Consanguinity, human evolution, and complex diseases

A. H. Bittles and M. L. Black

Centre for Comparative Genomics, Murdoch University, South Street, Perth WA 6150, Australia; and Centre for Human Genetics, Edith Cowan University, Joondalup Drive, Perth WA 6027, Australia

Edited by Diddahally R. Govindaraju, Boston University School of Medicine, Boston, MA, and accepted by the Editorial Board August 27, 2009

There is little information on inbreeding during the critical early years of human existence. However, given the small founding group sizes and restricted mate choices it seems inevitable that inbreeding would have been substantial. Currently, couples related as second cousins or closer (F ≥ 0.0156) and their progeny account for an estimated 10.4% of the global population. The highest rates of consanguineous marriage occur in north and sub-Saharan Africa, the Middle East, and west, central, and south Asia. In these regions even couples who regard themselves as unrelated may exhibit high levels of homozygosity, because marriage within clan, tribe, caste, or biraderi boundaries has been a long-established tradition. Mortality in first-cousin progeny is ≥3.5% higher than in nonconsanguineous offspring, although demographic, social, and economic factors can significantly influence the outcome. Improving socioeconomic conditions and better access to health care will impact the effects of consanguinity, with a shift from infant and childhood mortality to extended morbidity. At the same time, a range of primarily social factors, including urbanization, improved female education, and smaller family sizes indicate that the global prevalence of consanguineous unions will decline. This shift in marriage patterns will initially result in decreased homozygosity, accompanied by a reduction in the expression of recessive single-gene disorders. Although the roles of common and rare gene variants in the etiology of complex disease remain contentious, it would be expected that declining consanguinity would also be reflected in reduced prevalence of complex diseases, especially in population isolates.

Civil and Religious Regulation of Consanguineous Marriage

The roots of negative Western attitudes toward consanguinity extend back over 1,500 years. In the Eastern Roman Empire the legality of first-cousin marriage had been confirmed by the Emperor Arcadius in 400 AD (10), possibly in acceptance of the marriage regulations defined in the Old Testament Book of Leviticus 18:7–18. But according to the Venerable Bede writing in the early 8th century (11), in 597 AD Augustine the first Archbishop of Canterbury was advised by Pope Gregory I that first-cousin marriage was banned by sacred law, a somewhat overenthusiastic interpretation of Leviticus 18:6: “none of you shall approach to any that is near kin to him, to uncover their nakedness.” Depending on the translation of Bede consulted, Gregory I further advised that first-cousin unions “do not result in children” (11), an opinion that is factually incorrect (12), or that “the offspring of such marriages cannot thrive” (10), which also is at best an overstatement.

Until 1917 the Roman Catholic Church required dispensation for unions between couples related as first, second, or third cousins (equivalent to a coefficient of inbreeding, F ≥ 0.0039),...
with a wide range of reasons accepted as grounds for consanguinity dispensation, e.g., the small size of the local population, advanced bridial age, or lack of dowry (13). As a result of misunderstanding after the switch from the Roman to the Germanic system for calculating degrees of consanguinity, during the late 11th to the early 13th centuries the requirement for dispensation expanded to include fourth-, fifth-, and sixth-cousin marriages ($F \geq 0.00006$), a level of regulation that rapidly proved impractical at local level (10). Because Luther had attacked the dispensation requirements for consanguineous unions as representing the rules of the church rather than of divine intention, and as a revenue-raising device (10), after the Reformation the Protestant denominations largely accepted the Levitical marriage proscriptions with no restriction on first-cousin unions.

The Levitical guidelines also permit uncle–niece marriage ($F = 0.125$), which along with first-cousin marriages are still practiced in many Sephardi Jewish communities. Marriage regulations in Islam permit first-cousin and double first-cousin ($F = 0.125$) marriages, but uncle–niece unions are prohibited by the Quran. Contrary to common belief there is no encouragement of consanguinity within Islam, and although the Prophet Muhammad married his daughter Fatima to his ward and first cousin Ali, several hadith (sayings of the Prophet) endorse marriage between nonrelatives (14). It therefore seems that the strong preference for first-cousin marriage in most Muslim countries, principally the parallel paternal subtype, i.e., between a man and his father’s brother’s daughter, reflect both pre-Islamic Arab tradition and the rules introduced in the Quran enabling female inheritance of wealth (15).

First-cousin marriage is generally permitted within Buddhism, but the marriage regulations in Hinduism are more complex. According to the north Indian tradition believed to date back to 200 BC, pedigrees are examined over an average of seven generations on the male side and five generations on the female side to preclude a consanguineous union (16). Whereas in Dravidian south India, cross first-cousin marriage (between a man and his mother’s brother’s daughter) and more especially uncle–niece marriages are favored across all castes. Because of their customary nature, cross-cousin marriages were recognized by the government of India in the Hindu Marriage Act of 1955 and the legality of uncle–niece marriages was confirmed in the Hindu Code Bill of 1984 (17).

The Current Global Prevalence of Consanguineous Marriage
As illustrated in Fig. 1, based on detailed information accessible at the Global Consanguinity website (www.consang.net), closekin marriage continues to be preferential in many major populations, with the influence of religion apparent in the major regional differences in consanguinity prevalence across the globe (18). Despite anthropological reports indicating consanguineous marriage throughout sub-Saharan Africa, and in populous Asian countries including Bangladesh and Indonesia, little quantitative information on consanguinity is available from these regions. Nevertheless, current data indicate that some 10.4% of the 6.7 billion global population are related as second cousins or closer ($F \geq 0.0156$). Although the overall prevalence of consanguineous marriage seems to be declining, in some countries the present-day rates of consanguinity exceed those of the preceding generation, possibly reflecting greater overall survival to adulthood that in turn increases the numbers of marriageable biological relatives (19).

Large-scale emigration of people from countries where consanguinity is preferential to North America, Europe, and Oceania was an important demographic feature of the latter half of the 20th century. As previously indicated, first-cousin marriages ($F = 0.0625$) have the potential to cause legal problems for migrants and state law enforcement authorities in the United States because these unions are now either illegal or a criminal

![Consanguineous marriage (%)](img)

Fig. 1. Global distribution of marriages between couples related as second cousins or closer ($F \geq 0.0156$).
offense in 31 of 50 states (5, 6, 20), despite a unanimous recommendation in 1970 that all such state laws should be rescinded (21). In Western Europe there are at least 10 million resident migrants from regions where consanguinity is preferential, and it is the possibility that the progeny of consanguineous unions are more likely to be affected by recessive genetic disorders that has aroused greater controversy, for example, with calls by some legislators for a ban on first-cousin marriages in the United Kingdom’s Pakistani community (19, 22). Although a decline in first-cousin marriage has been observed in the Norwegian Pakistani community (23), no similar trend seems to have occurred in the United Kingdom’s Pakistani population (24) or in the Turkish or Moroccan communities in Belgium (25), and a rapid reduction in the preference for consanguineous unions by first- and second-generation migrant families in Europe appears improbable.

The Comparative Roles of Consanguinity and Endogamy in Genetic Studies

Intracommunity marriage is the norm in regions where consanguineous marriage is favored, usually contracted within long-established male lineages, e.g., within the clan (hamula) and tribe in Arab societies, within caste in India, and intrabiraderi in Pakistan. Because gene flow between communities is highly restricted in most traditional societies, adjacent villages of common consubstantial subcommunities may exhibit very different inherited disease profiles, reflecting local founder mutations and genetic drift (18). These characteristics have been demonstrated in tribe-specific single gene disorders in Saudi Arabia (26–28), the differential origins and expansion patterns of β-globin mutations in an Israeli Arab village (29), and village- and lineage-specific predisposing genes for visceral leishmaniasis in Sudan (30).

Under these circumstances and whether or not the parents are known to be consanguineous, a recessive founder or de novo mutation of chronic effect can rapidly increase in frequency within a particular community or subcommunity, resulting in the birth of an affected child. In communities with a high level of consanguineous marriage, the diagnosis of a recessive disorder in one or more members of the same family is generally indicative of a recent mutation, whereas the presence of a rare disorder in several families suggests an older mutational event or previous admixture through marriage with a person from another community (31).

Population substructure, whether caused by ethnic, geographical, religious, or social divisions, often results in variant marker allele frequencies in different subpopulations. The occurrence of type 1 errors, i.e., false positive results, is of major importance in case-control studies, association studies, and clinical trials (32, 33). Conflicting opinions have been expressed as to the impact of population stratification on genomewide studies with, for example, the claim that in the United Kingdom if persons of non-European ancestry are excluded “the extent of population stratification in the British population is generally modest” (34). Conversely, in the more homogeneous Icelandic population it was believed that population substructure had to be considered in the sampling strategy, with the implication that it would be of much greater importance in larger populations with more diverse genetic origins (35). Because genomic studies consistently report that a large majority (93–95%) of genetic variation is within-population (36), the latter opinion is unsurprising and highlights the need for vigilance in case-control studies to preclude spurious associations.

As discussed in the following sections, population stratification may also be of critical importance in the investigation of consanguinity-associated morbidity and mortality, with straightforward comparisons drawn between the progeny of first cousins versus unrelated parents of dubious validity unless both sets of parents are known to be members of the same clan, tribe, caste, or biraderi (19). For this reason, in many populations the clan or its hereditary social/occupational equivalent may be the most logical unit for genetic screening and genetic counseling programs, as exemplified by the distribution pattern of β-thalassemia in Oman where >50% of cases were diagnosed in just one of the 185 major tribes and subtribes (37, 38).

Consanguinity and Health

Within genetics, contemporary attention on consanguineous marriage continues to be largely focused on the expression and identification of rare autosomal recessive alleles, a recent example being a comparative study in Norway of progressive encephalopathy in Pakistani migrants and the indigenous population (39). But as indicated in Fig. 2, from an overall health perspective consanguinity is a much wider and more complex topic involving major social, economic, and demographic influences, differential reproductive behavior, and early- and late-onset morbidity and mortality. A thorough appreciation of the salient nongenetic variables is therefore essential in addressing the concerns of individuals, families, and communities with regard to reproductive choices, and in designing genetic education and genetic counseling programs for consanguineous couples.

The highest overall prevalence of consanguineous unions is in poor rural communities, which are typified by low levels of maternal education, early age at marriage and first birth, short birth intervals, and longer reproductive spans (15, 40–42). Each of these factors is independently associated with larger family sizes and higher rates of infant and early childhood mortality, with reproductive compensation for early losses a further complicating issue in assessing the overall health outcomes of consanguinity (12). Comprehensive genetic education and premarital genetic counseling programs can help to lessen the burden of genetic diseases in such communities, as reported in Israeli Arab and Bedouin villages (43–45). While in Middle Eastern countries such as Bahrain educational programs aimed at high school children, and through them their parents and relatives, have had a marked beneficial effect in reducing the incidence of sickle cell disease (46). There are, however, current limitations to the success of these initiatives in many low-income countries, in particular the lack of clinicians, genetic counselors, nurses, and scientific support staff with appropriate specialist training (47). Patients referred for genetic counseling may also expect directive advice as to whether or not to proceed with a pregnancy, with failure to provide an opinion interpreted as a lack of knowledge on the part of the clinician (48), and even when specific rulings have been provided by religious authorities permitting prenatal diagnosis of genetic diseases and elective termination of a pregnancy, this option may remain unacceptable to individual couples (15).

Consanguinity, Mortality, and Morbidity

To investigate the impact of consanguinity on deaths from ≈6 months gestation to an average of 10 years of age, a metaanalysis was conducted directly comparing prereproductive mortality in first-cousin versus nonconsanguineous progeny within specific populations. The study sample comprised 69 populations residing in 15 countries located across four continents, with a total sample size of 2.14 million (Table S1). An unweighted linear regression comparing mean mortality in first-cousin versus nonconsanguineous progeny in each population was plotted according to the standard equation $y = a + bx$. The results are presented in Fig. 3 as a scatter diagram and show a mean excess mortality at first-cousin level of 3.5% ($r^2 = 0.70; P < 0.00001$) that is consistent across the range of control mortalities, i.e., the level of excess consanguinity-associated mortality is independent of the basal (nonconsanguineous) death rate in each study population. The estimate of 3.5% excess deaths among first-cousin progeny compares with an earlier global estimate of 4.4%
excess mortality (49) calculated from 38 studies each of which was included in the present analysis, and it matches the 3.5% excess mortality derived for Italian data of the early to mid 20th century (13).

Initial estimates of the adverse effects of consanguineous marriage, expressed as lethal gene equivalents, had produced significantly higher values for consanguinity-associated mortality, mainly because of lack of control for the negative correlation between consanguinity and socioeconomic status (50). Although control for the effects of nongenetic variables was improved in the present study, the mean value of 3.5% excess mortality at the first-cousin level is an upper-level estimate that may be subject to further downward revision as data from better-designed studies become available.

The influence of first-cousin marriage on the prevalence of autosomal recessive single-gene disorders was examined as part of an investigation into consanguinity-associated morbidity in a Pakistani community in the United Kingdom (51). From the results of this 5-year prospective study it was calculated that there would be a ≈7/1,000 increase in autosomal recessive disorders per 0.01 increase in the mean coefficient of inbreeding (52). Thus, in a national population such as Pakistan where ≈50% of marriages were between first cousins ($F = 0.0625$) (53) some 22/1,000 extra single-gene disorders would be expected.

Unfortunately, the original study omitted control for population subdivision, which has been shown to be a notable feature of indigenous and migrant Pakistani populations (54–56), and as previously noted is typical of many more traditional populations. Wahlund effect predicts that subdivided populations characteristically exhibit higher than predicted levels of homozygosity. Given the known levels of population substructure associated with biraderi membership in Pakistan and the Pakistani community in the United Kingdom, nonconsanguineous couples are at higher risk of sharing the same recessive disease mutation than counterparts in populations where limited or no substructure exists. The consequent random consanguinity effect on the distribution and expression patterns of recessive disease genes means that in populations with significant subdivision the beneficial health outcomes that have been claimed through simply avoiding consanguineous marriage are almost certainly exaggerated and require reassessment (19, 57).

**Consanguinity and Complex Diseases**

There has been extended debate on the nature of the genetic contribution to complex diseases, i.e., whether the common disease/common variant or the common disease/rare variant hypothesis is more applicable (58), with the role of copy number...
variants also proposed (59, 60). Consanguinity would be expected to exert a greater influence on the etiology of complex diseases if rare autosomal recessive alleles were causally implicated, whereas if disease alleles that are common in the gene pool are involved then intrafamilial marriage would have a proportionately lesser effect. However, because both gene–gene interactions and numerous nongenetic factors in prenatal and postnatal life also contribute to the disease phenotype, a single all-embracing solution to the genetics of complex diseases is highly improbable.

Major genomewide analyses of diseases with onset primarily in childhood and adulthood have identified associations with specific chromosomal regions, e.g., for type 1 and type 2 diabetes (61, 62), although these studies have emphasized the large numbers of genes involved and the small increased risk that appears to be associated with most individual variants. Concern also has been expressed that concentration on the identification of gene variants via patients with the disease under study rather than full genome sequencing of randomly ascertained samples could lead to significantly inflated rates of false positives (63).

Investigations into the effects of consanguinity on congenital defects have produced quite varied results, in large part because of a lack of standardized assessment protocols and the different environmental and socioeconomic circumstances of the study populations. Using nonconsanguineous progeny as controls, estimates of the excess level of congenital defects in first-cousin offspring have ranged from 0.7% to 7.5% (64–68), but the Latin American Collaborative Study of Congenital Malformation based on 34,1902 newborns found a significant association with consanguinity only for hydrocephalus, postaxial polydactyly, and bilateral oral and facial clefts (69).

A different picture emerges from the large literature on congenital heart defects, which are conservatively estimated to have an incidence of 50/1,000 live births (70). Although a consistent positive association between consanguinity and disorders such as ventricular septal defect and atrial septal defect has been demonstrated, indicating the involvement of common variants, both positive and negative associations with patent ducus arteriosus, atrioventricular septal defect, pulmonary atresia, and tetralogy of Fallot have been reported in different populations (71–74), suggestive of community-specific founder mutations. Moreover, also possible that nonstandardized diagnostic protocols may have contributed to the variant findings reported by different study centers.

As yet relationships between consanguinity and complex diseases of adulthood have been significantly underinvestigated, and the few studies published have relied mainly on rudimentary sampling strategies, with simple consanguineous versus nonconsanguineous comparisons in disease prevalence and inadequate attention paid to possible genetic or demographic subdivisions. Accordingly, the results obtained often are contradictory, e.g., with both positive and negative associations reported between consanguinity and breast cancer (75–77), and consanguinity and heart disease (75, 78, 79). Long-term studies conducted on the Dalmatian islands in the Adriatic Sea have indicated a positive association between inbreeding and a very wide range of common adulthood disorders, including hypertension, coronary heart disease, stroke, cancer, uni/bipolar depression, asthma, gout, peptic ulcer, and osteoporosis (80–82). The data thus suggest virtually ubiquitous causal involvement of rare autosomal recessive genes in adult-onset disease in this population, with the more general corollary that increasing genomewide heterozygosity after a decline in consanguineous marriage should lead to a widespread reduction in the burden of common genetic diseases (83).

The Dalmatian studies have the very considerable advantage of demographically well-characterized populations with known ethnic origins, although the actual definitions used in assessing the comparative levels of inbreeding are genetically quite imprecise and principally reflect village endogamy rather than consanguinity per se. As previously discussed, until the early 20th century church dispensation would have been required for marriages between spouses related as third cousins or closer (F ≥ 0.0039) in these devoutly Roman Catholic communities (13). In the absence of church records indicating dispensation for marriages contracted within the prevailing consanguinity regulations, the consanguineous relationships examined may principally have been random rather than preferential in nature and reflected restricted marriage partner choices. The analysis of genealogical data covering four to five generations showed substantial levels of consanguinity in some communities, with mean coefficients of inbreeding ranging from α = 0.002 to 0.049 calculated at village level, indicating major variations in local marriage patterns driven by both the history and the geographical location of each settlement (80).

Pedigree-based estimates of consanguinity and the resultant levels of homozygosity have several limitations; in particular, they do not provide information on close-kin marriages that have occurred in distant generations and thus underestimate cumulative inbreeding effects, and with rare exceptions incorrectly ascribed paternity is not recorded. To complement the pedigree-based approaches previously adopted and avoid these difficulties, high-density genome scans were used to estimate individual autozygosity (Froh) from uninterrupted runs of homozygosity (ROH). An appropriate length threshold was empirically derived for ROH and the method was applied to data derived from residents of the Dalmatian islands, the Orkney islands off the north coast of Scotland, mainland Scotland, and the state of Utah (84). Initial comparisons of Froh values ranging from 0.5, i.e., with a minimum length threshold of 0.5 Mb, to 1 (length threshold 1 Mb) and 5 (length threshold 5Mb) with pedigree data from the Orkneys indicated good correlation with pedigree-based mean coefficients of inbreeding and so confirmed the applicability of the method for the direct assessment of autozygosity. The method has been further applied to investigate changes in autozygosity through time in two American study populations. The steady decreases observed in the size and frequency of ROH > 1 Mb in length in these populations were ascribed to expanded marriage pools and larger effective population sizes and indicate that diagnostic protocols may have contributed to the variant findings reported by different study centers.

When applied to behavioral disorders genomewide analysis has indicated the potential contribution of thousands of alleles of very small effect in schizophrenia and bipolar disorder, with significant genetic overlap between the two disease states (86, 87). At the same time, homozygosity mapping in autism (88) and a case-control study of bipolar disorder type 1 in consanguineous progeny (89) both implicated the causal expression of rare recessive genes. ROH similarly have been shown to be significantly more common in patients with schizophrenia spectrum disorders, suggesting the involvement of recessive alleles in the etiology of the disorder (90). Reverting to earlier comments on the relationship between endogamy and consanguinity, an association between consanguinity and Alzheimer disease was demonstrated in a genealogical study of the Saguenay region in Quebec (91), and multiple loci for Alzheimer disease were identified in a highly endogamous and consanguineous Israeli Arab kindred (92), in both cases indicative of founder mutations. Thus, from a more general perspective these results strengthen the argument that all association studies on complex diseases would benefit from a sound prior knowledge of community demographic and genetic structure.

Discussion

Although consanguinity is a highly complex and multifaceted topic (Fig. 2), the claimed social and cultural advantages, such as
Bittles and Black

Contrasting biological and social outcomes of consanguineous marriage (Fig. 4). Therefore, in conjunction with increasing genetic diversity and restricted allelic heterogeneity are generally expected in isolated founder populations, it also is salutary that a genomewide association analysis of obesity and other metabolic disorders in a Pacific island community, in which reduced haplotype diversity and extended linkage disequilibrium had already been demonstrated, failed to detect major contributory alleles and instead indicated the presence of common variants of small effect (108, 109).

Having largely been ignored for many years, the specific roles of population bottlenecks and consanguinity in influencing variation between and within populations are now receiving due attention, with special focus on homozygosity in identifying recent common ancestry via ROH analysis (110). The potential complexity of the interrelationships between consanguinity and human health and disease was highlighted by the reported association between consanguinity and predisposition to major infectious diseases (111). If these findings are substantiated, by ameliorating the risk of exposure to infectious agents a global decline in consanguinity could also providentially reduce the risk of inflammatory disease and hence the development of coronary disease in middle and old age (112).

Time will tell whether these as yet tenuous epidemiological connections can be sustained. In the interim, it is important to emphasize that in assessing the impact of consanguinity on any aspect of health a clear causal relationship needs to be established, rather than reliance on speculation driven solely by the presence of a close kin union in the family pedigree. At the same time, rigorous control for population stratification should be a prerequisite in the many populations where community substructures exist if confused and confusing conclusions are to be avoided.

ACKNOWLEDGMENTS. We thank the referees for constructive comments. A.H.B. is supported by National Science Foundation Grant 0527751.

Bittles and Black

PNAS | January 26, 2010 | vol. 107 | suppl. 1 | 1785
98. Shaw A, Hurst JA (2009) “I don’t see any point in telling them”: Attitudes to sharing genetic information in the family and carrier testing of relatives among British Pakistani adults referred to a genetics clinic. Ethn Health 14:205–224.