Specific Aspects of Consanguinity: Some Examples from the Tunisian Population

Lilia Romdhane a Nizar Ben Halim a Insaf Rejeb a Rym Kefi a
Yosra Bouyacoub a Mariem Ben Rekaya a Habib Messai a Olfa Messaoud a
Zied Riahi a,b Crystal Bonnet b Faten Ben Rhouma a Majdi Nagara a
Christine Petit b Ken McElreavey c Giovanni Romeo d Sonia Abdelhak a

a Biomedical Genomics and Oncogenetics Laboratory LR11IPT05, Institut Pasteur de Tunis, Université Tunis El Manar, Tunis, Tunisia; b Unité de Génétique et Physiologie de l’Audition, INSERM UMR 587, and c Human Developmental Genetics Unit, Pasteur Institute, Paris, France; d Unità Operativa di Genetica Medica, Dipartimento di Scienze Mediche e Chirurgiche, Policlinico Sant’Orsola-Malpighi, Bologna, Italy

Key Words
Consanguinity · Endogamy · Founder effect · Genetic disease · Founder mutation · Tunisian population

Abstract
Located at the cross-road between Europe and Africa, Tunisia is a North African country of 11 million inhabitants. Throughout its history, it has been invaded by different ethnic groups. These historical events, and consanguinity, have impacted on the spectrum and frequency of genetic diseases in Tunisia. Investigations of Tunisian families have significantly contributed to elucidation of the genetic bases of rare disorders, providing an invaluable resource of cases due to particular familial structures (large family size, consanguinity and share of common ancestors). In the present study, we report on and review different aspects of consanguinity in the Tunisian population as a case study, representing several features common to neighboring or historically related countries in North Africa and the Middle East. Despite the educational, demographic and behavioral changes that have taken place during the last four decades, familial and geographical endogamy still exist at high frequencies, especially in rural areas. The health implications of consanguinity in Tunisian families include an increased risk of the expression of autosomal recessive diseases and particular phenotypic expressions. With new sequencing technologies, the investigation of consanguineous populations provides a unique opportunity to better evaluate the impact of consanguinity on the genome dynamic and on health, in addition to a better understanding of the genetic bases of diseases.

Introduction
In clinical genetics, a consanguineous marriage commonly refers to a union between individuals related as second cousins or closer [1]. It has been estimated that at least 1,100 people live in communities with strong prefer-
ence for intra-familial unions and some 10.4% of the global population are the progeny of consanguineous parents [2]. Consanguinity rates vary around the world according to religion, culture and geographical localization. The highest prevalence of consanguineous marriage has been recorded in North Africa, the Middle East, Central and Southern Asia, with between 20 and more than 50% of all unions between second cousins or closer [1]. In population genetics, consanguinity may also refer to marital unions of individuals with at least one common ancestor, such as those occurring within population isolates, small towns, tribes and intra-communities. In mathematical terms, the probability that a child of consanguineous parents will be autozygous for a particular allele identical by descent (IBD) is given by the coefficient of inbreeding: $F$ [3]. In first cousin unions, a couple is predicted to have 1/8 of their genes IBD. As a consequence, their progeny would be expected to be homozygous at 1/16 of all loci on average, i.e. $F = 0.0625$ [4].

Among Arabs, intra-familial unions include double first cousins, first cousins, first cousins once removed and second cousins. Unlike in some other consanguineous societies, uncle-niece marriage is prohibited in Islam and therefore is absent among Arabs [5]. Tunisia is a North African country of about 11 million inhabitants. The majority of the population is defined as Arab Muslim, with a minority of Jews and Christians who contribute to the cultural diversity [6]. Throughout its history, the native population, the Amazigh (known as Berbers), has witnessed successive invasions by different ethnic groups which have shaped and impacted on the genetic landscape of the current population [6]. The main demographic and social features of the Tunisian population are the large size of families with more than one affected individual, different generations living in the same place and often an association with consanguinity. Demographic, socioeconomic, cultural and historical events have impacted on the distribution and the frequency of genetic diseases in Tunisia [6, 7]. In the present article, we focused on specific aspects of consanguinity in Tunisia, as a case study representing several features common to neighboring or historically related countries in North Africa and the Middle East.

**Frequency and Specific Aspects of Consanguinity**

A limited number of studies have estimated the frequency of consanguineous marriage in the Tunisian population. In 1975, the prevalence of unions between relatives was reported to be 19.2% [8, 9], and in a sample of 5,767 individuals, Riou et al. [10] observed 26.9% consanguineous marriages in Northern Tunisia. In the same region, more precisely in the governorate of Bizerte, a higher frequency of 39.3% consanguinity also was noted [11]. In two studies in the governorate of Monastir, in the central region of the country, consanguineous marriages were shown to be 20.1 and 24.8%, respectively [12, 13]. A more recent work conducted in different localities of the country indicated an overall rate of 29.8% consanguinity [14]. First-cousin unions ($F = 0.0625$) were the most common type of consanguineous union and ranged in prevalence between 16.7 and 23.0% of marriages [10–14]. A preference for the patrilateral parallel-cousin marriage (father’s brother’s children) was also highlighted [14].

Geographical endogamy, the matrimonial rule of unions between individuals belonging to the same region, is also still frequent in Tunisia, as illustrated in three studies in three regions of the country [15–17]. A frequency of 87.6% of endogamous matings was found, based on the origin (governorate) category of the parents, with 96.8% endogamy in rural areas of Nabeul in the preceding generation [15–17]. The same situation was found in the Berber village of Douiret, where the frequency of endogamous unions remains high (>96%) among the couples in the study, their parents and their grand-parents [unpubl. data]. All of these findings confirm that familial and geographical endogamy is still very common in Tunisia despite the impressive educational, demographic and behavioral changes that have taken place during the last four decades [14].

As in other North African and Middle Eastern countries a typical matrimonial pattern of choice is adopted in Tunisia with, in terms of priority, a man marrying his first cousin, a distant relative or a neighbor. This form of choice could be considered as an obligation, since any deviation with respect to these rules could lead to social reproach or even familial sanction [9]. Spousal selection can also be influenced by geographical proximity, especially in rural surroundings. Indeed, the distribution patterns of unions between relatives show a clear predominance in rural areas at 25.4% consanguinity, compared to 16.1% in urban regions [8, 9]. Kin marriage is thus often encouraged by multiple economic, social, affective and cultural factors, expressed in the traditional saying ‘Who would give his wealth to others?’, where wealth includes children, and with cousins promised to each other in marriage at a young age [9, 14]. Among unions between related individuals, the actual pattern of marital choice was supposed to be conserved from parents to children [16].
However, in a majority of cases, a collective society model appears to be followed with assumptions of dependence on adherence to patterns of parental behavior overstated [16].

**Founder Effect, Consanguinity and Genetic Disease Frequency**

Founder effect occurs when a small group of persons breaks off from a larger population and creates a new colony or, in the case of a bottleneck, which results from wars or epidemics. In the first case the subpopulation is formed by newcomers, more especially invaders, or less commonly autochthonous inhabitants escaping persecution by invaders, who withdraw to urban peripheries, mountains or deserts. Both types of situation are encountered in Tunisia, since the country has experienced invasions by different ethnic groups and has witnessed successive epidemic waves of plague and starvation throughout its history, with potential founder effects in different parts of the country [6].

Both large family and patriarchal tribe models based on common ancestry, and other features of Arab societies, are responsible for the specific distribution patterns of genetic diseases that are typical of Arab and many Middle Eastern populations [14, 18]. A cumulative history of consanguineous unions may prompt specific enrichment of founder mutations, a high prevalence of unique genetic disorders as well as relative homogeneity of the mutation spectrum of these populations. In Tunisia, the patterns of founder mutation distribution disclosed two classes of ancestral genetic lesions: those that could be shared with other populations and those that are Tunisia specific [6].

Allelic homogeneity is encountered in several genetic diseases in Tunisia [6]. Crigler Najjar syndrome type I and Xeroderma pigmentosum (XP) group C are caused by the founder mutations p.Q357R and p.V548AfsX25, respectively [19–21], that are shared with other Mediterranean populations. In other cases, some allelic heterogeneity is responsible for the expression of the morbidity phenotype, with a predominant founder mutation as the major cause of the disease along with rare familial private molecular lesions. For example, limb-girdle muscular dystrophy 2C (LGDM2C) is caused by three mutations one of which, the founder 521delT, is by far the most frequent [22, 23]. The same situation is encountered for the founder mutation p.R228X, encountered in more than 80% of XP group A patients. Most XP group A cases additionally are concentrated in the same geographical area of Central Tunisia [24, 25]. The third example is illustrated by the p.W1327X mutation responsible for glycogenosis type III that is frequently observed in the homozygous state and mainly occurs in a coastal region of Central Tunisia [26, 27]. Taking into account historical events in this region, and given the occurrence of the p.W1327X mutation among Turkish patients [28], this gene lesion could have been introduced to Tunisia by Turkish settlers, although confirmatory haplotype analysis would be needed to support this hypothesis.

Allelic homogeneity is not systematically found in consanguineous and isolated populations. Compound heterozygosity for the GJB2 founder mutations c.35delG and p.E47X in a consanguineous family originating from an isolate in Northern Tunisia challenges the classical concept of unique founder mutation segregation in such isolates [29]. Allelic heterogeneity within the same families has been observed for several diseases including ataxia telangiectasia, Stargardt disease [unpubl. data] and XP group A [30]. For the latter, a previously undescribed clinical presentation was observed due to the combination of a founder mutation and a new mutation that had arisen in one of the family branches [31]. In another case, compound heterozygosity resulted in a misleading linkage analysis result [32].

**Health Impact of Consanguinity**

Genetic diseases constitute a real health burden in developing countries. In a recent update of a previous study on the spectrum of genetic diseases in Tunisia [7], over 400 genetic disorders have been identified. In our studies, phenotypic and epidemiological particularities have been observed [unpubl. data]. In a large patient series from our laboratory, consanguinity is often present at a prevalence of ≥80%. Geographical endogamy is also apparent, with the same range of values reflecting the geographical and social isolation of patients’ families. Moreover, the results reveal the fragmentation of the Tunisian population into isolated communities, resulting in an increased risk of genetic disease expression [unpubl. data]. Although few studies have focused on the health impact of consanguinity due to the scarcity of epidemiological reports, some authors reported a correlation between intra-familial unions and genetic disorders such as polydactyly, deafness, degenerative spinocerebellar diseases, bipolar disorder, metabolic...
disorders and XP [33–35]. From a reproductive health perspective, consanguineous couples were found to have a higher fertility index than non-consanguineous couples. Although the rates of spontaneous abortions and stillbirths were not correlated with consanguinity, higher rates of neonatal and post-neonatal deaths, and deaths of under 5-year-old children were observed in consanguineous couples [13]. In this particular context of fertility and newborn mortality, additional factors such as the socioeconomic status of parents, pregnancy intervals, contraceptive use and residence should also be taken into consideration [13].

In our recent studies [unpubl. data], we evaluated the risk associated with consanguinity in a large class of Mendelian monogenic and multifactorial conditions. We found that parental consanguinity is associated with an 8–9-fold increased risk of expression for a recessive genetic disorder (OR = 8.53, CI = 6.70–10.86, p < 10−3). After correction for confounding demographic variables like age, sex and geographical origin, the associated risk remained high with an approximately 6-fold increase, which is in agreement with previous reports on consanguineous populations, mainly from the Middle East. On this basis, consanguinity must still be considered as a major factor in the expression of inherited conditions due to recessive genes [14]. The excess risk associated with familial endogamy increases significantly according to the degree of consanguinity [14], but also differs according to the specific hereditary pathology. Thus consanguineous marriage results in an approximately 7-fold relative risk in the expression of metabolic disorders. Within this disease group, both primary hyperoxaluria and glycogenosis type III show high ORs associated with consanguinity, i.e. ORs of 15 (OR = 15.01, CI = 1.81–124.19, p = 0.012) and 13 (OR = 12.87, CI = 3.68–45.01, p < 10−4), respectively [unpubl. data]. Similarly, the relative risk due to the 35delG mutation, the most frequent allele for nonsyndromic recessive deafness in Tunisia, is increased 10-fold for first-cousin offspring [36]. Lastly, the consanguinity-associated relative risk for XP group A expression was estimated to be 26 (OR = 25.77, CI = 11.15–59.56, p < 10−5) [unpubl. data].

Mining the scientific literature, available in bibliographical databases, and genetic data, stored in DNA variant repositories, is an alternative means of assessing both the implied influence of consanguinity on the genetic disease spectrum of the Tunisian population and phenotypic and distribution particularities [6, 7]. The current meta-analysis consisted first of collecting information on all genetic diseases described in the Tunisian population into a relational database, with both known and unknown molecular etiologies [unpubl. data]. This step was possible due to publications on genetic diseases affecting Tunisian individuals from both national investigations and international collaborations. As the excess risk that an autosomal recessive disorder will be expressed in the progeny of consanguineous unions is inversely proportional to the frequency of the disease allele in the gene pool, highly inbred Tunisian families with multiple affected members were extensively recruited in genetic linkage studies. Their investigation using homozygosity mapping largely contributed to the elucidation of the genetic basis of each disease [7]. Stratification of the database entries according to the mode of transmission indicated that 425 genetic disorders affecting Tunisian patients have been described, of which 61% were autosomal recessive and 50% were due to at least one mutation [unpubl. data]. Among the 600 molecular lesions identified in Tunisian patients, most often in the homozygous state, those derived from a common ancestor caused 104 genetic conditions of which 82% were recessively inherited [unpubl. data]. To our knowledge, this is the first meta-analysis that gives an overview of the current spectrum of genetic diseases in a consanguineous population. For comparative purposes, a similar exhaustive study on genetic diseases in a population with a low rate of consanguinity is warranted.

In a population where consanguinity is deeply rooted, intra-familial marriage does not alter the probability of an autosomal dominant disease being expressed but could affect the disease phenotype. Homozygosity for a mutation in the COL7A1 gene leading to an abnormal structural protein and responsible for dystrophic epidermolysis bullosa has caused a more severe phenotype of the disease in a consanguineous family [37]. Also, the high prevalence of familial hypercholesterolemia (FH) in Central and Southern Tunisia in combination with a high rate of consanguinity in these regions resulted in an elevated frequency of FH-homozygous patients who exhibited a more severe life-threatening phenotype than FH-heterozygous individuals [38, 39]. Another interesting example of how autosomal dominant mutation in the homozygous state could alter phenotypic expression is illustrated by a mutation in the CASR gene in another consanguineous family. The parents, who were heterozygous for the mutation, presented with familial hypocalciuric hypercalcemia whereas the proband, who harbored the same mutation in the homozygous state, had neonatal severe
hyperparathyroidism [40]. This case is an example of familial comorbidity, i.e. the co-existence of two or more genetic diseases in the same family. Seventy-five comorbid associations have been identified in Tunisian families with 41 cases of familial consanguinity noted [unpubl. data]. Intra-familial genetic heterogeneity in consanguineous families both highlights the difficulty of genetic counseling in such families and makes patient management challenging [unpubl. data].

**Consanguinity and Human Genome Structure**

New genome scan technologies, like high-density SNP microarrays and next-generation sequencing, allow the generation of genotyping data on hundreds of thousands of markers for a large number of individuals. By identifying long homozygous genomic regions called ‘runs of homozygosity’ (ROH), these methods also enable to study the extent of homozygosity within and between populations with various degrees of isolation and consanguinity [41]. If ROH are inherited from a common ancestor they are IBD and are termed autozygous. High-resolution studies of the human genome could provide insights into population history, cultural practices and genome organization, with homozygosity segment length reflecting the degree of shared parental ancestry and its age. ROH distribution across the genome is non-uniform and shows distinctive continental patterns [42]. Short ROH (tens of kb in length) are present in outbred individuals and those of intermediate size (hundreds of kb to several Mb) are seen in individuals whose parents are probably related [42]. Extremely long ROH, measuring tens of Mb, have been observed in individuals from populations with higher levels of background relatedness [42].

In an attempt to establish the genomic ancestry of North African populations, Henn et al. [43] studied the patterns of genetic variation in seven North African populations, using 730,000 SNP located across the genome. The presence of long ROH indicated that the Tunisian Berber population averaged almost twice as much of their genome in ROH than other North African populations, with a value of 230 kb versus 120 kb, respectively [43]. Analyses of autozygous genomic segments showed that a large majority of individuals in the analyzed populations were distantly related, as they presented less IBD than predicted in third cousins ($F = 0.0039$). The Tunisian Berber population nevertheless displayed an excess of pairs of individuals sharing 200–1,200 cM IBD. Surprisingly, the bimodal distribution of IBD segments between pairs of individuals of the Tunisian Berber sample was atypical, revealing that many first- and second-cousin genetic equivalent pairs were present in this sample, even though the volunteers declared themselves to be unrelated. This finding supposes that the pattern of ROH and pairwise IBD in the Tunisian Berbers probably results from endogamy due to geographical isolation or cultural marriage preferences [43]. Moreover, it emphasizes the problem of estimating inbreeding coefficients from pedigrees which ignore more ancient familial relatedness that is absent from most genealogies [41], or those originating from a population where inbreeding events have been and are both frequent and ancient [unpubl. data].

The present study is the first attempt to unravel genomic structure in terms of ROH pattern distribution and autozygous segments among an ethnic group of the Tunisian population; however, the findings cannot be extrapolated to the whole Tunisian population. Indeed, our preliminary investigation of Tunisian population samples from two urban coastal regions showed cumulative ROH ranging from 1.2–1.5 Mb [unpubl. data]. It therefore would be interesting to deepen the study of ROH genomic pattern distribution and variation among the Tunisian population using whole exome or genome sequencing, to further understand the relationships between consanguinity, founder effects, evolutionary processes and patterns of deleterious variations. An overview of the selection pressures on the genome mediated by consanguinity could highlight the ‘purging’ of deleterious mutations in the homozygous state, thus preventing the expression of corresponding genotypes in future generations. Nevertheless, the fact that some ‘healthy’ individuals are homozygous for deleterious mutations remains to be understood [44]. It has recently been reported that long ROH resulting from recent consanguinity are disproportionately enriched in more deleterious homozygous variants, which suggests that consanguinity not only increases the expression of recessive genetic diseases of significant fitness effect but also amplifies the burden of mildly deleterious homozygotes [45].

The unique arrangement of ROH in the genome, known as the autozygome in consanguineous populations, allows the use of an autozygome-guided exome strategy as an efficient means of unravelling the genetic basis of heterogeneous diseases in families from the Middle East and Tunisia [46–48]. Especially where traditional fluorescent marker linkage analysis yielded misleading
data that excluded a truly disease-causing gene [48, 49]. In our laboratory, whole exome sequencing has been used to investigate four consanguineous Tunisian families with nonsyndromic hearing loss, in whom no mutation was conventionally identified at the GJB2 locus. As the causal mutations in autosomal recessive diseases in consanguineous populations are usually homozygous, and these are generally easier to call by next-generation sequencing when the homozygosity scan filter is applied, the search for the causal variant is made easier. Indeed, after the exclusion of heterozygous SNPs and those on sex chromosomes, the putative causative substitution number was drastically reduced from 14,182 ± 1,793 to 6 ± 4, thus allowing the rapid identification of the causal mutations [unpubl. data].

**Conclusion**

Familial and geographical endogamy is widely practiced by many global populations, with the highest rates in the Arab world and in some South Asian countries [1]. Negative and positive effects of consanguinity have been identified in these consanguineous populations [18]. In Tunisia, the profile of genetic diseases in the general population displays a combination of founder effect and familial and geographical endogamy. This non-panmictic model of marital unions is a general phenomenon which highlights the fact that the Tunisian population is a meta-population of multiple subpopulations isolated at the reproductive level. Although several studies established a significant association between consanguinity and the expression of recessive genetic diseases, the health outcomes of consanguinity on other common conditions, like the chromosomal aberration Down syndrome and multifactorial non-communicable diseases, remain controversial.

In an attempt to improve the management of genetic diseases, it was suggested that patient registers be developed to provide a valuable tool for prevention and for better patient care at national and regional levels. Unfortunately, genetic services are still drastically lacking in Tunisia [6, 50, 51], and there is a lack of educational programs which could draw the attention of the general public to the potential negative health impacts of consanguinity. Morbidity and mortality due to genetic diseases could be reduced significantly by avoiding consanguineous unions. Nevertheless, rather than negatively criticizing consanguineous marriage in populations with a long-held tradition of intra-familial unions, ensuring access to preconception and premarital counseling services could provide a more acceptable way of reducing the burden of recessive diseases [52].

It seems questionable that a traditional culturally favored practice such as consanguineous marriage could lead only to negative outcomes, and indeed some authors have discussed potential biological benefits of consanguinity [53]. Assessment of the impact of consanguinity on health is difficult due to the interplay of numerous factors [5]. With the advent of the high-throughput sequencing technologies that are rapidly becoming cheaper and more widely accessible, re-thinking the design of consanguinity studies would be most welcome, and in turn lead to a better assessment of the potential beneficial effects associated with consanguinity.

**Acknowledgements**

This work was supported by the Tunisian Ministry of Public Health, the Ministry of Higher Education and Scientific Research (LR11IPT05) and by the E.C. Grant agreement No. 295097 for FP7 project Genomedika (www.genomedika.org). We are grateful to Prof. Alan Bittles and Dr. Katherine Montague for their valuable assistance that helped us to improve our manuscript.
Specific Aspects of Consanguinity in Tunisia


