.0000000000000627  
PMid:25714577
  56. N. Patel, D. Anand, D. Monies, S. Maddirevula, A. O. Khan, T. Algoufi, M. Alowain, E. Faqeih, M. Alshammari, A. Qudair, H. Alsharif, F. Aljubran, H. S. Alsaif, N. Ibrahim, F. M. Abdulwahab, M. Hashem, H. Alsedairy, M. A. Aldahmesh, S. A. Lachke, F. S. Alkuraya: Novel phenotypes and loci identified through clinical genomics approaches to pediatric cataract. *Hum Genet* 136, 205-225 (2017)  
DOI: 10.1007/s00439-016-1747-6  
PMid:27878435 PMCID:PMC5783298
  57. C. Shtir, M. A. Aldahmesh, S. Al-Dahmash, E. Abboud, H. Alkuraya, M. A. Abouammoh, S. R. Nowailaty, G. Al-Thubaiti, E. A. Naim, B. AlYounes, F. S. Binhumaid, A. B. AlOtaibi, A. S. Altamimi, F. H. Alamer, M. Hashem, M. Abouelhoda, D. Monies, F. S. Alkuraya: Exome-based case-control association study using extreme phenotype design reveals novel candidates with protective effect in diabetic retinopathy. *Hum Genet* 135, 193-200 (2016)  
DOI: 10.1007/s00439-015-1624-8  
PMid:26693933
  58. N. Patel, M. A. Aldahmesh, H. Alkuraya, S. Anazi, H. Alsharif, A. O. Khan, A. Sunker, S. Al-Mohsen, E. B. Abboud, S. R. Nowilaty, M. Alowain, H. Al-Zaidan, B. Al-Saud, A. Alasmari, G. M. Abdel-Salam, M. Abouelhoda, F. M. Abdulwahab, N. Ibrahim, E. Naim, B. Al-Younes, E. A. AlMostafa, A. Allssa, M. Hashem, O. Buzovetsky, Y. Xiong, D. Monies, N. Altassan, R. Shaheen, S. A. Al-Hazaa, F. S. Alkuraya: Expanding the clinical, allelic, and locus heterogeneity of retinal dystrophies. *Genet Med* 18, 554-562 (2016)  
DOI: 10.1038/gim.2015.127  
PMid:26355662
  59. M. Shagrani, J. Burkholder, D. Broering, M. Abouelhoda, T. Faquih, M. El-Kalioby, S. N. Subhani, E. Goljan, R. Albar, D. Monies, N. Mazhar, B. S. AlAbdulaziz, K. A. Abdelrahman, N. Altassan, F. S. Alkuraya: Genetic profiling of children with advanced cholestatic liver disease. *Clin Genet* 92, 52-61 (2017)  
DOI: 10.1111/cge.12959  
PMid:28039895
  60. G. M. Shankar, M. Abedalthagafi, R. A. Vaubel, P. H. Merrill, N. Nayyar, C. M. Gill, R. Brewster, W. L. Bi, P. K. Agarwalla, A. R. Thorner, D. A. Reardon, O. Al-Mefty, P. Y. Wen, B. M. Alexander, P. van Hummelen, T. T. Batchelor, K. L. Ligon, A. H. Ligon, M. Meyerson, I. F. Dunn, R. Beroukhim, D. N. Louis, A. Perry, S. L. Carter, C. Giannini, W. T. Curry, D. P. Cahill, F. G. Barker, P. K. Brastianos, S. Santagata: Germline and somatic BAP1 mutations in high-grade rhabdoid meningiomas. *Neuro Oncol* 19, 535-545 (2017)  
DOI: 10.1093/neuonc/nox094  
PMCID:PMC5737364
  61. T. A. Bale, M. Abedalthagafi, W. L. Bi, Y. J. Kang, P. Merrill, I. F. Dunn, A. Dubuc, S. K. Charbonneau, L. Brown, A. H. Ligon, S. H. Ramkissoon, K. L. Ligon: Genomic characterization of recurrent high-grade astroblastoma. *Cancer Genet* 209, 321-330 (2016)  
DOI: 10.1016/j.cancergen.2016.06.002  
PMid:27425854

62. W. L. Bi, P. Horowitz, N. F. Greenwald, M. Abedalthagafi, P. K. Agarwalla, W. J. Gibson, Y. Mei, S. E. Schumacher, U. Ben-David, A. Chevalier, S. Carter, G. Tiao, P. K. Brastianos, A. H. Ligon, M. Ducar, L. MacConaill, E. R. Laws, S. Santagata, R. Beroukhim, I. F. Dunn: Landscape of genomic alterations in pituitary adenomas. *Clin Cancer Res* 23, 1841-1851 (2017)  
DOI: 10.1158/1078-0432.CCR-16-0790  
PMid:27707790 PMCID:PMC5380512
63. W. L. Bi, N. F. Greenwald, S. H. Ramkissoon, M. Abedalthagafi, S. M. Coy, K. L. Ligon, Y. Mei, L. MacConaill, M. Ducar, L. Min, S. Santagata, U. B. Kaiser, R. Beroukhim, E. R. Laws, I. F. Dunn: Clinical identification of oncogenic drivers and copy-number alterations in pituitary tumors. *Endocrinology* 158, 2284-2291 (2017)  
DOI: 10.1210/en.2016-1967  
PMid:28486603 PMCID:PMC5505210
64. W. L. Bi, N. F. Greenwald, M. Abedalthagafi, J. Wala, W. J. Gibson, P. K. Agarwalla, P. Horowitz, S. E. Schumacher, E. Esaulova, Y. Mei, A. Chevalier, M. Ducar, A. R. Thorner, P. van Hummelen, A. Stemmer-Rachamimov, M. Artyomov, O. Al-Mefty, G. P. Dunn, S. Santagata, I. F. Dunn, R. Beroukhim: Genomic landscape of high-grade meningiomas. *NPJ Genom Med* 2 (2017). DOI: 10.1038/s41525-017-0014-7  
DOI: 10.1038/s41525-017-0014-7
65. M. Abedalthagafi, W. L. Bi, A. A. Aizer, P. H. Merrill, R. Brewster, P. K. Agarwalla, M. L. Listewnik, D. Dias-Santagata, A. R. Thorner, P. Van Hummelen, P. K. Brastianos, D. A. Reardon, P. Y. Wen, O. Al-Mefty, S. H. Ramkissoon, R. D. Folkerth, K. L. Ligon, A. H. Ligon, B. M. Alexander, I. F. Dunn, R. Beroukhim, S. Santagata: Oncogenic PI3K mutations are as common as AKT1 and SMO mutations in meningioma. *Neuro Oncol* 18, 649-655 (2016)  
DOI: 10.1093/neuonc/nov316  
PMid:26826201 PMCID:PMC4827048
66. H. Mollanoori, S. Teimourian: Therapeutic applications of CRISPR/Cas9 system in gene therapy. *Biotechnol Lett* 40, 907-914 (2018)  
DOI: 10.1007/s10529-018-2555-y  
PMid:29704220

**Key Words:** Genetics, Congenital, Consanguinity, Saudi human genome project, Precision medicine, NGS, Review

**Send correspondence to:** Malak Abedalthagafi, King Fahad Medical City and King Abdulaziz City for Science and Technology, P.O. Box 59046, Riyadh 11525, Kingdom of Saudi Arabia, Tel: 966 11 4813202, Fax: 966 1 4813230, E-mail: malthagafi@kacst.edu.sa