Genetic disorders are not equally distributed over the geography of the Arab region. While a number of disorders have a wide geographical presence encompassing 10 or more Arab countries, almost half of these disorders occur in a single Arab country or population. Nearly, one-third of the genetic disorders in Arabs result from congenital malformations and chromosomal abnormalities, which are also responsible for a significant proportion of neonatal and perinatal deaths in Arab populations. Strikingly, about two-thirds of these diseases in Arab patients follow an autosomal recessive mode of inheritance. High fertility rates together with increased consanguineous marriages, generally noticed in Arab populations, tend to increase the rates of genetic and congenital abnormalities. Many of the nearly 500 genes studied in Arab people revealed striking spectra of heterogeneity with many novel and rare mutations causing large arrays of clinical outcomes. In this review we provided an overview of Arab gene geography, and various genetic abnormalities in Arab populations, including disorders of blood, metabolic, circulatory and neoplasm, and also discussed their associated molecules or genes responsible for the cause of these disorders. Although studying Arab-specific genetic disorders resulted in a high value knowledge base, approximately 35% of genetic diseases in Arabs do not have a defined molecular etiology. This is a clear indication that comprehensive research is required in this area to understand the molecular pathologies causing diseases in Arab populations.

Keywords: Arab populations, neolithic, population genetics, gene geography, genetic disorders, neoplasms
A DEFINITION OF ‘ARAB POPULATIONS’

The term “Arabs” indicates a panethnicity of peoples of various ancestral origins, religious backgrounds, and historic identities. It is possible to define the geographical area inhabited by Arabs using one of the two following approaches:

1. The linguistic approach is a relaxed definition and it includes all populations speaking the Arabic language and living in a vast area extending from south of Iran in the east to Morocco in the west including parts in the south-east of Asia Minor, East, and West Africa.

2. The political definition of Arabs is more conservative as it only includes those populations residing in 23 Arab States, namely: Algeria, Bahrain, Comoros, Djibouti, Egypt, Eritrea, Iraq, Jordan, Kuwait, Lebanon, Libya, Mauritania, Morocco, Oman, Palestine, Qatar, Saudi Arabia, Somalia, Sudan, Syria, Tunisia, United Arab Emirates (UAE), and Yemen.

In the subsequent parts of this paper, it is the political definition that would mainly be used to define the term “Arab region” or simply “the region”. In all cases, the Arab geocultural unit is the largest in the world after Russia and Anglo-America. The size of this unit exceeds 375 million people and spans more than 14 million square kilometers.¹

PALEOLITHIC OUT-OF-AFRICA MIGRATIONS

Archeological excavations, historical records, and molecular analyses, mainly based on the study of uniparental Y-chromosome and mitochondrial DNA (mtDNA), provided considerable information regarding the early evolutionary history of modern humans in the vast geographical region embracing Arab populations. The advent of genomic methodologies based on the simultaneous analysis of hundreds of thousands of single nucleotide polymorphisms allowed the drawing of conclusions on the genetic structures of Arab populations with a higher resolution.²

DNA evidence indicates that modern humans originated in East Africa about 200-100 kiloyears (kyr) ago then established regional populations throughout the continent.³ Archeological artifacts excavated from Taforalt in today’s Morocco indicate that human inhabitation of modern day’s Maghreb region (i.e., modern day Morocco, Algeria, Tunisia and Libya) dates back to some 82 kyr ago.⁴ At that time, settlements in the region were characterized by developed cultural manifestations that could only be present in Europe 40 millennia later.⁵ According to the Recent Out-of-Africa model, members of one branch of anatomically modern humans left Africa to the Near East some 70-45 kyr ago.⁶⁷ Phylogenies constructed on the basis of mtDNA comparisons are indicative for two possible migration routes in this episode of human history (see Figure 1):

1. A major route laid across Bab-el-Mandeb straits in the Red Sea linking modern day Eritrea and Djibouti in Africa to Yemen, hence, probably making the Arabian Peninsula as the initial
staging post in the first successful migration of anatomically modern humans out of Africa 70-60 kyr ago. Y-chromosome diversity studies in modern Saudi males support this view as 14% of them exhibit a pool typical of African biogeographic ancestry. High diversity in the Y-haplogroup substructure in samples from the region extends the geography of this active route to include southern Arabia, South Iran, and South Pakistan. This route has possibly maintained its important role in influencing gene flow from Africa along the coastal crescent-shaped corridor of the Gulf of Oman and could have facilitated human dispersals into the region until nearly 2500 years ago.

(2) Another route followed the Nile from East Africa, heading northwards and crossing through the Sinai Peninsula into the Levant and resulted in a noticeable gene flow during the Upper Paleolithic and Mesolithic periods between 40-14 kyr ago. Recent data from Alu/short tandem repeat compound systems and genome-wide polymorphisms are in support with this view with 4-15% of the Levantine groups harboring African ancestry while this influence barely reaches 1-3% in Southern Europeans. Human populations in the Near East then branched in several directions, some heading north into Europe and others heading east into Asia.

Paleoanthropological evidence and mtDNA variation analysis indicate that both the Levantine corridor and the Horn of Africa served, repeatedly, as migratory passageways between Africa and Eurasia. Some of the oldest known genetic mutations that could have followed this route include: (1) the delta F508 (c.1521_1523delCTT) mutation of the CFTR gene, which is responsible today for a majority of cases with cystic fibrosis in Europe, and (2) the p.Glu6Val sickle cell mutation associated with the Benin haplotype and frequently observed in the western coastal region of the Arabian Peninsula, the Levant, Egypt, and in the Maghreb region.

Some studies also support the view that regions near, but external to northeast Africa, like the Levant, the southern-Arabian Peninsula, or Mesopotamia could have served as incubators for the early diversification of non-African lineages and the development of local cultural techniques. Again, the p.Glu6Val sickle cell mutation provides a supportive evidence for this view since the mutation associated with the Arab/Asian haplotype seems to be restricted to the eastern coastal regions of the Arabian Peninsula with milder presence in Mesopotamia and the Levant.

THE EARLY FARMERS

Around 12 kyr ago, Neolithic human populations adapted some developed agricultural technologies that allowed them to cause a far-reaching shift in subsistence and lifestyle. Improvement of the climatic conditions in the area along with the practice of agriculture helped in the establishment of major historical settlements with sizeable densities that could have contributed enormously to the genetic makeup of modern Arab populations. Yet, farming was almost always associated with settlements near mosquito-infested soft and marshy soil causing large malarial outbreaks. Infectious agents that favored humid conditions could have also played major roles in the selective advantage to a variety of other genetic traits. On the other hand, adapting to an active lifestyle along with calorie-restricted diets, common in communities at the time, could have provided protective features that suppressed the expression of celiac disease, type 2 diabetes, and inflammatory bowel disease.

In this phase of human history, the Arabian Peninsula, Sub-Saharan Africa, the Levant and Iran saw local population expansions from refugia that could have participated in the building of the primitive Arabian population. Y-chromosome and mtDNA haplogroup data support this view. For example, approximately 62–69% of today’s males in Saudi Arabia share common structures with those in the near east and this demonstrates a possibly important role for the Levant in shaping the Neolithic dispersal of human settlements in the Gulf. This genetic evidence is consistent with archeological
interpretations of the expansion of sedentary Natufian hamlets in the Levant during the wet phase 15–13 kyr. Male lineage estimates for these prominent Levantine haplogroups indicate a north to south influence with a history of almost 12 kyr in Saudi Arabia, 11 kyr in Yemen, mainly in the western region, and only at nearly 7 kyr in Qatar and the UAE. Detailed analyses hint to a major terrestrial colonization for the eastern Arabian Peninsula, which was followed by subsequent population isolation from the western Arabian Peninsula and demonstrating significant genetic affinities to near-eastern populations. Many of the earliest disease-causing genetic mutations might have followed these steps. In particular, the c.208-2A > G mutation in the human amnionless homolog (AMN) gene, found in 15% of Imerslund-Gräsbeck syndrome cases, could have emerged in the region around 13.6 kyr. Today, this mutation is responsible for over 50% of the Imerslund-Gräsbeck syndrome cases among Arabic, Turkic, and Sephardic Jewish families. On the other hand, studies of mtDNA variability confirm a notable sub-Saharan African female-driven flow in the Arabian Peninsula. An Iranian influence also existed, but this was weakened by the presence of barriers to gene flow posed by the two major Iranian deserts and the Zagros mountain range.

Analysis of the pattern of Y-chromosome and mtDNA variations in North Africa provides evidence of the relatively young population history of North Africa mainly influenced by a strong demic expansion of Neolithic pastoralists from the Levant and possible admixture with original settlers. Some of these earliest civilizations in the Maghreb region include immigrant Berbers who originated from the Sahara 10,000 years ago and left considerable gene imprints in the gene pool of the populations inhabiting the area between modern day Mauritania and southern Egypt. Nearly 2,000 years later, Mesolithic Capsians became the next influential genetic stock in the region.

**MAJOR EVENTS IN ANCIENT HISTORY**

In the Arabian Peninsula, Semitic-speaking peoples of Arabian origin migrated into the valley of the Tigris and Euphrates rivers in Mesopotamia some 7,000-5,500 years ago. Analysis of Y chromosome and mitochondrial DNA in Iraqi Marsh Arabs revealed a prevalent autochthonous Middle Eastern component for both male and female gene pools, with weak Southwest Asian and African contributions. The detailed analysis of genome-wide variation patterns among Qatars indicate that the Southwest Asian influence is derived from Greater Persia rather than from China while the African stock has a sub-Saharan origin and not a Southern African Bantu origin. Data from the neighboring Bahraini and Emirati populations reveal an increasing North-to-South influence of the Southwest Asian component with a high contribution of 23% and 24%, respectively. This could also explain the exceptionally high frequencies of the Asian sickle cell mutation in the region extending from Kuwait to the United Arab Emirates.

Archeological evidence further indicates that another group of Semites left Arabia around 4,500 years ago during the Early Bronze Age and settled along the Levant and mixed in with the local populations there. Some 3,500 years ago, the Phoenician civilization of Lebanon became a developed entreprensing maritime trading culture. Phoenician traders spread across the Mediterranean and established major cities and colonies that harbored their pathologic or polymorphic gene variations. Among the pathologic gene variations that could have followed Phoenician footsteps are (1) the IVS-I-110 (c.93-21G > A) beta-globin gene mutation, the most frequently encountered beta-thalassemia mutation among Arabs, and (2) the p.G542X mutation in the CFTR gene, a frequently observed cystic fibrosis mutation in the Mediterranean basin. Results of the Genographic Consortium from Y-chromosome variations indicate that as many as 1 in 17 men living today on the coasts of North Africa and southern Europe may have a Phoenician direct male-lineage ancestry. The genetic pool was further enriched in Mesopotamia through Persians while Romans gained a 600 year-long period of settlements throughout most of the region that were subsequently replaced by the Byzantines.

**MAJOR EVENTS IN MEDIEVAL HISTORY**

Soon after the rise of Islam 1,400 years ago, the Arab Caliphates unified the region flanking the Mediterranean and amalgamated the dominant ethnic identity that persists today in the Near East, the Levant, the Maghreb, and Andalusia in the Iberian Peninsula. The Arabian Peninsula gained an increasing role and linked distant populations of China and India to communities of the Mediterranean and beyond. During this period, demographical dynamics were predominantly governed by cultural change in endogenous populations rather than demic influences with significant gene flow. This view
is strongly supported by Y-chromosome analysis of Muslim expansion in India and mtDNA haplogroups in the Sinai Peninsula and North Africa.\textsuperscript{24,65,66} During the 11\textsuperscript{th}-13\textsuperscript{th} centuries CE, the Levant witnessed major Crusader settlements that could have caused remarkable genetic drifts and bottlenecks and introduced western European lineages.\textsuperscript{58} In the 16\textsuperscript{th} Century CE, the impact of the western European gene stock extended to the eastern Arabian Peninsula where major parts, including today’s Bahrain, fell under the authority of the Portuguese for nearly 150 years. This presence left clear impressions in the mutational spectrum of common disorders in the eastern Arabian Peninsula as in the frequent observation of the western Mediterranean Codon 39 (c.118C \(\rightarrow\) T) b-thalassemia mutation;\textsuperscript{67-69} reviewed in Obeid and Tadmouri.\textsuperscript{27} On the contrary, some other disorders from the region have possibly spread out to geographically distant locations under this Portuguese influence as demonstrated in the increasing evidence noted with regard to the world distribution of Machado-Joseph disease.\textsuperscript{70,71} During the 13\textsuperscript{th}-19\textsuperscript{th} centuries CE, Ottomans controlled much of the lands surrounding the Mediterranean then expanded their influence to cover all the Arabian Peninsula and further contributed to the enrichment of the genetic pool in the region.\textsuperscript{72} After the 19\textsuperscript{th} century, areas of the Maghreb were colonized by France, Spain and Italy while the Levant, Egypt, and the Arabian Peninsula where colonized by France and England. Despite this long trail of historical admixtures, genetic isolates persisted in the Arab region. Some of these isolates include the inhabitants of the Island of Jerba in Tunisia,\textsuperscript{73} the Bedouins of Sinai,\textsuperscript{65} the dwellers of the Dead Sea region in Jordan,\textsuperscript{74} the Druze of the Levant,\textsuperscript{58} and the Kurdish population of Northern Iraq.\textsuperscript{75}

**THE GENETIC HETEROGENEITY OF ARABS**

Arab populations display some of the highest rates of consanguineous marriages in the world including a large proportion of first cousin marriages.\textsuperscript{76} At a macrogenomic level, this norm permits the reunion of ancestral chromosomal segments in a homozygous pattern referred to as the autozygome.\textsuperscript{77} At a microgenomic level, however, populations in the region exhibit exceptionally high levels of variance within those runs of homozygosity.\textsuperscript{2} This variance seems to follow a sexually asymmetric model with higher heterogeneity recorded among the female groups while paternal lineages are mostly of autochthonous origin.\textsuperscript{29,78} In either way, this variance leads phenotypically to a wide array of more than 1,100 genetic disorders described in the region of which 44\% are confined to a single population or region, a diversity of affected body systems and of clinical outcomes, and a diversity of disease incidence and geographical distributions (reviewed in Tadmouri\textsuperscript{79}).

While the common practice of consanguinity seems to have also contributed to the preponderance of more autosomal recessive (60\%) than autosomal dominant (28\%) disorders in the region,\textsuperscript{76} it is probably the large spectra of pathological gene mutations associated with many genetic disorders in the region that emphasizes the genetic heterogeneity of Arab populations at its best. The following disease families represent few examples of a continuously growing list of disorders related to a long list of mutations many of which have possibly originated in the region.

**BLOOD DISORDERS**

**b-Thalassemia**

b-thalassemia syndromes are a group of hereditary disorders characterized by a genetic deficiency in the synthesis of beta-globin chains. A meta-analysis of 6,652 b-thalassemia alleles from 17 Arab populations indicated the presence of 73 out of the \(~ 250\) b-globin gene mutations occurring worldwide. In contrast to many world populations, this heterogeneity seems to be a common observation in many Arab populations irrespective of the size of pooled b-thalassemia alleles. This case is clearly demonstrated in Algeria, Egypt, Morocco, Tunisia, and the United Arab Emirates exhibiting the largest heterogeneity with more than 20 b-thalassemia mutation types described in each population so far (reviewed in Obeid and Tadmouri\textsuperscript{27}).

**Glucose-6-Phosphate Dehydrogenase (G6PD) deficiency**

G6PD deficiency is an X-linked inherited disorder caused by a defect or deficiency in the production of an important red blood cell enzyme called G6PD. G6PD deficiency may cause the sudden destruction of premature red blood cells leading to hemolytic anemia since the body cannot compensate for the destroyed cells. In Tunisia, the African G6PD*A variant is the most prevalent among G6PD patients and
causes a severe phenotype hemolytic anemia following the ingestion of fava beans. This mutation is also followed by the G6PD*Mediterranean (c.563C > T; p.Ser188Phe) and the G6PD*Aures (c.143T > C; p.Ile48Thr) mutations. The later, was originally described in Algeria and then in Saudi Arabia.81,82 The analysis of mildly affected males, revealed the presence of the association of c.1311C > T, a newly described silent mutation in the exon 12, with the c.93T > C polymorphism in the intron 11 and two single intronic base deletions: IVS-V-17 (-C) and IVS-VIII-43 (-G).83 In Sudan, the G6PD*B variant represents the most common type of enzyme in all the population groups. However, the mutant G6PD*A1 enzyme, but with normal activity, is prevalent among individuals of African descent. Among the deficiency-causing variants G6PD*Mediterranean and G6PD*A1 are the most common.83 The genetic heterogeneity of G6PD further continues in the Arabian Peninsula. In the United Arab Emirates, G6PD*B+ is the major allele described among non-deficient subjects while the G6PD*Mediterranean mutation is the most common cause of G6PD deficiency among Emiri patients.84 Other mutations detected include: the African G6PD*A (c.202G > A) and the G6PD*Aures mutations.84 This spectrum of mutations seems to be common with neighboring Kuwait, where the G6PD*Mediterranean and the African G6PD*A1 genotypes are the most common followed less frequent G6PD*Chatham and G6PD*Aures alleles.85 The Saudi population is also no exception, the G6PD*A−, G6PD*Mediterranean, and G6PD*B+ are the major variants producing a severe deficiency state among affected individuals. These variants exhibit a significant difference in their frequencies, with the highest recorded in areas that were endemic to malaria and have high frequencies of sickle cell disease and b-thalassemia, namely, the Eastern and the Southern Regions.86,87 In neighboring Jordan, molecular screening of G6PD alleles revealed a higher incidence of the disease in Jordan Valley, known for its historically higher rates of malaria, when compared to the Amman area and has also shown the existence of six mutations: the c.563C > T G6PD*Mediterranean mutation (53%), the African G6PD*A1 (c.376A > G + 202G > A; p.Asn126Asp + Val68Met) mutation, G6PD*Chatham (c.1003G > A; p.Ala335Thr), G6PD*Valladolid (c.406C > T), G6PD*Aures (c.143T > C), and G6PD*Asahi (c.202G > A).88 Molecular screening of G6PD alleles in Iraqi Kurdish males indicated that the G6PD*Mediterranean variant was the most common (88%), followed by the G6PD*Chatham variant (c.1003G > A; 9%).89 In a study of 21 unrelated individuals with G6PD*Mediterranean,90 confirmed that almost all patients from Saudi Arabia, Iraq, Iran, Jordan, Lebanon, and Palestine share the c.563C > T mutation.

**METABOLIC DISORDERS**

**Cystic fibrosis**

Cystic fibrosis is a multi-system life threatening inherited disorder that primarily affects the lungs and digestive system. The spectrum of cystic fibrosis mutations in Arab populations reveals a major difference from worldwide observations. For examples, more than 70% of cystic fibrosis patients with European ancestry show the delta F508 (c.1521_1523delCTT) mutation of the CFTR gene. In Arab patients these figures are far from being homogenous. A comprehensive meta-analysis of 827 alleles with cystic fibrosis and encompassing 15 Arab populations revealed a wide spectrum of 56 CFTR gene mutations responsible for the disease in the region (unpublished observations). This heterogeneity seems to continue at regional level as well. For instance, the cystic fibrosis population of the Arabian Peninsula exhibit 17 CFTR mutations. In Saudi Arabians, the 3120 + 1G > A (c.2988 + 1G > A) CFTR mutation is the most common, while in Kuwait it is replaced by the delta F508 mutation. In neighboring Bahrain, three mutations other mutations seem to prevail, these are: 2043delG (c.1911delG), 548A > T (c.416A > T), and 4041C > G (c.3909C > G). This battery of mutations is replaced in Qatar by the commonly observed c.3700A > G (p.I1234V) mutation in the CFTR gene. The picture further changes in Oman and the United Arab Emirates where the c.1647T > G (p.S549R) mutation is common and the delta F508 occurs at relatively low frequencies, but exclusively in patients of Baluchi descent (reviewed in Obeid and Tadmouri27).

**Lipoid congenital adrenal hyperplasia**

This is a severe genetic disorder of steroid hormone biosynthesis, in which the production of all adrenal and gonadal steroids is significantly impaired by a severe defect in the conversion of cholesterol to pregnenolone. Worldwide, lipoid congenital adrenal hyperplasia is caused by nearly 35 mutations in the steroidogenic acute regulatory (StAR) protein gene. Collective results of 20 Arab patients from
Libya, Egypt, Palestine, Jordan, Kuwait, Qatar, Saudi Arabia, and Yemen indicate the presence of 12 mutations including five novel ones in the SSTAR gene (reviewed in Obeid and Tadmouri).

DISORDERS OF THE CIRCULATORY SYSTEM

An extensive survey on genetic disorders in Arab people indicated that there are at least 27 disorders of the circulatory system known to run in Arab families. However, unlike blood disorders and common metabolic abnormalities, appreciation of the genetic etiologies of diseases of the circulatory system has only occurred in the last decade. This resulted in the presence of scanty information that hints to specific genetic signatures characteristic of Arab patients with cardiovascular disorders.

Congenital heart disease (CHD)

CHD is a structural abnormality of the heart or intra-thoracic great vessels. It is the most common birth defect worldwide representing one third of all congenital malformations presenting in the neonatal period. Arabs are liable to have more children with congenital defects including CHD because of high fertility rates. The presence of small isolated communities in different parts of the Arab world with the common practice of consanguinity is another evidence of high incidence of CHD (e.g., Armenians, Bedouins, Druzes, Jews, Kurds, Nubians, Berbers, Tebo, and Twareq). A molecular study in Lebanese CHD patients identified a differential duplication of a 44-bp intronic segment within the Rel-family transcription factor gene, NFATC1, suggestive that this gene could be a potential ventricular septal defect-susceptibility gene. In a prospective study involving 60 Jordanian babies with cleft lip and/or cleft palate, 47% had CHD. However, no chromosomal studies were performed in these patients.

Coronary artery disease (CAD)

A study of Arabs living in Kuwait, found a strong association between a C to G substitution substitution in the 3-prime untranslated region (3'UTR) of the APOC3 gene with coronary artery disease. The population in the study included adults from Kuwait, Jordan, Palestine, Lebanon, Syria, Egypt, and Iraq. In Saudi individuals, CAD was also found to be associated with the 3'UTR allele of the APOC3 gene, but also other associations were found with the MTHFR c.677C > T variant, a platelet glycoprotein receptor IIIa (PIA1/PIA1) genotype, and the null-genotypes of GSTT1 and GSTM1. In support of a probable specificity of the genotypic etiology of coronary artery disease in Arabs, no association was found with the lipoprotein lipase (LPL) polymorphisms (LPL-HindIII and LPL-PvuII), the infrequent band of 3.2-kb of the apolipoprotein A-I/C-III; the insertion/deletion sites in the polymorphic region of intron 16 of the angiotensin I-converting enzyme (ACE) gene, the p.W64R polymorphism of the b3-adrenoceptor (b3-AR) gene, PvuII polymorphism in the LPL gene, and the c.677C > T and c.1298A > C variants of the MTHFR gene.

Hypertrophic cardiomyopathy (HCM)

HCM is characterized by an abnormal thickening of the heart muscles, resulting from mutations in one of several genes that result in defects in the protein component of the cardiac muscles. An apical hypertrophic cardiomyopathy in father and daughter of a Lebanese Christian family has been reported. In both, identical segments of the left ventricle were involved by the hypertrophic process with differing degrees of severity. In an analysis of data pertaining to all patients less than 50-years of age in Qatar, six of 42 Qatars were diagnosed with HCM, making it the most encountered cardiomyopathy in this group following dilated cardiomyopathy. HCM occurred in two peaks: one below 15-years of age, and the other between 36 and 50-years of age. About 27% of the children (between 1- and 15-years) were found to have HCM. The prevalence rate of HCM was calculated as 3.1 per 100,000 of the population.

Arterial tortuosity syndrome

Probably, the earliest account of the disease in the region dates back to year 2000 with the description of 12 patients from eight different families in Saudi Arabia. The first mutations associated with the disease, however, were reported six years later in patients of Moroccan origin who had homozygosity for the c.510G > A (p.W170X) and for a frameshift c.961delG (p.V321fsX391) mutation in the SLC2A10 gene. In Qatar, two mutations, a novel p.R105C and a recurrent p.S81R, were recently described in the SLC2A10 gene in seven patients from two unrelated families.
Other disorders

In two consanguineous Saudi families, long QT syndrome (LQTS) was described as segregating with a novel homozygous splicing mutation in the KCNQ1 gene. The observation of the same mutation in both families indicated that this could be a founder mutation. On the contrary, Naxos disease, a rare cardiomyopathy disorder, failed to exhibit linkage with the previously identified plakoglobin gene in two Saudi patients indicating that the disease might have a private signature in the region.

NEOPLASMS

Neoplasms are not typically regarded as population-specific disorders. However, several aspects of these disorders differ by race and ethnicity. Among Arabs, several types of cancers show many distinct features that are quite different from those seen in other populations worldwide. Very preliminary data from the CTGA (Catalogue for Transmission Genetics in Arabs) Database for genetic disorders in Arab populations indicate the presence of at least 55 cancer types in Arab people. Breast, ovarian, lung, and colorectal cancers are the main cancers that run in Arab families. Cancer susceptibility genes for many of these cancers have been reported. Yet, other cancers with familial types such as prostate, pancreatic, and testicular cancers did not reveal specific cancer-susceptibility genes at this time.

Breast and ovarian cancer

Broadly speaking, 90% of breast cancer cases are sporadic and the processes leading to gene mutations in such cases are not well-understood. Defined genetic predisposition accounts for only about 5–10% of inherited breast cancer types. In either familial or sporadic cases, multiple genetic etiologies, related to mutations in oncogenes and tumor suppressor genes, characterize breast carcinomas in Arab patients. A large fraction of inherited cases of breast cancer are usually associated with mutations of the BRCA1 and BRCA2 genes. Other genes have also been implicated, such as: BRCAT, BRCA3, TP53, BRIP1, PTEN, and STK11 genes. In sporadic breast cancer, increased susceptibility has been blamed on the mutation of low penetrance genes including TNFA, HSP70-2, and TNFRII. These private signatures of the disease in the region have probably contributed to the peculiar clinical characteristics of the disease in Arab women particularly the earlier mean age of onset, which is at least a decade earlier than in women of other ethnicities, and the more aggressive course of the disease.

According to a study by Rouba et al., the proportion of BRCA1 and BRCA2 mutations could be higher in Arab women when compared to other populations. In Morocco, five deleterious mutations in the BRCA1 gene where encountered in families with breast/ovarian cancer, including the novel compound deletional c.2805delA/2924delA mutation. In Algerian women, four of 11 familial cases were associated with BRCA1 alterations. In neighboring Tunisia, the prevalence of breast cancer is calculated to be between 16% and 38%. There, four BRCA1 mutations have been identified including a novel Tunisian-specific c.212+2insG mutation and a frequently observed c.798_799delTT Tunisian and North African founder mutation. In Egypt, the p.Arg841Trp BRCA1 disease-associated mutation was detected while a novel p.Glu1373X mutation in exon 12 of the BRCA1 gene was identified in ovarian or breast cancer patients in Arab kindred from East Jerusalem. An extensive analysis of familial breast cancer in Lebanon revealed the presence of 38 BRCA1 sequence variants, many of which are novel. Adding to this heterogeneity, two other unclassified BRCA1 variants, p.Phe486Leu and p.Asn550His, were detected in Saudi patients.

In the case of BRCA2 gene, the scene is far from being different. Four mutations in BRCA2 gene cause breast/ovarian cancer in Moroccan families including three novel ones (c.3381delT/3609delT; c.710delE/A733delA, and c.2753insG/3463insG). The same study also identified a large number of distinct polymorphisms and unclassified variants in BRCA2 as well as in BRCA1 that were described for the first time. In four unrelated Tunisian families, two novel c.1313dupT and c.7654dupT mutations in exons 10 and 16 of the BRCA2 gene were reported. In an Arab patient of Palestinian descent with breast cancer, the c.2482delGACT novel BRCA2 truncating mutation was observed. An extensive analysis of familial breast cancer in Lebanon revealed the presence of 40 BRCA2 gene sequence variants, many of which are novel. In Saudi patients, an unclassified p.Asp1420Tyr BRCA2 variant was detected. This array of region-specific mutation seems to extend to Arab Diasporas as well. For example, the c.5804delA4 mutation in exon 11 of BRCA2 gene was seen in nearly half of the carriers of...
deleterious mutations in Arab American women. This mutation has not been previously associated with a particular Arab ethnicity and may represent a founder mutation of recent origin.125

Another frequently mutated gene in Arab breast cancer patients is the TP53 gene. In fact, the frequency of TP53 mutations among Saudi patients is one of the highest in the world. The list of mutations include seven novel ones of which five are found in exon 4 of the TP53 gene. In brief, tumors from Arab breast cancer patients have a high prevalence (29%) of TP53 mutations in exons 4 and 5, whereas the smallest proportion of TP53 mutations (10%) is found in exon 7. Also, an excess of G:C → A:T transitions (49%) at non-CpG sites was noted, suggesting exposure to particular environmental carcinogens such as N-nitroso compounds.126 In addition, several single nucleotide polymorphisms in Arab patients seem to be specific to the indigenous populations and could be associated with increased risk of breast cancer. Examples include: the p.Pro72Pro in the TP53 gene and the c.309GG in the MDM2 gene in Saudi women, the c.-251A IL8 allele in Tunisian women, and the c.1298A > C DNA polymorphism in the MTHFR gene in patients of Syrian ancestry.127-129

In western societies, mutation of the TP3 gene is highly associated with epithelial ovarian cancers (50–80%), however, only 32% Arab patients with this neoplasm exhibit TP3 gene mutations. Instead, PIK3CA amplification, but not PIK3CA mutation, is the single most common genetic alteration in Arab cases (60%) and is mutually exclusive with gene mutations in both PI3 Kinase and MAPK pathways (PIK3CA, KRAS, and BRAF).130-132 This finding is suggestive for a significant role of the dysregulated PI3K/Akt pathway in the pathogenesis of ovarian cancers.132

Colorectal carcinoma (CRC)

This type of neoplasm is a further example demonstrating a genetic heterogeneity in the region in which not only different alleles of the same gene are involved, but also several genes seem to be of importance for the emergence of this ailment. In Moroccan patients with attenuated polyposis, the homozygous p.Tyr165Cys and c.1186_1187insGG mutations of the MYH gene were reported133,134 whereas in neighboring Tunisia, a large deletion involving exon 6 of the MLH1, a DNA mismatch repair, gene was observed in a family with six patients diagnosed with a colorectal or an endometrial cancer and characterized by a severe phenotype and an early onset.135 Another study in Tunisians demonstrated a significant association between the p.E1317Q, p.D1822V, and p.I1307K variants of the adenomatous polyposis coli (APC) gene with colorectal carcinoma risk.136 The p.I1307K mutation seems to have a long history in the region as demonstrated in the repeated observation of the allele among many populations in the region. In 1999, the p.I1307K mutation was first described among Ashkenazi and Yemenite Jews.137 A study on the general population demonstrated a carrier frequency of the allele in Yemenite Jews of approximately 5%.138 A more extensive analysis showed the p.I1307K mutation existed in Sephardi Jews of Syrian, Egyptian, Moroccan, Yemeni, and Palestinian origins, as well as in Muslim and Christian individuals of Arab descent. This study also demonstrated that the ancestor of modern p.I1307K alleles existed some 2.2-2.95 kya.139 The portrait of colorectal carcinoma further gets more interesting with the presence of a recent study that investigated the methylation patterns in colorectal carcinoma from Egypt and Jordan and showed that differing gene methylation patterns and mutation frequencies are also involved, hence, indicating dissimilar molecular pathogenesis and probably reflecting different environmental exposures.140

Prostate cancer

In Tunisians, a significantly increased prostate cancer risk was associated with the VEGF-634 (GC + CC) combined genotype while the VEGF-634C allele was associated with high histological grade. However, the VEGF-1154A/-634G haplotype was negatively associated with prostate cancer risk and high tumor grade.141 No association was observed between the p.N700S TSP1 polymorphism and prostate cancer risk or severity. Yet, subjects carrying one copy of the MMP9-1562T allele exhibited a threefold higher risk of developing prostate cancer.142

Other neoplasms

The CYP3A4*2C, GSTT1 null, and GSTP1 TT genotypes demonstrated significant association with diffuse large B-cell lymphoma (DLBCL) in the Saudi population.143 The CYP3A4 c.4887C > A genotypes CA, AA and variant allele A were demonstrated to have significant differences and greater risk of developing papillary thyroid cancer in Saudi patients compared to wild type genotype CC. Also, in thyroid cancer,
GSTT1 null showed higher risk while GSTM1 null showed protective effect. Tunisian smokers carrying this later allele had an approximately 2.2-fold high risk of bladder cancer. Furthermore, individuals carrying at least one copy of the methionine synthase (MS) c.2756A > G variant allele and heterozygous for the c.1298A > C MTHFR polymorphism displayed a 2.33 and 1.8 times increased risk of developing bladder cancer, respectively.

FINAL NOTE
A multitude of studies reviewed in this paper clearly indicate that the Arab region was an important milieu for the early adaptations of modern human populations to the out-of-Africa environment. The experiences learned in that period certainly have allowed human populations to establish further settlements and cover many areas in the rest of the world. The tidal movements of historical populations in and out of the Arab region allowed the area to become an important bridge for the flow of genes between Africa, Asia, and Europe. This characteristic made the area a focal point of attraction for many population geneticists seeking to fill the gap in the interpretation of benign or lethal genomic variations in world populations.

While we could be fascinated with the extent of the genetic heterogeneity that characterizes Arab population, understanding the genetic structure of populations and exploring their biogeographical heterogeneities may also yield a better understanding of the genetic processes and, eventually, disease etiologies in the region. In many instances, studying Arab families, with Arab-specific genetic disorders, has resulted in a high value knowledge base and linked many genes to well-defined phenotypes and helped a great deal in global genome annotation efforts. Yet, many of the nearly 500 genes studied in Arab people revealed striking spectra of heterogeneities with many rare and novel mutations causing large arrays of clinical outcomes, thus, considerably complicating proper counseling and diagnosis for many disorders. Unfortunately, the materialization of large-scale personalized medical genomics may not be expected in the near future especially because of the presence of hundreds of genetic disorders in Arabs with no defined molecular determinants and because of the restricted economies to sustain genomic research throughout the region.

REFERENCES

[1] US Census Bureau. http://www.census.gov/population/international/data/idb/informationGateway.php. visited: 23.2.2014.
[2] Hunter-Zinck H, Musharoff S, Salit J, Al-Ali KA, Chouache L, Gohar A, Matthews R, Butler MW, Fuller J, Hackett NR, Crystal RG, Clark AG. Population genetic structure of the people of Qatar. Am J Hum Genet. 2010;87(1):17–25.
[3] Liu H, Prugnolle F, Manica A, Balloux F. A geographically explicit genetic model of worldwide human-settlement history. Am J Hum Genet. 2006;79(3):230–237.
[4] Fernandes V, Alshamali F, Alves M, Costa MD, Pereira JB, Silva NM, Cherni L, Harich N, Cerny V, Soares P, Richards MB, Pereira L. The Arabian cradle: Mitochondrial relics of the first steps along the southern route out of Africa. Am J Hum Genet. 2012;90(2):347–355.
[5] Bouzouggar A, Barton N, Vanhaeren M, d’Errico F, Collett S, Higham T, Hodge E, Parfitt S, Rhodes E, Schwenninger L, Stringer C, Turner E, Ward S, Moutmir A, Stamboul A. 82,000-year-old shell beads from North Africa and implications for the origins of modern human behavior. Proc Natl Acad Sci U S A. 2007;104(24):9964–9969.
[6] Relethford JH. Genetic evidence and the modern human origins debate. Heredity. 2008;100(6):555 – 563.
[7] Fernandes V, Alshamali F, Alves M, Costa MD, Pereira JB, Silva NM, Cherni L, Harich N, Cerny V, Soares P, Richards MB, Pereira L. The Arabian cradle: Mitochondrial relics of the first steps along the southern route out of Africa. Am J Hum Genet. 2012;90(2):347–355.
[8] Bailey GN, Flemming NC, King GCP, Lambeck K, Momer G, Moran LU, Al-Sharkei A, Vita-Finzi C. Coastlines, submerged landscapes, and human evolution: The Red Sea Basin and the Farasan Islands. J Island Coastal Archaeol. 2007;2:127 – 160.
[9] Cerny V, Mulligan CJ, Ridl J, Zaloudkova M, Edens CM, Hajek M, Pereira L. Regional differences in the distribution of the sub-Saharan, West Eurasian, and South Asian mtDNA lineages in Yemen. Am J Phys Anthrop. 2008;136(2):128–137.
[10] Abu-Amero KK, Hellani A, Gonzalez AM, Lamuga JM, Cabrera VM, Underhill PA. Saudi Arabian Y-Chromosome diversity and its relationship with nearby regions. BMC Genet. 2009;10:59.
[11] Cadenas AM, Zhivotovsky LA, Cavalli-Sforza LL, Underhill PA, Herrera RJ. Y-chromosome diversity characterizes the Arabian Peninsula. BMC Genet. 2008;9:169.
[12] Cann RL, Stoneking M, Wilson AC. Mitochondrial DNA and human evolution. Nature. 1987;325(6099):31 – 36.
[13] Ingman M, Kaessmann H, Pääbo S, Gyllensten U. Mitochondrial genome variation and the origin of modern humans. Nature. 2000;408(6813):708–713.
[14] Luis J, Rowold DJ, Regueiro M, Caeiro B, Cinnioglu C, Roseman C, Underhill PA, Cavalli-Sforza LL, Herrera RJ. The Levant versus the Horn of Africa: Evidence for bidirectional corridors of human migrations. Am J Hum Genet. 2004;74(3):532–544.
[15] Pérez-Miranda AM, Alfonso-Sánchez MA, Peña JA, Herrera RJ. Qatari DNA variation at a crossroad of human migrations. Hum Hered. 2006;61(2):67–79.
Laarabi FZ, Cherkaoui Jaouad I, Baert-Desurmont S, Ouldim K, Ibrahimi A, Kanouni N, Frebourg T, Sefiani A. The first
Bougatef K, Marrakchi R, Ouerhani S, Sassi R, Moussa A, Kourda N, Blondeau Lahely Y, Najjar T, Ben Jilani S,
Drucker L, Shpilberg O, Neumann A, Shapira J, Stackievicz R, Beyth Y, Yarkoni S. Adenomatous polyposis coli I1307K
Genetic disorders among Arab populations
Teebi AS.
Abubaker J, Bavi P, Al-Haqawi W, Jehan Z, Munkarah A, Uddin S, Al-Kuraya KS. PIK3CA alterations in Middle Eastern
Wang Y, Helland A, Holm R, Kristensen GB, Børresen-Dale AL. PIK3CA mutations in advanced ovarian carcinomas.
AlShatwi AA, Hasan TN, Shafi G, Alsaif MA, Al-Hazzani AA, Alsaif AA. A single-nucleotide polymorphism in the TP53
Uhrhammer N, Abdelouahab A, Lafarge L, Feillel V, Ben Dib A, Bignon YJ, Mégarbane A. Prevalence of BRCA1 and
Tazzite A, Jouhadi H, Madel W, Aretini P, Falaschi E, Collavetti A, Benider A, Caligo MA. BRCA1 and BRCA2 germline
Tariq M, Khodjet El Khil H, Jalalbert T, Mahfoudh W, Chouchane L, Ayed FB, Bignon YJ, Elgaied AB. Complete mutation
Mahnoufi N, Bouaouina N, Ahmed SB, Gabbouj S, Shan J, Mathew R, Uhrhammer N, Bignon YJ, Troudi W, Elgaied
Mahfoudh W, Bouaouina N, Ahmed SB, Gabbouj S, Shan J, Mathew R, Uhrhammer N, Bignon YJ, Troudi W, Elgaied
Mahfoudh W, Bouaouina N, Ahmed SB, Gabbouj S, Shan J, Mathew R, Uhrhammer N, Bignon YJ, Troudi W, Elgaied
El-Harith el-HA, Abdel-Hadi MS, Steinmann D, Dork T. BRCA1 and BRCA2 mutations in breast cancer patients from
Abu-Halaf N, Badrriyah M, Abi-Halaf N, Badrriyah M, Abeliovich D, Peretz T. A novel BRCA-1 mutation in Arab
Badrriyah M, Abeliovich D, Peretz T. A novel BRCA-1 mutation in Arab
Badrriyah M, Abeliovich D, Peretz T. A novel BRCA-1 mutation in Arab
Ahmed SB, Gabbouj S, Shan J, Mathew R, Uhrhammer N, Bignon YJ, Troudi W, Elgaied
Al-Qasem AJ, Toulimat M, Eldali AM, Tulbah A, Al-Yousef N, Al-Daihan SK, Al-Tassan N, Al-Tweigeri T, Aboussekhra A.
Al-Harith el-HA, Abdel-Hadi MS, Steinmann D, Dork T. BRCA1 and BRCA2 mutations in breast cancer patients from
Fraker T, Ghabreau L, Alachkar A. Association of polymorphisms in one-carbon metabolizing genes with ovarian
Kandilci A, Omar S, Kapran Y, Dizdaroglu F, Bondy ML, Amos CI, Issa JP, Levin B, Hamilton SR. Differing DNA
Kanj A, Colvin D, Pepe S, El-Naggar A, El-Shabrawi-Cala E. Prevalence and spectrum of BRCA1 in Jordanian and
Am J Hum Genet. 2003;73(6):1250–1260.
Clin Genet. 2012;82(3):265-274.
Cancer. 2009;115(12):2875–2878.
Clin Genet. 2012;81(1):26-31.
J Clin Oncol. 2009;27(24):4025–4032.
Clin Cancer Res. 2005;11(8):2875–2878.
Mol Biol Rep. 2013;40(1):243–253.
Cancer. 2007;108(2):562–568.
Cancer Res. 2003;63(23):8281–8287.
Cancer Genet Cytogenet. 2007;172(1):38–44.
Al-Dayel F, Al-Rasheed M, Ibrahim M, Bu R, Abubaker J, Al-Jomah N, Mohamed GH, Moorji A, Uddin S, Siraj AK, Al-Kuraya K. Polymorphisms of drug-metabolizing enzymes CYP1A1, GSTT and GSTP contribute to the development of diffuse large B-cell lymphoma risk in the Saudi Arabian population. *Leuk Lymphoma.* 2008;49(1):122–129.

Siraj AK, Ibrahim M, Al-Rasheed M, Abubaker J, Bu R, Siddiqui SJ, Al-Dayel F, Al-Sanea O, Al-Nuaim A, Uddin S, Al-Kuraya K. Polymorphisms of selected xenobiotic genes contribute to the development of papillary thyroid cancer susceptibility in Middle Eastern population. *BMC Med Genet.* 2008;9:61.

Ouerhani S, Tebourski F, Slama MR, Marrakchi R, Rabeh M, Hassine LB, Ayed M, ElGaaied AB. The role of glutathione transferases M1 and T1 in individual susceptibility to bladder cancer in a Tunisian population. *Ann Hum Biol.* 2006;33(5-6):529–535.

Ouerhani S, Oliveira E, Marrakchi R, Ben Slama MR, Sfaxi M, Ayed M, Chebil M, Amorim A, El Gaaied AB, Prata MJ. Methylenetetrahydrofolate reductase and methionine synthase polymorphisms and risk of bladder cancer in a Tunisian population. *Cancer Genet Cytogenet.* 2007;176(1):48–53.

Ozçelik T, Kanaan M, Avraham KB, Yannoukakos D, Mégarbane A, Tadmouri GO, Middleton L, Romeo G, King MC, Levy-Lahad E. Collaborative genomics for human health and cooperation in the Mediterranean region. *Nat Genet.* 2010;42(8):641–645.