Published in final edited form as:

Trends Genet. 2013 July ; 29(7): 412–418. doi:10.1016/j.tig.2013.01.007.

Amish Revisited: Next Generation Sequencing Studies of Psychiatric Disorders Among the Plain People

Liping Hou¹, Gloria Faraci¹, David T.W. Chen¹, Layla Kassem¹, Thomas G. Schulze², Yin Yao Shugart³, and Francis J. McMahon¹

¹Human Genetics Branch, National Institute of Mental Health Intramural Research Program, National Institutes of Health, US Dept of Health and Human Services, Bethesda, MD USA

²Dept of Psychiatry and Psychotherapy, University of Göttingen, Göttingen, Germany

³Unit on Statistical Genomics, National Institute of Mental Health Intramural Research Program, National Institutes of Health, US Dept of Health and Human Services, Bethesda, MD USA

Abstract

The rapid development of next-generation sequencing (NGS) technology has led to renewed interest in the potential contribution of rarer forms of genetic variation to complex, non-Mendelian phenotypes, such as psychiatric illnesses. Although challenging, family-based studies offer some advantages, especially in communities with large families and a limited number of founders. Here we revisit family-based studies of mental illnesses in traditional Amish and Mennonite communities -- known collectively as the Plain people. We discuss the new opportunities for NGS in these populations, with a particular emphasis on investigating psychiatric disorders. We also address some of the challenges facing NGS-based studies of complex phenotypes in founder populations.

Keywords

bipolar disorder; identity by descent; homozygosity; Mennonite; Anabaptist

Genome sequencing in distant relatives

The increasing power and decreasing cost of genome sequencing have fueled widespread interest in NGS technology to discover rare alleles that contribute to risks for human illnesses. NGS studies have already identified many rare variants involved in single-gene mendelian disorders 1, 2, 3 and 4 as well as alleles involved in rare and genetically-heterogeneous conditions such as brain malformations [5]. The discovery of rare copy-number variants conferring high risk for disorders such as autism and schizophrenia [6] has also led to a growing recognition of the importance of rare forms of genetic variation in psychiatric illnesses. Variants with rare or moderate allele frequency may confer greater risk for disease than common variants because high-risk alleles that contribute to early-onset diseases may reduce reproductive fitness and be driven to low frequencies by purifying selection [7]. High-risk alleles can be very informative because they implicate genes that may play a more important etiologic role than low-risk variants. Such genes provide excellent targets for functional genetic studies and new therapies. Although each rare high-risk allele can account for little heritability, sets of high-risk alleles could collectively

Corresponding author: McMahon, F.J. (mcmahonf@mail.nih.gov), MD, Chief, Human Genetics Branch and Genetic Basis of Mood and Anxiety Disorders Section, National Institute of Mental Health Intramural Research Program, 35 Convent Drive, Rm 1A202, Bethesda, MD 20892-3719 tel 301-451-4455 / fax 301-402-9081.

account for substantial fractions of a common disease [6] and may act on a polygenic background to modulate clinical features.

As high-throughput sequencing rapidly advances, whole-genome sequencing (WGS) is growing in popularity relative to exome sequencing which captures only the transcribed ~1% of the genome. WGS requires no capture procedures and provides information on both transcribed and intergenic regions, where much regulatory variation resides. However, WGS is still two- to threefold more expensive than exome sequencing, and it requires more sophisticated quality control, data management, and bioinformatics. Much of the variation identified by WGS is difficult to interpret - but this is changing. For example, the Encyclopedia of DNA Elements (ENCODE) project has begun to shed some light on the functional consequences of genetic variation outside coding regions [8].

Extended families are a valuable asset for genome sequencing studies. Because the same chromosomal segments recur among relatives, otherwise rare alleles can be observed repeatedly in multiple individuals, reducing false positive findings due to sequencing errors that can be difficult to identify in singletons. Statistical imputation can be used to infer sequence information for relatives who are not sequenced but genotyped with inexpensive single-nucleotide polymorphism (SNP) arrays [9], reducing sequencing costs. Families may also provide a genetically more homogeneous set of individuals from whom shared variants conferring risk for disease may be identified [10]. This approach (Figure 1) complements the traditional case control design that is more often used in sequencing studies and has been so valuable for genome-wide association studies. Extended families that are geographically stable are becoming less common in developed countries, prompting a renewed interest in traditional societies that still produce large families who live together. These isolated founder populations provide a unique opportunity to discover important alleles involved in genetically complex phenotypes such as mental illnesses.

Brief history of the Plain people and their origins

The Plain people of North America are one such population. The Plain people collectively encompass several groups descended from the Alsace region, where Switzerland, France, and Germany converge. Persecuted for their rejection of Catholicism and the Protestant Reformation, they migrated to North America, starting in the early 1700s. A relatively small group of founders settled initially in Pennsylvania, where they enjoyed religious tolerance and liberty, and have now spread all over the continent. Endogamy is the rule, and very few people join the community from the outside, leading to high rates of consanguinity and an increased incidence of several autosomal recessive disorders [11].

The largest and best-known Plain people are the Old Order Amish and 'Horse-and-Bugg' Mennonites. They typically reject automobiles, public utilities, and all forms of electronic entertainment in favor of a low-technology agrarian lifestyle focused on family and community. The largest Plain populations now live in Holmes County (Ohio), Lancaster County (Pennsylvania), and northern Indiana (Figure 2). With large families, Plain populations are growing and spreading across North America. Amish communities are organized into church districts comprising around 20–30 families. Marriage within the church district is common. Baptism occurs in early adulthood and is voluntary, but ~90% of Amish young people stay in the community of their birth. Strong community ties make possible a system of mutual support and assistance in times of need [12, 13 and 14], reducing stress and social isolation, but also narrowing the range of acceptable expressions of individuality. This creates a unique backdrop in which to study psychiatric disorders and their causes.

Genetic studies in the Plain people

In 1964, Victor McKusick (Figure 3) and colleagues published a now-famous paper on medical genetics in the Amish [15]. They noted several reasons why the Old Order Amish would be a ‘useful’ population for genetic studies:

1. “The Old Order Amish constitute a defined population...There can be no question of who is presently Amish...”
2. “It is a closed population. The Amish do not proselytize and marrying outside the faith is forbidden. Although some leave the group, very few ‘new’ persons enter it [who are] easily identified by their name...”
3. “The origins of the populations are western European and are rather well known.”
4. “Genealogic records are extensive. Most Amish can trace their complete ancestry back to the immigrants from Europe.”

Owing in part to these advantages and in part to Dr McKusick’s personal relationship with the Amish, a large list of mendelian conditions have been described and many have been solved genetically. A quick search of Online Mendelian Inheritance in Man reveals over 50 examples. These include diseases rarely seen outside the Amish community, such as cartilage-hair hypoplasia (#250250), Ellis Van Creveld syndrome (#225500), and several inborn errors of metabolism, as well as more widespread diseases such as ataxia-telangiectasia (#208900), progressive familial intrahepatic cholestasis (#211600), and several types of muscular dystrophy. This work continues at centers such as The Clinic for Special Children in Strasburg, PA, where treatments for some of these inherited disorders are beginning to be developed [16, 17 and 18].

Progress in the study of non-mendelian disorders has been more limited, mirroring the slower progress in other populations for these more challenging disorders. One clear success is the identification of a common allele in the gene encoding a cytochrome P450 enzyme (CYP2C19) that has an impact upon response to anti-platelet therapy in Amish patients with heart disease [19]. Importantly, this gene may also have relevance for anti-platelet therapy in other populations.

The culture and lifestyle of the Amish and other Plain peoples can also create difficulties for research. Most Old Order Amish communities eschew electronic forms of communication, requiring researchers to find and evaluate potential volunteers in the field. The total Amish population is about 275 000, which limits the number of potential volunteers, even with relatively efficient case-finding methods. German and Swiss dialects are spoken at home (although English is taught at school) and formal schooling usually ends after the 8th grade. This creates educational and language barriers that can complicate informed consent, volunteer enrollment, and evaluation, especially for psychiatric disorders. ‘Horse-and-Buggy’ and other Old Order Mennonites are less genetically homogeneous and more mobile than the Amish, but they share a common ancestry and can be easier to reach and diagnose, adding to the pool of potential study participants. Other Plain populations, such as the Hutterites, are small and typically live in remote areas, limiting access.

Amish and Mennonite population structure

The history of Amish and Mennonite settlements in the US is well known. Extensive genealogical records have been kept by the Plain people [20], and an electronic genealogical database is curated at the NIH [21]. However, few studies have examined the precise genetic relationships between the Amish in Lancaster County and those who arrived later and settled in the mid-West or elsewhere in North America. One thing is clear, however: each of the

Plain populations is based on a small number of founding couples, who themselves represented a fraction of the ancestral population in Europe.

These facts account for two of the major population genetic forces that have shaped the Plain populations: founder effects and genetic drift. The founder effect is the concentration of unique, private, or otherwise low-frequency alleles in isolated populations whose ancestry goes back to a few founding individuals, who presumably carried those same alleles. Founder effects and endogamy explain much of the increased autosomal recessive disease burden long observed among the Plain people. Genetic drift refers to the random fluctuations in allele frequencies that occur in relatively small, isolated populations. Genetic drift alone can create large differences in allele frequencies, complicating the identification of disease alleles in founder populations. Footprints of these same population dynamics probably exist in everyone [22], albeit obscured by mixing between populations.

One unique advantage of inbred populations for disease gene mapping is the phenomenon of autozygosity. Inbreeding can increase the incidence of rare recessive disorders. Each isolated population has its own genetic load that depends on which alleles were present in the founders or arose through new mutations. Rare alleles that act recessively to influence disease risk tend to reside within longer stretches of homozygosity that can be easily mapped using standard SNP arrays. This strategy has been used successfully to identify causal alleles for several autosomal recessive disorders prevalent in Plain populations [23]. Although common psychiatric disorders do not exhibit obvious recessive inheritance, recessive alleles may play a role as components of the risk. For example, excess runs of homozygosity have been reported in the mentally-ill offspring of first-cousins in Japan [24].

Family-based sequencing in Plain kindreds

The Amish and other Plain people are well suited for research approaches based on large kindreds. Large families increase the opportunities for ascertaining extended relatives with the same disorder. Because most Plain people marry within their community, even individuals belonging to different nuclear families are often related, forming an extended kindred ideal for genetic studies. Furthermore, because the Plain communities were each founded by relatively few original settlers, genetic diversity is reduced, which may mean smaller sets of risk alleles involved in common disorders. Individuals also tend to stay in the area where they were born, facilitating longitudinal and offspring studies 25 and 26.

Although their large sizes offer many advantages for genetic research, Plain kindreds also tend to be very complex. Individuals are usually related in several ways, creating so-called ‘inbreeding loops’ that complicate traditional segregation and linkage analyses. Dense SNP arrays enable alternative approaches that rely on detecting long chromosomal segments inherited from a common ancestor. These identity-by-descent (IBD) methods are discussed next.

IBD and shared segment analysis

One of the big challenges for studies based on large-scale sequencing is the enormous number of variants that are present. Each individual is likely to carry many thousands of variants. Of these, some 10 000 are predicted to be potentially functional based on bioinformatic analysis. Because common diseases are thought to be genetically heterogeneous, many genes are potentially involved, and narrowing down the list of candidate variants to a few ‘causal’ mutations is challenging.

Under many genetic models, causal alleles can be expected to reside within regions of IBD shared by affected relatives. This is because mutations are passed on from previous

generations together with a flanking stretch of DNA that gradually diminishes from generation to generation due to recombination. IBD regions represent chromosomal segments that are inherited without recombination from a common ancestor, and therefore all of the variants within IBD regions are expected to be in strong linkage disequilibrium. In distant relatives, IBD regions constitute a small portion of the genome (Figure 4), and can thus greatly reduce the search space for causal alleles [10]. This approach has been used to map both monogenic and complex disorders in the Plain people [27] and in other founder populations 28 and 29. In complex disorders, not all distantly related cases may share the same causal alleles, but regions of IBD that are shared more often by pairs of distantly related cases are good places to look for causal alleles. Several studies have shown that IBD analysis might be sufficiently robust to detect loci involved in genetically heterogeneous traits where standard genetic linkage analysis has been impossible or ambiguous 30 and 31.

Statistical approaches to identifying IBD are rapidly evolving ([32] for recent review). It is generally argued in the literature that IBD detection is robust to inbreeding and actually benefits from the long stretches of IBD observed in founder populations [33]. Statistical issues also arise with regard to testing for association of rare variants with phenotypes in families. Traditional family-based association tests are limited by family size, and typically assume that all affected relatives share the same causal allele. Newer methods are under development that can handle more extended pedigree structures [34] and collapse variants in the same gene[35], thus enabling burden tests that are more powerful for rare allele studies. Such methods generally require sequencing of both affected and unaffected individuals.

Several important questions require further study. How closely related should people be when considering IBD sharing? Are close relatives better than distant relatives in some situations? What about clusters of more closely related individuals who share both genes and environment? These very interesting and important questions currently lack a precise answer. In our preliminary data we have found that IBD sharing among the Amish can be reliably measured with SNP arrays in relatives as distant as third cousins, after which it becomes difficult to distinguish from noise (Hou, L. *et al.*, unpublished).

NGS can help fine-map regions of IBD sharing in several ways. First, because it gives complete information on variable sites, NGS can better define the boundaries of shared segments. Second, because it detects alternative alleles that can be annotated for function, NGS helps narrow the focus to the variants most likely to play a causal role in the disease. Third, by allowing nucleotide-level comparisons between segments shared IBD by distant relatives, NGS can help to identify variants that have arisen *de novo* in recent generations and may contribute to disease penetrance in individuals who carry risk haplotypes (similar to the ‘Clan Genomics’ model discussed in [36]).

Psychiatric illness among the Amish and other Plain people

We have thus far enumerated several advantages of applying NGS methods to Plain populations, including (i) a homogenous environment, which increases the proportion of phenotypic variance due to genes, (ii) long stretches of linkage disequilibrium, which facilitates the detection of chromosomal segments shared IBD, (iii) enrichment for recessive alleles, and (iv) presumably reduced heterogeneity of causative alleles owing to bottlenecks and founder effects. Although these characteristics make the Plain people a useful population for many complex human diseases, we argue that there are additional features that make this population promising for studying mental illnesses.

For example, the generally stable and predictable rural lifestyle, together with low rates of substance abuse, reduces non-inherited influences on thoughts and behavior that can

complicate diagnosis. Family informants are abundant, available, and cooperative, often providing crucial diagnostic information. Inpatient care is usually provided by a relatively small network of psychiatric hospitals, facilitating focused ascertainment. Furthermore, because of the close-knit community structure, aberrant behavior rarely goes undetected.

There is a long history of bipolar disorder studies among the Amish in Lancaster County, PA 37 and 38. The genetic component of these studies generally used traditional linkage methods 39, 40 and 41. Although important in their time, these studies encountered many of same problems as other linkage studies of complex phenotypes, such as small sample size, low marker density across the genome, and lack of mature statistical tools to handle linkage analysis in pedigrees with very complex structures. Some intriguing recent studies have taken fresh approaches. For example, copy-number variants that may have contributed to illness in a large multiplex Old Order Amish pedigree were recently examined [29]. Another study examined circadian rhythms - known to be disturbed in individuals with bipolar disorder - in cultured cells derived from Amish cases and matched controls [30]. Sample size remains a limiting factor in all of these studies, however, and replication will be critical.

The presentation of major psychiatric disorders among the Amish seems to be generally similar to that in other North American populations. The published literature is dominated by Egeland's pioneering work on mood disorders among the Old Order Amish in Lancaster County in the late 1900's (Figure 5). Despite an over-representation of males among identified cases (bipolar disorder usually shows no sex differences in prevalence), these studies found that clinical symptoms, age at onset, and illness course were all broadly consistent with mood disorders within non-Amish populations. Our preliminary observations within the mid-Western Amish in Ohio and Indiana generally agree with this literature, although we observe lower rates of psychosis and comorbid substance-use disorders among the Amish than in other populations (Kassem, L. *et al.*, unpublished data).

Studies of psychiatric disorders among the Plain people also face several challenges. Large epidemiologic studies of psychiatric disorders have not, to our knowledge, been performed in this population, and therefore the true prevalence and geographic distribution of disorders are unknown. Diagnostic instruments and criteria have generally been validated in people drawn from modern industrialized societies, where thoughts and emotions are often expressed in different ways. Intensive social supports among the Plain people may ameliorate impairment and disability, but also obscure typical yardsticks of impairment and incapacity. For example, bipolar disorder (especially the episodes of elevated mood and overactivity known as mania), can be more difficult to diagnose where access to credit cards, alcohol, and automobiles is limited, protecting manic patients from some of the most serious consequences of their behavior.

Concluding remarks

The Amish and other Plain people have long contributed to our understanding of rare genetic disorders. With the advent of NGS we are now able to test the proposition that those populations may also have much to teach us about common non-mendelian disorders, and in particular psychiatric disorders. Although too few suitable cases may exist for powerful studies of very rare risk alleles, a problem which applies mainly to uncommon illnesses, this difficulty should fade over time as the Plain population continues to expand rapidly, and new cases arise in each generation. Other challenges include causal alleles that may be too rare or psychiatric disorders that are too heterogeneous to enable true discoveries in existing sets of cases. The best way to approach this problem is to increase sample size. The relative homogeneity of Plain populations facilitates increasing sample size without a proportional increase in heterogeneity, but genetic differences between geographically separated

populations need to be considered [10]. In addition, not all distant relatives may share causal alleles. In that case, researchers can return to the original probands and collect more closely related cases, or adapt one of the collapsing methods that consider the burden of deleterious alleles within a gene or group of genes [42], with appropriate corrections for relatedness. Another potential pitfall is that candidate variants may be unique to the Plain population and may not be seen again in outbred case control samples. In that situation, it may still be valuable to resequence in other people the same genes that contain rare variants associated with disease in the Plain population [43]. Even though the same rare variants will probably not be present in outbred populations, the variants identified in the Plain populations may mark a gene involved in the disease. If so, then study of that gene in outbred populations with the same disease should uncover additional causative variants. Finally, the list of candidate variants may be too long, even after filtering and bioinformatic annotation. This can be addressed by increasing the stringency of bioinformatics filters, considering convergent data from other kinds of studies (e.g., differential gene expression), and expanding the sample to include more distantly related individuals. Of course, high-risk alleles for common psychiatric disorders may simply not exist. In that case, the data collected from the Plain population will still be valuable for studying the phenotypic effects of common low-risk alleles in extended kindreds.

Most of these pitfalls apply equally to any population in which common complex diseases are studied. The unique history, genetic structure, and relative accessibility of the Plain people for medical research present grounds for optimism that significant progress can be made. Psychiatric genetic studies may particularly benefit from the homogenous environment, long stretches of IBD, enrichment for recessive alleles, reduced genetic heterogeneity, close-knit families, and lower prevalence of substance abuse among the Plain people. Innovation in human genetics is often driven by new technologies, such as NGS. The application of these new technologies to well-studied populations with a record of supporting progress in research may be the next innovative step that advances this challenging field.

Acknowledgments

Supported by the NIMH Intramural Research Program. We thank Daniel Weaver (Holmes County, OH) and Dr Richard Stevick (Messiah College, Lebanon, PA) for critical review of the manuscript.

References

1. Lupski JR, et al. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med*. 2010; 362:1181–1191. [PubMed: 20220177]
2. Musunuru K, et al. Exome sequencing, ANGPTL3 mutations, and familial combined hypolipidemia. *N Engl J Med*. 2010; 363:2220–2227. [PubMed: 20942659]
3. Ng SB, et al. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nature genetics*. 2010; 42:790–793. [PubMed: 20711175]
4. Sobreira NL, et al. Whole-genome sequencing of a single proband together with linkage analysis identifies a Mendelian disease gene. *PLoS Genet*. 2010; 6:e1000991. [PubMed: 20577567]
5. Bilguvar K, et al. Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations. *Nature*. 2010; 467:207–210. [PubMed: 20729831]
6. Sebat J, et al. Rare structural variants in schizophrenia: one disorder, multiple mutations; one mutation, multiple disorders. *Trends Genet*. 2009; 25:528–535. [PubMed: 19883952]
7. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? *American journal of human genetics*. 2001; 69:124–137. [PubMed: 11404818]
8. Consortium EP, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012; 489:57–74. [PubMed: 22955616]

9. Li Y, et al. Genotype imputation. *Annual review of genomics and human genetics*. 2009; 10:387–406.
10. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet*. 2010; 11:415–425. [PubMed: 20479773]
11. Strauss KA, Puffenberger EG. Genetics, medicine, and the Plain people. *Annual review of genomics and human genetics*. 2009; 10:513–536.
12. Kraybill, DB. *The Riddle of Amish Culture*. The Johns Hopkins University Press; 2001.
13. Hostetler, JA. *Amish Society*. The Johns Hopkins University Press; 1993.
14. Hurst, CE.; McConnell, DL. *An Amish Paradox: Diversity and Change in the World's Largest Amish Community*. The Johns Hopkins University Press; 2010.
15. McKusick VA, et al. Genetic Studies of the Amish, Background and Potentialities. *Bull Johns Hopkins Hosp*. 1964; 115:203–222. [PubMed: 14209042]
16. Morton DH, et al. Diagnosis and treatment of maple syrup disease: a study of 36 patients. *Pediatrics*. 2002; 109:999–1008. [PubMed: 12042535]
17. Strauss KA, et al. Management of hyperbilirubinemia and prevention of kernicterus in 20 patients with Crigler-Najjar disease. *Eur J Pediatr*. 2006; 165:306–319. [PubMed: 16435131]
18. Beck SJ, et al. A controlled trial of working memory training for children and adolescents with ADHD. *J Clin Child Adolesc Psychol*. 2010; 39:825–836. [PubMed: 21058129]
19. Shuldiner AR, et al. Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. *JAMA*. 2009; 302:849–857. [PubMed: 19706858]
20. Fisher, AL. *Fisher Family History: Descendants and History of Christian Fisher*. Eby's Quality Printing; PA: 1988. p. 1757–1838.
21. Lee WJ, et al. PedHunter 2.0 and its usage to characterize the founder structure of the Old Order Amish of Lancaster County. *BMC Med Genet*. 2010; 11:68. [PubMed: 20433770]
22. Schaffner SF. The X chromosome in population genetics. *Nat Rev Genet*. 2004; 5:43–51. [PubMed: 14708015]
23. Puffenberger EG, et al. Genetic mapping and exome sequencing identify variants associated with five novel diseases. *PLoS One*. 2012; 7:e28936. [PubMed: 22279524]
24. Kurotaki N, et al. Identification of novel schizophrenia loci by homozygosity mapping using DNA microarray analysis. *PLoS One*. 2011; 6:e20589. [PubMed: 21655227]
25. Shaw JA, et al. A 10-year prospective study of prodromal patterns for bipolar disorder among Amish youth. *J Am Acad Child Adolesc Psychiatry*. 2005; 44:1104–1111. [PubMed: 16239857]
26. Egeland JA, et al. A 16-Year Prospective Study of Prodromal Features Prior to BPI Onset in Well Amish Children. *J Affect Disord*. 2012; 142:186–192. [PubMed: 22771141]
27. Brooks P, et al. Robust physical methods that enrich genomic regions identical by descent for linkage studies: confirmation of a locus for osteogenesis imperfecta. *BMC genetics*. 2009; 10:16. [PubMed: 19331686]
28. Houwen RH, et al. Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. *Nature genetics*. 1994; 8:380–386. [PubMed: 7894490]
29. Morrow EM, et al. Identifying autism loci and genes by tracing recent shared ancestry. *Science*. 2008; 321:218–223. [PubMed: 18621663]
30. Ercan-Sencicek AG, et al. L-histidine decarboxylase and Tourette's syndrome. *N Engl J Med*. 2010; 362:1901–1908. [PubMed: 20445167]
31. Krawitz PM, et al. Identity-by-descent filtering of exome sequence data identifies PIGV mutations in hyperphosphatasia mental retardation syndrome. *Nature genetics*. 2010; 42:827–829. [PubMed: 20802478]
32. Browning SR, Browning BL. Identity by descent between distant relatives: detection and applications. *Annual review of genetics*. 2012; 46:617–633.
33. Te Meerman GJ, et al. Perspectives of identity by descent (IBD) mapping in founder populations. *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*. 1995; 25(Suppl 2):97–102. [PubMed: 8590355]

34. Thornton T, McPeck MS. ROADTRIPS: case-control association testing with partially or completely unknown population and pedigree structure. *American journal of human genetics*. 2010; 86:172–184. [PubMed: 20137780]
35. Shugart YY, et al. Weighted pedigree-based statistics for testing the association of rare variants. *BMC Genomics*. 2012; 13:667. [PubMed: 23176082]
36. Lupski JR, et al. Clan genomics and the complex architecture of human disease. *Cell*. 2011; 147:32–43. [PubMed: 21962505]
37. Egeland JA, Hostetter AM. Amish Study, I: Affective disorders among the Amish, 1976–1980. *Am J Psychiatry*. 1983; 140:56–61. [PubMed: 6847986]
38. Egeland JA, et al. Description of Amish Study data set. *Genet Epidemiol*. 1989; 6:195–199. [PubMed: 2786484]
39. Egeland JA, et al. Bipolar affective disorders linked to DNA markers on chromosome 11. *Nature*. 1987; 325:783–787. [PubMed: 2881209]
40. Kelsoe JR, et al. Re-evaluation of the linkage relationship between chromosome 11p loci and the gene for bipolar affective disorder in the Old Order Amish. *Nature*. 1989; 342:238–243. [PubMed: 2682265]
41. Ginns EI, et al. A genome-wide search for chromosomal loci linked to mental health wellness in relatives at high risk for bipolar affective disorder among the Old Order Amish. *Proc Natl Acad Sci U S A*. 1998; 95:15531–15536. [PubMed: 9861003]
42. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *American journal of human genetics*. 2008; 83:311–321. [PubMed: 18691683]
43. Liu DJ, Leal SM. Replication strategies for rare variant complex trait association studies via next-generation sequencing. *American journal of human genetics*. 2010; 87:790–801. [PubMed: 21129725]

Figure 1.

One Family-based Sequencing Strategy (adapted from Cirulli & Goldstein [10] with permission). Affected individuals are shaded. Red stars represent the causal variant, stars of other colors represent variants that are shared by the sequenced individuals but do not segregate with disease in the family. The most distantly related cases are sequenced first, with follow-up genotyping of candidate variants in additional relatives. Genes harboring segregating variants are then sequenced in new families and case-control cohorts.

Figure 2.

Amish settlements within the continental US. An Amish settlement is defined by three or more Amish households or two or more Amish households with a known bishop, minister, or deacon (source: <http://digitalunion.osu.edu/r2/summer07/eellis/Mapshome.html>).

Figure 3.
Victor McKusick (1921–2008). ‘Father of Medical Genetics’ and pioneer of genetic studies in the Amish (Johns Hopkins Chesney Medical Archives).

Figure 4. Identity by Descent. Colored bars represent founder chromosomes. Yellow-boxed segments indicate regions of the genome that descend from the common ancestors and are shared by some individuals in the current generation.

Figure 5.
Janice Egeland, who pioneered many studies of psychiatric disorders among the Lancaster County Amish.