Genomics of immune diseases and new therapies

Michael Lenardo, **Bernice Lo**, and **Carrie L. Lucas**

NIAID Clinical Genomics Program, Molecular Development of the Immune System Section, Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA

Abstract

Genomic DNA sequencing technologies have been one of the great advances of the 21st century – decreasing in cost by 7 orders of magnitude and opening up new fields of investigation throughout research and clinical medicine. Genomics coupled with biochemical investigation has allowed the molecular definition of a growing number of new genetic diseases that reveal new concepts of immune regulation. Also, defining the genetic pathogenesis of these diseases has led to improved diagnosis, prognosis, genetic counseling, and, most importantly, new therapies. We highlight the investigational journey from patient phenotype to treatment using the newly defined XMEN disease caused by the genetic loss of the MAGT1 magnesium transporter. This disease illustrates how genomics yields new fundamental immunoregulatory insights as well as how research genomics is integrated into clinical immunology. At the end, we discuss two other recently described diseases: PASLI (PI3K dysregulation) and CHAI/LATAIE (CTLA-4 deficiency) that show journeys from unknown immunological diseases to new precision medicine treatments using genomics.

"In almost all things, what they contain of useful or applicable nature is hardly perceived unless we are deprived of them, or they become deranged in some way."

William Harvey (1657)

Introduction

Medical genetics became a medical subspecialty in the 1950's with the hope that understanding the genetic basis of disease would accelerate diagnosis, prognosis, and treatments (1). The first genetic (primary) immunodeficiency, Bruton's agammaglobulinemia, together with the first successful etiological therapy, intravenous immunoglobulin replacement, were defined in 1952 (2). However, four decades passed before the genetic and molecular technologies advanced enough to permit the identification of the causative gene, Bruton tyrosine kinase, *BTK* (3, 4). Within the past 5 years, new technologies for the investigation of genes has resulted in a cornucopia of newly defined immunodeficiencies and immunoregulatory diseases and an explosion of knowledge on human immune regulation (5, 6).

Genomics is the investigation of nucleotide sequence variants in the entire nuclear DNA content of an organism and their effect on phenotype. Up until recently, the molecular state of the genome in a patient's cell was essentially unknowable. Next generation sequencing (NGS) technologies have now made it routine. Molecular genetic analysis in humans has

also been accelerated by new techniques that allow biochemical and molecular analyses of human genetic defects with a rigor approaching that previously possible only in model organisms. These include efficient cell cultivation, improved electroporation techniques to introduce DNA into cells, and the ability to alter gene expression with interfering RNA or, more recently, genomic editing with techniques such as the clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 system (7). Gene mapping and candidate gene analyses for primary immunodeficiencies have been largely replaced with NGS (5, 6). Computational approaches have streamlined the identification and validation of causal gene variants. While cross-validation of human disease genes with genetically altered mice, a species that shares 99% of its genes with humans, is a mainstay of immunology, there are many examples in which human disease cannot be modeled in the mouse, such as caspase-8 deficiency or CTLA-4 haploinsufficiency (8, 9). Thus, we are now technologically capable of diving into an ocean of 3 billion nucleotides with a reasonable hope of swimming back to the surface with pearls of fundamental knowledge about the healthy and diseased immune system. This essay is about our experience making such deep sequence explorations.

Approaches to study the genetic contribution to disease

The first DNA sequence of a human genome introduced the possibility that genomics could yield new insights into cellular regulation and disease mechanisms (10). However, knowledge about function required comparative genomics in which DNA sequence variations between different individuals, especially those with specific diseases, were investigated. Early on, cost prevented the wide deployment of DNA sequencing, so alternative genomic screening strategies were developed. The hypothesis emerged that common single nucleotide polymorphisms (SNP) would correlate with common diseases (the common SNP-common disease hypothesis) (11). Relatively inexpensive hybridization chips were used to probe common well-defined SNPs in order to correlate predominant allele frequencies with a variety of diseases including most common autoimmune diseases, a procedure called genome wide association (GWA). For the SNPs on these chips, MHC class II alleles consistently showed significant and reproducible linkage disequilibrium with immunological diseases, but other associations showed only slight linkage (12). Specifically, very few SNPs in the coding portion of the genome had a large effect for any disease (11–13). Another complication of GWA studies is that many SNPs correlated with disease were in non-coding regions, had modest contribution to disease, and were not medically actionable (11–13). This was puzzling and left a large part of the inheritance inferred from other analyses, such as twin studies, unexplained. Also, these data were hard to interpret since they were based on statistical correlation, and determining whether they actually caused disease, though finally established in some cases, was difficult (11, 12). For certain complex diseases, GWA studies suggest that a group of gene variants, each with a small effect, act in concert to produce the disease phenotype (11–13).

As the cost of sequencing DNA dropped precipitously and access to NGS technology became widely available, emphasis shifted to direct gene sequence investigation. Currently genomic investigation uses NGS of exonic DNA captured by hybridization chips, "the exome", or the entire nuclear DNA of a cell from a human subject, "the genome." The study of individual exomes and genomes, now increasing exponentially and including a panoply of

disease cohorts, yielded a big surprise. Every human genome sequenced contains a surprisingly immense number of rare variants compared to the reference genome and these are usually private, meaning that they are restricted to the immediate blood relatives of the family, parents, siblings, and children (14). Hence, the genomic "software" for every newborn is unique with many variants – generally bugs but occasionally improvements – that affect the cellular programs. Furthermore, in contrast to the weak effects of common SNPs, rare single nucleotide variants (SNVs) and copy number variants (CNVs) can have highly penetrant and deleterious effects on phenotype (14). Single nucleotide changes that are called SNPs are found at frequencies greater than 1% in a population and those with frequencies less than that, usually exceedingly low or unique to a single individual, are called SNVs. How can we interpret the difference in disease heritability between common SNPs and rare SNVs? Any DNA change arising in a family with strong deleterious effects would be eliminated from the population as it is passed on to successive generations because of decreased reproductive fitness. Thus, common SNPs would have been selected for weak disease effects in order to be successfully transmitted within the population. A caveat to this assertion is that common SNPs could not have been selected against by recent specific environmental influences such as drugs, toxins etc. and therefore, some may still have unexpectedly potent deleterious effects. On the other hand, most rare variants would be limited to one or a small number of generations in a family depending on how severely it affects reproductive fitness. The concept of rare, family-restricted variants that have severe phenotypes has been called "clan genomics" and pursuit of these variants has had a major impact on genomic investigation of diseases (14). Rare variation with large effect sizes is likely contributing a significant proportion to the "missing heritability" of complex traits and disease (14, 15).

Why have an astronomical number of private variants (thousands or more per genome) with low minor allele frequencies (MAFs), been uncovered in human genomic sequences around the world. This is likely because of our rapid evolutionary success and the dramatic increase of human populations in a comparatively short span of evolutionary time (16). There may also be increasing retention of disease susceptibility variants that formerly would have been selected out but now persist because of better health through hygiene and medicine. The other feature of clan genomics is homozygosity of private variants through consanguinity. Early in human history, certain isolated hunter-gatherer clans likely practiced a high level of endogamy, causing homozygous retention of private variants through consanguinity (17). As small populations moved or suffered from disease or starvation, the gene pool would be subject to genetic drift and founder effects through population bottlenecks. Even today, this tendency to consanguinity has persisted through social, economic, or religious forces. Large regions of the world, comprising roughly 10% of humanity, have frequent consanguineous reproduction that retains and promotes homozygosity of private SNVs, especially deleterious loss-of-function (LOF) mutants, without dispersing them throughout the regional or national population (16, 17).

Powerful and evolving NGS technologies, now provide the principal approach to identifying disease-associated SNVs and short (< 50 base pairs (bp)) insertions and deletions (Indels) (5, 6, 15, 18, 19). Whole genome sequencing of paired ends (WGS), comparative genomic hybridization (CGH) arrays, and chromosome analysis can identify structural variations

(SVs, defined as > 50 bp) including CNVs and more complex genomic rearrangements (18, 19). Individual genomes may harbor nearly 10,000 SNVs, affecting almost 2% of the genome, which will be an important source of disease variants (18–20). These new genomic technologies provide an unbiased method to identify gene mutations in Mendelian, monogenic, and *de novo* inheritance patterns. Not all monogenic disorders can be considered Mendelian, since *de novo* gene variants that are congenital but not inherited, and by definition private, cause important disease phenotypes (18, 20). Family-based WGS studies have estimated that each individual's genome contains many germline *de novo* mutations (DNMs) and that these are an important source of disease mutations (20–22). These mutations, found in the affected proband but not in the parents or siblings, are potentially more deleterious because they have not been subject to natural selection. DNA sequence analysis of selected cell populations can also uncover somatic mutations (SM) such as *FAS* gene mutations that are present in a subpopulation of CD4-, CD8- T cells but not in the germline of certain autoimmune lymphoproliferative syndrome (ALPS) patients (23). Somatic reversion can also ameliorate disease phenotype by "correcting" the genetic defect in appropriate cell lineages (24).

In consanguineous families, NGS of affected individuals together with parents and other affected or unaffected relatives combined with homozygosity mapping can efficiently identify disease-associated variants. Non-consanguineous, large multigenerational, and multiplex pedigrees can also be used to identify rare inherited variants. NGS has caused immunology research to emphasize forward genetics, in which a gene(s) variant is sought to explain a specific phenotype, usually a disease. Forward genetics utilizing NGS provides the best unbiased approach to correlating genotypes with phenotypes and has already added a substantial number of new entities to over 250 genetic diseases of the immune system (25). The new technology has also uncovered thousands of disease-associated human gene variants (14, 18, 19, 25). As more comprehensive datasets are assembled, correlating gene variants with various immunological pathologies and infection susceptibility will reveal the genetic landscape of immune diseases.

NGS data coupled with bioinformatics analyses have become increasingly valuable for diagnosing known genetic disorders but are generally not adequate on their own to conclusively implicate new genetic variant(s) in disease pathogenesis (26). When computational approaches do succeed, it is usually when there is a clearly defined clinical phenotype that occurs in multiple unrelated individuals and maps to the same gene in which nucleotide changes cause similar biochemical effects (27, 28). Additional biochemical and molecular validation is essential. This is because there are a large number of variants but the availability of affected samples for rare diseases is often limited. Also, there are surprisingly many severe LOF variants found in ostensibly normal people, almost 100 genes per individual with 20 being homozygous (29). The problem of inadequate validation has been documented in population studies showing that hundreds of variants reported in the human gene mutation database as disease alleles and lacking adequate biochemical validation could be found in normal individuals (15, 26, 30). Although statistical and bioinformatics tools will improve, the most convincing association of gene variants with disease will be based on molecular and biochemical validation in human cell lines or experimental organisms such as mice, zebrafish, or others (26, 31). Validation requires addressing three major issues: How

does the nucleotide alteration affect the protein biochemically? How does the gene mutation fit a specific model of inheritance? This can be complicated by reduced penetrance or expressivity of the genetic change. Finally, how does the specific biochemical change derange the cellular pathways in a manner that potentially explains the observed phenotype/ disease? As described below, validation is greatly simplified when a "cellular" phenotype that clearly relates to disease manifestations can be defined.

In the remainder of this review, we will walk through the genomic process of disease gene discovery starting with patient phenotype and ending with the molecular definition of a new disease and its therapy. We will use the example of identifying specific mutations in the *MAGT1* gene as causative of "X-linked immunodeficiency with magnesium defect, Epstein-Barr virus (EBV) infection, and neoplasia" (XMEN) disease to illustrate the gene discovery process (32). These examples will be in italics. Knowing the identity of the XMEN disease gene offered new theoretical and basic insights into magnesium regulation as well as a direct etiological therapy for the most important consequence of the disease, uncontrolled active EBV infection (33, 34). At the end of the review, we will describe two additional diseases for which knowing the specific gene defect led directly to specific improvements in therapy for life-threatening immune diseases.

Clinical and laboratory phenotype

The process begins with the chief complaint that the patient brings to medical attention. It can be severe or, as was the case of XMEN patients, a mild increase in susceptibility to childhood infections (32). In the realm of immune disease, the phenotype often relates to deficiencies in immune responses manifested by viral, bacterial, fungal, or parasitic infections (2, 8, 35, 36). Conversely, it can involve overactive immune responses leading to cellular or antibody-mediated autoimmunity or atopic disease such as asthma (37, 38). Immune homeostasis can also be deranged with the accumulation of immune cells and/or the distortion of immune subsets. Often, there are combinations of these abnormalities that defy obvious explanation due to the complex web of regulatory interactions in immune responses, for example, severe immunodeficiency with excessive IgE (37). Reasons to believe an immune abnormality is primarily genetic include early-onset, occurrence in a multiplex family, marked severity of disease compared to the normal population, and a characteristic "syndrome" that emerges in unrelated families and cannot be attributed to environmental causes. Clinical lab examination, depending on the depth of analysis available, can uncover abnormal lymphocyte subsets and basic functional assays in B and T cell activation and proliferation. The collective clinical data becomes a foundation of the phenotype upon which the research lab builds.

The first two XMEN patients were brothers who had a clinical history of mild immunodeficiency with repeated but generally self-limited respiratory and gastrointestinal infections (32). Overall, the boys were developmentally normal and were otherwise apparently healthy on clinical exam. Clinical laboratory investigation revealed the fact that the boys had strikingly elevated EBV levels that ranged between 100,000 and 1 million copies per mL blood compared to normal individuals who are often completely negative but can occasionally have up to 800 copies per mL (34). Thus, the clinical phenotype, mild

infectious history and astronomical levels of EBV in developmentally normal males, was defined by these analyses. Two affected males in the same family suggested a genetic basis.

Because of the general availability of cells in immune diseases through blood sampling, more elaborate tests can focus on defining the functional defects characteristic of the disorder. A key goal is to define a "cellular phenotype" as a benchmark of the disease. This clarifies the genetic model and provides a molecular confirmation of candidate genetic variants. Later, it can be used as a diagnostic test for individuals suspected of having the disease. Immunophenotyping by flow cytometry for lymphocyte subsets can document derangements in lymphoid development or survival, proliferation, and differentiation. Though this cannot survey tissue-resident immune cells, it can assess many distinct cell types.

In the case of XMEN patients, flow cytometry revealed a modest decrease in peripheral naïve CD4 T cells that inverted the CD4:CD8 ratio and was accompanied by a decreased number of CD31+ recent thymic emigrants (32). These alterations were mild, however, and would not have been expected to cause severe infectious consequences, especially the chronic active EBV infection. Functional analyses were more revealing by showing that both CD4 and CD8 T cells were impaired in upregulating activation markers including CD69, CD25, and CD95 after T cell receptor (TCR) cross-linking (32). Interestingly, this deficit was overcome by using a combination of second messenger mimics (phorbol ester and ionomycin), implying a defect in proximal TCR signaling. By contrast, the patient B cells were normal in number and response to B cell receptor and toll-like receptor stimulation (32).

Modes of inheritance

With the cellular and clinical phenotype in hand, NGS was a logical approach to deciphering XMEN disease, but evaluating SNVs depends on the presumptive mode of inheritance determined from the family pedigree (18, 19, 25, 26). The principal modes are derived from classic Mendelian recessive (both alleles abnormal and skips generations) and dominant (one allele abnormal and affects successive generations) concepts as well as the possibility of haploinsufficiency, X-linked disorders, and *de novo* occurrence of mutations. In current databases of Mendelian monogenic immune diseases, autosomal recessive (AR) disorders are 3–4 times more common than autosomal dominant (AD) diseases partly because they were the easiest to solve using classic approaches (25). Immunodeficiencies due to AR LOF mutations can be caused by homozygous alleles or two different LOF alleles (biallelic or compound heterozygous) (37, 38). In countries with high consanguinity, AR disorders involve LOF alleles that are usually homozygous. AD disorders are increasingly being reported and arise mainly in two genetic forms: activating - in which the variant allele has hyperactive function; and dominant-interfering - in which the variant allele impairs the function of the normal allele. AD disease, the other side of the Mendelian coin, have been increasingly identified because NGS can uncover causative mutations in small pedigrees or large numbers of unrelated people even with marked variations in penetrance. PASLI disease (p110δ-activating mutation causing senescent T cells, lymphadenopathy and immunodeficiency, also known as activated PI3K-δ syndrome, APDS) is an AD disease due

to heterozygous mutations that upregulate phosphoinositol-3′ kinase activity due to changes either in the gene encoding the leukocyte-restricted p110δ enzymatic subunit or the ubiquitous p85α regulatory subunit (27, 28, 39, 40). How the enzymatic "gain of function" (GOF) causes lymphocyte proliferation, differentiation, and senescence with metabolic changes that prevent adequate immune responses will be described in greater depth below. The second type of AD mutations is dominant-interfering (DI) (41, 42). These mutations usually produce an anomalous protein that interferes with the function of the normal protein. For example, in ALPS, apoptosis signaling is inhibited by heterozygous alleles encoding defective FAS proteins that bind to the normal protein and hold it hostage in defective receptor complexes (41, 42). Similarly, in AD autoimmune polyendocrine syndrome (APS), mutations in the plant homology domain 1 (PHD1) generate defective AIRE proteins that sequester the normal AIRE protein in abortive transcriptional complexes in nuclear speckles (43). Rarely there can be genetic gain-of-function mutations that lead to immunodeficiency which some have called autosomal dominance of the "third" kind (44). One specialized Mendelian mode of inheritance, combining heterozygosity and LOF, is called haploinsufficiency (HI), for "half sufficiency" because the normal phenotype depends on full diploid gene dosage and a single defective allele reduces the amount of protein by half (45). These behave genetically like AD alleles. For example, patients with "CTLA-4 haploinsufficiency with autoimmune infiltration" (CHAI) disease have heterozygous, germline LOF mutations in *CTLA4* (9, 46). These decrease CTLA-4, resulting in abnormal T cell infiltration and damage to organs such as the lung, brain, and bowel (9). Another class of mutations involves the X chromosome, termed X-linked recessive (XLR). These predominantly affect males who inherit hemizygous defective alleles from their mothers. There are several highly penetrant immunological XLR disorders including X-linked SCID, IPEX, and XMEN diseases (32, 47, 48). These resemble recessive mutations because they are inherited through an unaffected mother. Finally, other important classes of mutations that are not, strictly speaking, Mendelian, are DNMs, found in the proband but absent in parents and siblings. The DNA replication error rate is estimated to be 1 per 10^8 base pairs, generating 30–100 DNMs in each generation and serving as the source for unique private SNVs. These increase with paternal age and can also be mosaic by occurring post-zygotically (20, 49, 50). *De novo* mutations can be experimentally identified by NGS of trios comprising the parents and affected child (50, 51). A variant type of DNM occurs in somatic cells. SM usually contribute to genetic disease by conferring a proliferative or survival advantage as well as an abnormal phenotype in the affected cells. Cancer can be thought of as the chief example of this.

For XMEN disease, the occurrence of disease in two boys suggested the possibility of an XLR mode of inheritance. This hypothesis was supported by the skewing of X chromosome inactivation in the unaffected mother such that nearly all of her hematopoietic cells showed inactivation of the X chromosome that she transmitted to both boys (32). This provided strong prima facie evidence for a mutation on the X chromosome that was deleterious to her hematopoietic cells (hence only cells with the normal X chromosome could be found) and responsible for X-linked disease in her sons. Solution hybridization reagents targeting the X chromosome were used to prepare exons from the boys' and the mother's DNA for NGS.

Sequencing of the captured DNA revealed a 10-base-pair deletion in a magnesium transporter gene called MAGT1 located on the X chromosome (32).

Phenotype variation

In reverse genetics, inbred mouse and rat strains aid the genetic engineering of disease models because it reduces background variation and generates consistent phenotypes. Although 10% of the world is consanguineous, modern human populations are generally highly outbred (16). Thus, in humans, the phenotypes of gene mutations can be highly variable because they shine through heterogeneous genomic atmospheres. This genetic admixing, together with environmental influences, makes it difficult, particularly with small numbers of patients with a given gene variant, to precisely define disease phenotype. Coupled to this vast tumult in gene mixing, human genetics has a powerful phenotype detection mechanism provided by the medical and research professions that globally detect rare and variable phenotypes. This has increased the pace of gene discovery and validation but also revealed how disease associated with a specific gene can present in many different guises.

In human genetics, variation associated with a specific genotype, particular for disease genes, was classically described in two ways: penetrance and expressivity (52–55). Penetrance refers to whether or not a phenotype, in a sort of all or none fashion, was present in individuals with the same gene mutation. Expressivity was used to describe pleiotropy associated with specific gene mutations. With the dizzying array of human conditions associated with genomic variations, the difference between these two concepts has been blurred. Penetrance probably best describes syndromes in which there is a highly stereotypical set of disease features where not everyone with a specific gene mutation gets the disease. For example, in ALPS, healthy parents harboring a dominant-interfering *FAS* mutation can pass the mutation to a child who develops a clinically significant disease (41). In this situation, where unaffected individuals harbor identical gene mutations to those with full-blown clinical disease, one would say the penetrance is incomplete. Mathematically, penetrance measures the proportion of affecteds in a given population of individuals with a specific genotype (55). In ALPS (AD) and CHAI disease (HI), penetrance ranges around 50 to 60% (9, 41, 45). No molecular explanation currently exists for this surprisingly low penetrance. It is also important to distinguish incomplete penetrance of a disease-causing variant from risk alleles for complex diseases that are identified in GWA studies and cannot, on their own, serve as a primary cause of disease. Variations in expressivity are ubiquitous; essentially every genetic disease associated with a specific gene shows variation in the severity and type of clinical manifestation. As a consequence, this term is used less often in contemporary genomics.

There are a multitude of causes for variation in penetrance and expressivity (Figure 1) (52–55). Some identified causes intrinsic to the genome are: the molecular nature of the gene mutation and the degree to which it is damaging to protein function; additional mutations either in the same gene (*cis*) or other genes (*trans*); inter-individual gene expression differences even between alleles; gene dosage; digenic disorders in which there are actually two obligatory causative mutations; sex dependence; age; imprinting; somatic mosaicism,

and epigenetic modification. Of special significance are "modifier" genes (see below) that are often unlinked to the primary disease-causing gene but critical determinants of the genetic "atmosphere" that influences penetrance and expressivity. For example, low expression levels of the affected gene can lead to a more severe abnormal phenotype and this can be determined by genetic background (55). In addition there are extrinsic causes due to environmental influences on the affected individual (52–55). A good example of this is genetic susceptibility to viral pathogens or mycobacteria that is silent until exposure to the specific pathogen causes severe disease; individuals who never encounter the pathogen will have zero disease penetrance (56, 57). Differences in penetrance and expressivity create disease disguises that make it difficult to implicate a specific gene variant as the cause of a genetic disease. Violations of the mode of inheritance and inconsistencies in the pedigree, uncertainty about which disease characteristics are the most relevant, and confounding effects from disorders that are actually due to a different gene variant are significant challenges for genomic detective work. Only with further investigation of the specific genotype and phenotype, especially the accumulation of more patients with the same genotype and phenotype does the empiric association between the two become convincing.

There are approximately two dozen XMEN patients from different countries and ethnic backgrounds, and the clinical and immunological phenotype is remarkably consistent. Although females can be carriers, there are no known males harboring silent MAGT1 mutations, perhaps indicating that the disease is 100% penetrant on the cellular level with variable age of apparent symptom onset (34). As noted above, patients showed a decrease in CD4+ T cells, mild immunodeficiency, and a severe loss of immunity against EBV. Accompanying lab findings include a reduction in activation markers such as CD25, and a reduced level of NKG2D on CD8+ T cells and natural killer cells. Yet, how these features fit together to explain a mechanism of disease was a puzzle.

Genotype variation and bioinformatics

Once the genome has been interrogated by CGH arrays or NGS, variations can be defined by comparing the patient sequence to the standard reference assembly as well as family members (5, 6, 15, 25, 51, 58–60). As alluded to above, a large number of SNVs will be identified which leads to the problem of which variants are potentially disease-causing. Exonic sequences will reveal alterations in the coding sequences of particular genes. WGS can also be used to detect alterations in non-coding regions in human disease (58–61). However, the definition of disease-causing alterations in non-coding regions is still in its infancy because it involves casting a searchlight across immense stretches of the genome with complicated regulatory structures that are still only dimly perceived. Even within coding regions, analysis is challenging because there are thousands of variants, many private, to be sifted through (14, 26). The problem was documented in the first analyses of the thousand exome project (26, 30). Even considering only severe LOF variants such as nonsense, indels, and splice site-disrupting SNVs, it was estimated that genomes from healthy humans could contain up to 100 LOF variants of which 20 would be homozygous (29). A similar analysis of missense mutations from this data set showed that 281–515 missense substitutions were predicted to be protein-damaging and LOF (see below) and between 40 and 85 would be homozygous (30). Further compounding the problem was the

fact that a shockingly large number (hundreds) of variants in the human disease mutation database were found in normal people and that the original publication of these variants had no biochemical validation of disease involvement (30). Thus, NGS interpretation is prone to false-positive disease correlations (26, 30). To surmount these problems, recommendations for variant interpretation emphasize weighing multiple lines of evidence. When available, statistical arguments across many families is certainly strongly supportive, but difficult to deploy (26). But this is difficult to deploy. Most disease phenotypes are very rare and it is difficult, without the guidance of an identified gene, to assemble a well-defined patient group. For more common conditions, general and less specific phenotypes are binned together (e.g., common variable immunodeficiency) and later prove to have multiple genes involved, thereby defying a uniform approach (62, 63). Indeed, in defining over a dozen new diseases, our program has not been able to use a statistical study design.

The approaches we prefer to use are: 1) judicious choice of candidate variants by a stratification process (51, 58, 60); and 2) biochemical and molecular validation. These have the virtues of proven success and providing a molecular understanding of disease that benefits treatment and prognosis. There are a variety of software tools that are continuously evolving to stratify gene candidates. We will discuss the guiding principles for variant evaluation and will leave the reader to consult with bioinformatics experts for specific recommendations for genomics projects. The stratification of candidate variants is essentially a guesstimate of the probability that a given variant is disease-causing generally using a population of affecteds that is too small to allow a statistical correlation. First, it is crucial to ensure that the variant is not a sequencing or annotation error (51). NGS technologies are high-throughput approaches designed to generate bulk but not complete sequence information and a sequence variant can be most accurately described if it has been redundantly sequenced in the dataset, usually expressed as the read number at that site or "fold coverage." 30 is a good number; 100 is better, but beyond that is probably unnecessary (51). Medium-sized indels and CNVs can be hard to judge with confidence. Large CNVs or rearrangements will require CGH arrays or chromosome banding analyses. Even with good coverage, once a variant becomes highly suspicious, it should be confirmed by conventional Sanger sequencing on a separate sample from the patient.

Another initial consideration is whether or not the mutated gene is expressed in cell types with an abnormal phenotype relevant to the disease. It is generally safe to assume that the altered gene must be expressed in the affected cells, especially if there is a strong cellular phenotype. Further filtering of gene candidates can then be prioritized considering the presumptive mode of inheritance. Depending on the clarity of the clinical phenotype, the pedigree provides the roadmap for navigating the mutant allele among affecteds and unaffecteds (58). Extensive NGS has shown that nucleotide variation is not uniformly distributed in the genome and a variant further gains significance if it falls in a conserved region of the genome, indicating that the function of that region does not tolerate much variation (20). Conversely, certain types of genes (a random example of this would be olfactory receptors) frequently harbor LOF mutations and are an unlikely cause of immune disease (58). Also important is whether the MAF shows that the variant is unique or very rare in the variant databases (ExAC, dbSNP, etc.). There is no precise cutoff but < 0.1% could be allowed since the reported SNV universe is so vast and covers so many different

groups of normal and diseased individuals that there is a chance that the variant would have been reported. For example, heterozygous carriers of AR mutant alleles will be present in "normal" control populations because they will have no disease. Nonetheless, a high MAF decreases the chance a variant is disease-causing. Together these strategies can drastically shrink the list of possibilities from thousands to less than 100 (47, 51, 56, 58–60).

The common assumption for a disease-causing variant is that it will have a severe deleterious effect on gene function (26, 31, 51, 58, 64–67). Largely due to the wide use of exome sequencing, private variants with big effects have been located mainly in protein-coding regions of the genome. These currently account for about 85% of all Mendelian disease variants (63). Generally, nonsense mutations together with deletions (large > small) and splice variants, especially if they are frame-shifting, are more likely to be deleterious then missense mutations (26, 29). Base changes that simply substitute an alternate codon for the same amino acid, termed synonymous, are usually ignored. The problem then becomes how do you tell if a non-synonymous change is likely to be deleterious just by looking at it. Fortunately, ingenious software has been developed to consider various parameters of the non-synonymous change to determine if it is likely to be damaging to the protein (64–67). These include whether the mutation is in an evolutionarily conserved residue, whether it is a marked chemical change, e.g., substituting an acidic residue for a basic residue, whether it affects buried residues, whether it induces a steric clash between amino acids, and how it affects surface electrostatics, among other considerations. Generally, if the functional or conserved parts of the protein structure are drastically compromised, then it is judged to be likely damaging to the protein. Widely used examples of these programs, which are constantly evolving, are SIFT, PROVEAN, CUPSAT, PolyPhen2, MutPred, MutationTaster, GERP and PhyloP, and CADD and variations thereof (64–67). Studies have shown that these programs have great value in prioritizing candidate disease variants (64, 66). Also facilitating disease variant identification are databases that aggregate reference exome data (e.g., http://exac.broadinstitute.org/). These techniques can reduce the candidate variant pool usually to one or a handful of possibilities.

The search for a pathogenic variant in XMEN disease was simplified to a great extent by the hypothesis that the gene was X-linked based on the mother's skewed X chromosome inactivation (32). However, when the variants were prioritized, no apparently protein-damaging candidate variant emerged that was present in both boys but heterozygous in the mother. The absence of a suitable variant raised a suspicion about the coverage of genes on the X chromosome by the sequencing analysis. In fact, the software was designed to exclude mismatches over 2 bp as erroneous sequences and would miss larger deletions because they would be consigned to the "junk pile" of sequences from the sequencing run. If such an exclusion had occurred, then there would be genes covered in the mother that had no coverage in the children. In fact, two genes met this criterion. When the reads were examined carefully, two non-overlapping partial sequences of the MAGT1 gene were found that were separated by a 10 bp gap. If this 10 bp region was removed from the X chromosome reference sequence and the boys' sequences were reanalyzed, then a large number of sequences would be recovered from the software's "junk pile" of discarded sequences, showing a 10 bp deletion in the genomic DNA of both boys (and heterozygous in the mother) that spanned the gap between sequences that had passed the quality control

filter. Thus, by examining the "coverage" of genes, a problem with the software filters was detected and overcome. Further analysis of the deletion showed that it spanned an intron-exon junction causing a frameshift and premature stop codon in the protein (32). This variant was predicted by bioinformatics to be a deleterious mutation.

Biochemical validation

The critical final step in the journey and, unfortunately, often the longest and most labor-intensive, is validation (26, 68, 69). This is the bottleneck in genomics research that connects DNA sequencing and computational analysis to clinical application (Figure 2). This includes both making an argument for causality and developing a molecular understanding of the disease mechanism. If a strong cellular phenotype has been defined, then cellular knockouts or knockdowns using CRISPR, siRNA, or shRNA to decrease expression of the gene or transfection of the mutant allele into normal cells might yield data that the gene variant can recapitulate the cellular phenotype associated with disease (7, 8, 28, 32, 40, 68, 70). Also, introducing a normal gene version or knocking out a dominant variant might be sufficient to correct the abnormal phenotype in the patient cells (28, 32, 71). It is crucial to carry out biochemical investigations of the putative mutant protein since protein levels correlate poorly with mRNA in human cells (72). If a cellular phenotype is not apparent or the disease involves a developmental process, say, the differentiation of T cells, then more elaborate investigations are needed (26, 69, 73). Mouse models often reflect the cardinal features of human genetic diseases but are often incomplete (74). Moreover, they may not reflect complicated human genetic disease mechanisms such as haploinsufficiency (9). Researchers can also use emerging new technology to induce pluripotent stem cells (iPS) as a renewable source of cells to test gene variants. iPS can be differentiated into relevant cell populations *in vitro*, allowing gene experimentation in a physiological context (57, 68). These approaches can build the case that the variant is potentially disease-causing. The discoveries may then be ultimately confirmed by more patients and research as well as a corresponding phenotype in a relevant animal model (26, 68, 73).

Validation of the first XMEN mutation was guided by a predicted premature stop codon in exon 7 (out of 10), a situation that typically leads to nonsense-mediated decay of the mRNA and accounts for loss of protein in roughly 30% of genetic diseases (75). Indeed, in both affected boys, the mRNA was found to be depleted by polymerase chain reaction of cDNA, and a protein blot showed absence of the MAGT1 protein (32). A strong cellular phenotype in the form of defective TCR but not B cell receptor (BCR) signaling was observed in the patient cells and could be recapitulated by knocking down the MAGT1 protein in normal T cells. Furthermore, artificial expression of a normal MAGT1 gene transferred into patient T cells was able to correct the activation defect. Also, background literature on the gene indicated that it was a magnesium transporter, so Mg^{2+} levels and mobilization in immune cells were evaluated after stimulating T cells through the TCR. These investigations showed that there was a rapid TCR-gated Mg^{2+} flux that was required for a rapid and optimal Ca^{2+} flux. Further biochemical experiments revealed that the Mg^{2+} influx rapidly activated phospholipase C gamma 1 (PLCγ1) which cleaves phosphoinositide lipids into inositol triphosphate (IP_3) and diacylglycerol (DAG) (32). The reduction of IP_3 and DAG generation explained the blunted Ca^{2+} flux and impaired protein kinase C θ phosphorylation,

respectively. Together, these defects short-circuited the induction of nuclear factors such as NFAT and NF-κB that control the transcriptional program underpinning normal T cell activation. Other pathways of TCR signaling involving p38 and ERK were Mg^{2+} independent, indicating an uncoordinated but not completely deficient molecular orchestration of signaling circuits needed for a healthy T cell response. Correspondingly, the patients have a relatively minor immunodeficiency. Experiments also showed that BCR activation of the orthologous enzyme PLCγ2, does not involve any Mg^{2+} flux, explaining why the number of B cells and their function in XMEN patients was the same as healthy subjects (32). Thus, the molecular validation of the MAGT1 gene was consistent with the patients' cellular phenotype.

Determining the disease mechanism for a specific variant or the gene itself may be straightforward or may involve substantial investigation. This is the art of experimental immunology, creatively combining cellular, molecular, and biochemical analyses with an understanding of disease manifestations. The central guidepost is the patient phenotype as defined by astute clinicians. It is very helpful if a mouse model is available since many cell populations can be interrogated and well-established autoimmune disease models are available. Also, viruses, bacteria, and parasites can be used to test immune responses. If one can connect the mutant to well-studied molecular pathways, then there will be more reagents and fundamental knowledge to speed up the investigation. However, if the disease-implicating evidence from the genetics and the cellular phenotype is strong, then it would be a mistake to shy away from gene mutations that occur in little known pathways or "don't make sense." This may yield immunological insights with the greatest novelty and impact, but the investigative journey could be long and frustrating. Genomics is an unbiased exploration of genotype-phenotype interactions – so it has extraordinary potential to take the investigator into new uncharted territory.

The hallmark of XMEN disease is uncontrolled EBV infection that creates a strong susceptibility to B cell lymphoma even in childhood (33, 34). This selective pathogen susceptibility was difficult to explain on the basis of a relatively mild TCR signaling defect. Further analyses of antiviral mechanisms revealed a severe reduction of the NK cell activating receptor, NKG2D, which was observed in numerous XMEN patients who suffered uncontrolled EBV infection. Moreover, B cells infected with EBV showed increased expression of specific ligands that mark the cells for cytolytic destruction via NKG2D. With a deficit of NKG2D, these infected cells became invisible to anti-viral immunity and escaped elimination. MAGT1 deficiency reduced intracellular basal free Mg^{2+}, which impaired proper surface expression of NKG2D. Remarkably, adding excess Mg^{2+} to XMEN lymphocytes reversed the intracellular deficiency, resulting in rescue of NKG2D surface expression and cytolytic function against EBV-infected targets (34).

If we stop for a moment and ponder our journey through the incredibly powerful new genomic technologies, we realize that human genetic variation in disease is a gold mine for insight into molecular function. The depredations of immune function caused by genetic mutations have the potential to instruct us to how the immune system operates. As pointed out above, careful work must be done to be sure that the variant under investigation isn't "fool's gold." However, the full value of striking gold isn't achieved until it is worked into a

lustrous understanding of an immune mechanism. Currently in medical genetics, we believe that there is an overemphasis on generating DNA sequence from patients and insufficient emphasis on the biochemical work to understand the molecular and pathophysiological importance of gene variants that are detected. We have argued that if one assumes that 200 families with unknown genetic immunological diseases, for instance parent/child trios, will generate 200 terabytes of DNA sequence data and this will require 1–2 bioinformatics specialists to generate candidate variants, then it will likely require 10 to 12 biochemists and immunologists to biochemically validate the causative variants (69). However, many clinical sequencing groups have not prioritized biochemical efforts so that the mutant proteins can be convincingly validated or investigated in sufficient depth to understand the molecular mechanisms of disease and possibly identify therapeutic strategies. Hence, in-depth functional validation and biochemical studies are essential future goals for human genetic studies. Further, the study of the immune system has the unique and powerful advantage of being accessible for study and treatment, much more so than in cardiology, neurology, ophthalmology, etc. This level of effort will fashion the genetic "gold" first into amulets of understanding and then, hopefully, into better diagnosis and treatment.

Genetic etiology of disease and complex adaptive system theory

In the pre-NGS era of medical genomics, diseases were divided into Mendelian, meaning monogenic (or possibly digenic, but not more), and complex, meaning polygenic (13, 63, 76). As we pointed out above, even Mendelian diseases reflect the contribution of multiple genes. Conversely, NGS has revealed that genetic diseases heretofore thought to be polygenic – autism, mental retardation, and others, are actually heterogeneous collections of individuals with private, highly penetrant deleterious mutations (20, 77). Immunogeneticists have discovered that common variable immunodeficiency (CVID) can be explained in the same way (63). It is likely that rare, highly penetrant mutations may account for other immunological diseases including the common autoimmune conditions. This is where the nascent field of systems biology can play an influential role. By assembling the affected genes in a common phenotype into a regulatory network, systems analyses may be able to place new variants (private, highly deleterious and penetrant) into a logical framework of disease (78).

Understanding the effect of genes on cellular phenotype requires a model for understanding how the linear one-dimensional information in DNA generates a three-dimensional dynamic system. The protein product of a disease gene does not work in isolation but is part of a large interactive system (76). A useful concept for understanding genetic interactions is to view the cell as a complex adaptive system (CAS) (79). A CAS is a collection of diverse elements that self-organizes according to simple rules into a "system." Although analogous to many different interactome models, a CAS focuses on defining features that enable the system to adapt to new and unpredictable circumstances. Modeling has shown that CAS can generate large intricate systems with spectacularly complex behavior that surpasses the action of individual parts in surprising and unexpected ways (80). Intuitively, this seems to be an appropriate way to model how molecules form the structures and systems of immune cells.

CAS are known to exhibit three key properties: robustness, adaptability, and "tipping." Robustness refers to stability in the face of unpredictable perturbations, such as genetic alterations, through a network of interacting elements (81). Functionally speaking, robustness – the tendency of the system to remain stable – and adaptability – internal real-time changes (adaptations) to preserve function – are two sides of the same coin. To be clear, adaptability here does not refer to Darwinian adaptations through natural selection, but immediate adjustments in biochemical responses in the life of a single cell or organism to the effects of mutations under different cellular conditions. NGS has shown that individual human genomes tolerate a large number of private deleterious mutations in individuals that appear to be healthy (29, 30). How the CAS in different individuals can adapt to LOF or otherwise disease-causing changes to robustly preserve function is the key to understanding disease pathogenesis. Recent examples of adaptability to genetic loss or severe environmental stress on cells have been described (82, 83). In these instances, the random occurrence of new mutations and the unpredictability of pathogens and the environment, means that there can be no foreordained plan of adaptations to achieve robustness in the system. These same issues underlie any explanation of how penetrance and expressivity can vary so greatly even to the extent that relatives of patients harboring the same genetic defect can remain healthy (41). From the perspective of a CAS, the wide variation in the parts of the genetic system: gene families, allelic differences, polymorphisms and rare variants promotes rather than detracts from robustness. This diversity of components may permit the system to adapt in real time. For any given gene, the complexity of the system is built upon gene families and non-synonymous allelic differences that will create different protein homologues which could be further diversified by post-translational modifications to create a network of similar protein forms. Unlike the classical notion of "one gene-one protein" which implies a solitary gene function, the diversity of protein forms we now know can exist for any gene fits a CAS that uses different forms under distinct conditions of metabolism, invading pathogens, or immune processes to enable the system to operate most effectively.

In a similar vein, the interaction of a network of different genes, often referred to "modifiers", could achieve robustness by allowing adaptability to genetic defects. This model is very different than modeling the cell as a machine or an electronic circuit where each part is specifically engineered to play a specific role under controlled operating conditions (conceptually the same as the one gene-one protein concept). Machines can be highly complicated but lack the diversity that is crucial to manifesting complex adaptive behavior. When unexpected severe circumstances arise, a machine breaks but a CAS adapts (79). The other CAS characteristic, "tipping", refers to what happens when the various indirect parts begin to negatively synergize. Here the interconnectedness of the system destroys robustness. Similar to an icy interstate at night where drivers realize too late that they can't control their cars and one after another slide into a disastrous pile up. Events like lymphoid cancers or apoptosis may reflect tipping of the cellular CAS. As the broad genetic landscape is expressed into the cell interactome, it will be fascinating to understand how the cellular CAS adapts to genetic and environmental perturbation to preserve health or "tip" into disease. Inbred mouse models may provide valuable insights into human disease by allowing experimental evaluation of different alleles and interacting genes in controlled systems (76).

Genetic changes affect the cellular CAS in ways that belie the traditional paradigm of Mendelian versus complex genetic diseases. Moreover, understanding how these primary determinants of disease have molecular effects on the south will also give great insights into disease phenotype (76). Sometimes a variety of different genes affect the same critical component of the system and adaptability fails. For example, a variety of mutant alleles of several receptor and cytokine genes that all interfere with production or response to interferon-γ, cause the clinical entity "Mendelian susceptibility to mycobacterial disease" (MSMD) (56). On the other hand, the same gene can account for two entirely different diseases. Job's Syndrome, which features impaired Th17 differentiation, immunodeficiency, abnormally elevated IgE, and connective tissue abnormalities is caused by germline dominant-interfering STAT3 variants (84). By contrast, STAT3 heterozygous GOF variants have been identified in patients with lymphoproliferation and early-onset autoimmunity (85–87). A similar dichotomy has also been found in diseases related to STAT1 (88).

Although we have argued that the adaptability inherent in cells can create a continuum between Mendelian and complex disorders, the ends of the continuum can still be clearly defined. For Mendelian disorders, specific gene identification is achieved by each mutation being necessary and sufficient, within the bounds of penetrance and expressivity considerations, for the disease. Generally the mutations are deleterious to protein function, rare, and occur with a mode of inheritance pattern consistent with the disease phenotype (22, 31, 51). The mutation reveals the relevant biological pathway and explains the phenotype. By contrast, complex polygenic diseases, which could include multiple sclerosis, type I diabetes, and rheumatoid arthritis, etc., have disease variants that do not have the same properties as Mendelian disease variants. The associated variants in complex diseases are neither necessary nor sufficient, not rare, and often not in the protein-coding regions of the genome (12–14). Typically, they are noncoding and located in regulatory regions. Nevertheless, the definition of genes involved in Mendelian disorders will likely expose critical pathways that are involved in complex diseases (89). Classic studies have shown that the pathways detected in monogenic disorders, such as hypercholesterolemia, later prove to be valuable targets for pharmaceutical intervention for more common related abnormalities (90).

Clinical diagnosis and incidental genetic findings

The use of NGS for identifying gene variants has now started to be an accurate and cost-effective method that complements traditional molecular diagnostic tests in clinical genomics (91, 92). However, once variant implication in disease moves from the research laboratory into clinical medicine, then the DNA sequencing lab must adhere to the Clinical Laboratory Improvement Amendment (CLIA) requirements and other legal and ethical considerations pertain (93). In order to report a variant as the etiology of a disease, the lab must have a high degree of confidence in the variant assessment (91, 92). As the catalogue of pathogenic variants grows and bioinformatics tools improve, it will be increasingly possible to identify disease-causing variants accurately.

However, WES or WGS generate large amounts of data that extend beyond a specific set of gene candidates for immunological diseases and provide a wealth of genomic information

that is potentially important for the health of the patient. Potential risk alleles, spanning a wide range of diseases, was found in every healthy individual in a reanalysis of data from the 1000 exome data set (94). For example, an NGS study carried out to discover a gene mutation causing an immune disease might reveal a breast cancer susceptibility allele. Thus, the medical genetics community has begun to focus on how to integrate WES/WGS into medical care by enabling the reporting to physicians of actionable variants (95). However, many issues regarding how to integrate this process into medical care have been raised (96). Databases of clinically important variants such as ClinVar and Online Mendelian Inheritance in Man (OMIM) will allow access to collected genomic information. This will require ongoing epidemiological analysis of variants at the level of the laboratory and hospital. It will also create a need for allied genetic counselors who can interpret the genomic information and help physicians provide appropriate diagnostic and prognostic information for the patients. The Center for Disease Control and Prevention (CDC) has established a working group for developing standards for the use of NGS in clinical diagnosis (97).

Targeted therapy, precision medicine, and common immune diseases

An important aspiration in the exploration of the genetic basis of immunological disease is that it would provide a molecular approach to new therapeutics. This has been termed precision medicine (98). In the pre-NGS era, bone marrow transplantation was the mainstay of treatment for genetic diseases of the immune system. We believe that the future of genomics will provide two types of new therapies: 1) direct molecular interventions based on the pathway of the mutant protein and 2) genome editing combined with lymphocyte or hematopoietic stem cell transplantation (HSCT). Several recent examples show how gene identification has led directly to new therapeutic concepts for immune diseases. Identification in a group of patients with atopy, immunodeficiency, autoimmunity, and neurocognitive deficits of LOF mutations in phosphoglucomutase 3 (PGM3), which is an essential enzyme for the production of UDP-GlcNAc is one such example (38, 99, 100). How the defective protein glycosylation selectively alters immunity is unknown, but addition of GlcNAc to cells from the PGM3 patients restored intracellular UDP-GlcNAc levels (38). Clinical studies are underway to determine if exogenous non-diabetogenic sugars could be used therapeutically in these patients to ameliorate the immune disorders (38). Although HSCT is life-saving, sugar therapy might be useful as a simple, safer, and inexpensive therapy, especially if the patient is not a good transplant candidate (100). Below we recount additional cases of etiological treatments derived from gene identification in XMEN, CHAI/ LATAIE, and PASLI diseases.

After the initial discovery of XMEN disease, it wasn't clear that a direct etiological treatment was possible. The major pathogenesis of XMEN disease appeared to be due to the failure to control a single pathogen, EBV. Magnesium is an obligatory element in all organisms and its utilization is tightly regulated (101). The most important biological form of magnesium is the divalent cation Mg^{2+}, which is a cofactor for many metabolic reactions including all ATP-requiring kinase reactions, glycolysis, and nucleic acid enzymology. In addition to a defect in TCR signaling, MAGT1 deficiency also substantially decreases cytosolic free basal Mg^{2+} to less than 50% of normal. While studying the biochemical mechanism of reduced NKG2D expression, which results in uncontrolled EBV infection, it

was determined that the decrease in free basal Mg^{2+} was the cause for reduced expression of NKG2D (33). Further experimentation revealed that the reduction of surface NKG2D was due to ubiquitinylation and degradation. Importantly, it was also found that the free basal Mg^{2+} in the cell was in equilibrium with extracellular Mg^{2+}. This led to a series of experiments demonstrating that culturing T or NK cells from XMEN patients in supraphysiological concentrations of Mg^{2+} (> 1 mM), restored free basal Mg^{2+} to normal levels and rescued the surface expression and function of NKG2D (33). This then provided the rationale for therapeutic Mg^{2+} supplementation for the XMEN patients. Indeed, chronic Mg^{2+} supplementation resulted in a significant decrease in the fraction of EBV-infected B cells and blood levels of EBV (33, 34). The hope is that Mg^{2+} supplementation, a treatment suggested directly by the gene identity, will be a simple, low-cost way to improve anti-viral immunity, suppress EBV, and avoid lymphoma in XMEN patients and perhaps other patients with chronic active EBV.

CHAI and LATAIE disease

In a cohort of patients with severe immune dysregulation and solid organ lymphocytic infiltration, we identified one subset of patients with an AD mode of inheritance and another subset with an AR mode of inheritance. Molecular investigation revealed heterozygous, LOF mutations in *CTLA4* in the patients with AD disease, which we named "CTLA-4 haploinsufficiency with autoimmune infiltration" (CHAI) disease (9, 46). In patients that looked very similar but lacked *CTLA4* gene mutations, we detected biallelic (homozygous or compound heterozygous) LOF mutations in the lipopolysaccharide–responsive vesicle trafficking, beach- and anchor-containing (LRBA) gene causing recessive disease, which we termed "LRBA deficiency with autoantibodies, Treg defects, autoimmune infiltration, and enteropathy" (LATAIE) (102–104). Both subsets of patients share the predominant characteristics of autoimmune cytopenias, lymphocytic infiltration of multiple non-lymphoid organs, lymphadenopathy/splenomegaly, and hypogammaglobulinemia. There was variable expressivity in clinical presentation, in which the severity of the mutation appears to play a role. The clinical phenotype ranged from severe, life-threatening interstitial lung disease, enteropathy, autoimmune cytopenias, and CVID to an ALPS-like phenotype (i.e., autoimmune cytopenias with lymphadenopathy and/or splenomegaly) with no other autoimmune manifestations. The fact that a very similar clinical picture resulted from mutations in two different genes with two different modes of inheritance raised the critical question of how the two molecular pathways intersected.

CTLA-4 is a critical immune tolerance checkpoint molecule. It is a cell surface receptor protein on activated T lymphocytes that binds and controls access to the CD80 and CD86 co-stimulatory molecules that are expressed on the surface of antigen presenting cells (APCs) with which they are interacting. By constraining co-stimulation, it acts as a checkpoint by down-regulating T cell responses. The importance of this checkpoint control is demonstrated by *Ctla4* homozygous knockout mice that develop fatal, multiorgan lymphocytic infiltration and destruction (105, 106). Almost a carbon copy, patients with CHAI disease develop extensive and destructive lymphocytic infiltrates in multiple organs, mainly the gut, lungs, and/or brain (9, 46). However, discovering the gene causing CHAI disease was complicated by two factors. First, it resulted from an AD-HI mode of

inheritance rather than AR. Second, the penetrance was incomplete. WES was performed with DNA from affected and unaffected family members and analyzed using an AD model. All CHAI individuals had the *CTLA4* mutation, but not all mutation-positive individuals in the family had disease even though all mutation-positive individuals had reduced CTLA-4 protein. Only after finding multiple families with mutations and biochemically demonstrating that the mutations impaired CTLA-4 protein function or expression was *CTLA4* concluded to be the disease causative gene.

In addition to conventional T cells, CTLA-4 is a key inhibitory receptor that is constitutively expressed on regulatory T cells (Treg). In certain contexts, CTLA-4 is crucial for proper Treg suppressive function and limits costimulation by multiple molecular mechanisms, e.g., by negative signaling, by competing with the costimulatory molecule CD28 for the shared ligands CD80 and CD86, or by removing these ligands from APCs through trans-endocytosis (107, 108). Consistent with a failure of these functions of CTLA-4, we and others found that patient T cells were hyper-proliferative and their Tregs had an abnormal phenotype and were specifically impaired in suppressive function (9, 46). Surprisingly, several patients had low antibody levels rather than high, as would have been expected from the CTLA-4 null mice, thus illustrating the potential limitations of using mouse models for predicting human disease phenotype (106). CHAI patients also had elevated numbers of anergic or exhausted CD21^{lo} B cells that are putatively autoreactive since they are often enriched for self-reactive BCRs and are found at increased levels in autoantibody-mediated autoimmune diseases (109, 110). Overall, the clinical presentation of the patients revealed a critical role for CTLA-4 in regulating both B and T cell homeostasis.

These findings still left unresolved the pathogenesis of AR LATAIE disease. *LRBA* mutations were first discovered to be disease-causative in 2012 by genetic linkage analysis of several consanguineous families with hypogammaglobulinemia and autoimmunity, followed by sequencing of positional candidate genes (104). Shortly thereafter, other patients with LRBA deficiency were identified by NGS (102, 103). All mutations caused loss of LRBA protein; however, the molecular function of LRBA or its role in the immune system were initially unknown. A study of the biochemical basis of LATAIE disease not only revealed both the function of LRBA and why LATAIE patients resembled CHAI patients, but also revealed targeted therapies (111). LATAIE patients with life-threatening interstitial lung disease, which was refractory to conventional immunosuppressants, were treated with abatacept (a CTLA4-immunoglobulin fusion protein) for several years before their genetic etiology was known. With this treatment, the patients' autoimmune conditions and lung function dramatically improved. A clue to the connection between this remarkable treatment effect and *LRBA* mutations was the tantalizing homology of LRBA to vesicle trafficking molecules (112–114). This link led to the hypothesis that LRBA plays a key role in controlling protein trafficking of CTLA-4, which is known to be tightly regulated through recycling vesicles (111, 115). Consistent with this hypothesis, loss of LRBA, both in LATAIE patients and other experimental settings, was found to cause a profound post-translational loss of CTLA-4 protein.

CTLA-4 is stored in intracellular vesicles and cycles, after T cell stimulation, to the cell surface where it must reside to perform its inhibitory function (115, 116). LRBA was shown

to bind to CTLA-4 and, in its absence, CTLA-4 transited to lysosomes and was rapidly degraded. Inhibition of lysosomes with the drug chloroquine could rescue the loss of CTLA-4 in LRBA-deficient cells. Thus, LRBA appears to protect CTLA-4 from lysosomal degradation and therefore helps maintain intracellular pools of CTLA-4 for rapid mobilization to the cell surface for inhibitory function. Understanding the biochemical link between LRBA and CTLA-4 also explained why abatacept was such an effective treatment for the LRBA-deficient patients. By traveling through the blood to sites of T cell activation and providing a therapeutic cap on costimulation by blocking CD80 and CD86, abatacept could make up for the loss of endogenous CTLA-4 on the surface of conventional and regulatory T cells. Altogether, the discovery of CHAI and LATAIE, both autoimmune diseases resulting from CTLA-4 deficiency, emphasizes how critical this molecule is for immune tolerance and homeostasis, since reduction of CTLA-4 or loss of a single allele can throw the immune system out of balance. It also provides a vivid illustration of how research on HI genetic disorders can lead to a greater understanding of immunologic processes.

Uncovering the genetic and molecular etiology of CHAI and LATAIE disease provided novel insights into immune regulation as well as a rational basis for precision medicine treatments for the diseases. The lymphoproliferation and autoimmunity in CHAI and LATAIE diseases both stem from a loss of CTLA-4. Thus, treatment with abatacept as a CTLA-4 replacement therapy, clinically demonstrated to be effective in LATAIE disease, could also be effective for CHAI disease and perhaps other autoimmune diseases involving insufficient CTLA-4. Additionally, LRBA was found to regulate the turnover and degradation of CTLA-4 in lysosomes. Since the lysosomal inhibitor chloroquine could prevent lysosomal loss of CTLA-4 in LRBA-deficient cells, it or its pharmacologically safer alternative hydroxychloroquine could also be effective in treating CHAI or LATAIE diseases by boosting the levels of CTLA-4 protein. Interestingly, hydroxychloroquine is already used as a therapy for rheumatoid arthritis and lupus (117–119), so it is worth investigating whether its efficacy may be partly attributed to an augmentation of CTLA-4. If it were proven to be effective, hydroxychloroquine would be a very inexpensive alternative to abatacept. In summary, knowing not only the genes, but their biochemical and molecular roles in the cell, suggest molecular theories on which to base precision medicines that can have dramatic efficacy for previously undefined diseases.

PASLI disease

Multiple patients were identified who all suffered from an AD immunodeficiency syndrome with recurrent sinopulmonary infections, immunoglobulin synthesis defects, predisposition to EBV and/or CMV viremia, and lymphoproliferative disease (27, 28, 120–122). NGS of such patients revealed a first group of these individuals that carry mutations in the *PIK3CD* gene encoding the leukocyte-enriched p110δ subunit of phosphatidylinositol-4′,5′-bisphosphate 3′-kinase (PI3K) (27, 28, 120–122). In unrelated families, one of three germline, heterozygous mutations resulting in amino acid substitutions N334K, E525K, or E1021K, were identified. Intriguingly, the former two substitutions align precisely with hyperactivating N345K and E545K substitutions in p110α, resulting from somatic mutations in tumor cells (27, 28). Molecularly, these two changes have been shown to disrupt inhibitory contacts between the regulatory p85α protein and the catalytic p110 proteins of

the PI3K complex will Each of these mutations would be expected to augment enzymatic activity. This would explain the dominance of the heterozygous mutant allele. The recurrence of such mutations strongly suggest that they reveal amino acid vulnerabilities in enzyme activity regulation. We named this specific constellation of disease features associated with this genetic etiology p110δ-Activating mutations causing Senescence, Lymphadenopathy, and Immunodeficiency (PASLI, the disease has also been called Activated PI3K-δ Syndrome, or APDS) (27, 28). Since the initial descriptions, approximately 100–150 patients have been identified with *PIK3CD* mutations causing PASLI disease, and the initial list of 3 mutation sites has been expanded to mutations in 5 DNA sites that cause specific amino acid substitutions that increase enzyme activity. The prevalence of heterozygous, GOF mutations in *PIK3CD* suggests that p110δ-hyperactivating substitutions represent a significant contributor to disease within the population of immunodeficient lymphoproliferative individuals.

Shortly after discovery of heterozygous *PIK3CD* mutations, a second set of patients with a similar clinical phenotype, i.e., PASLI disease, were shown to harbor heterozygous splice site mutations in the ubiquitously expressed *PIK3R1* gene (39, 40). *PIK3R1* encodes the binding partner for p110δ called p85α that ensures stability, localization, and regulation of the p110δ protein. Notably, all the patients (approximately 30–40 individuals identified to date) shared an identical splice site mutation that results in skipping of exon 11 and production of an in-frame transcript encoding a protein lacking important residues in the inter-SH2 domain of p85α. Loss of these residues alters binding of the p85α regulatory subunit to the p110δ catalytic subunit such the former still guides stability and localization but cannot regulate p110δ enzyme activity, resulting in PI3K hyperactivation. Given the similar end result in molecular effect, i.e., hyperactivation of PI3K due to unrestrained p110δ enzyme activity, it is apparent why the clinical and cellular phenotypes within the immune system would be similar in patients with p110δ and p85α alterations. However, unlike p110δ, p85α expression is not restricted to the immune system. Despite ubiquitous p85α expression, these patients with *PIK3R1* mutations do not show dramatic non-immune phenotypes due to p110α or p110β hyperactivation. This remains an active area of investigation and may shed light on unique features of the association between p85α and p110δ (as opposed to p110α or p110β). Patients with *PIK3CD* or *PIK3R1* mutations can be referred to as PASLI-CD or PASLI-R1, respectively.

PI3Ks are a family of crucial signaling enzymes that transduce signals from tyrosine kinase receptors and G protein-coupled receptors by phosphorylating the hydroxyl group of the 3′ position of the inositol ring of phosphatidylinositol. This leads to the generation of phosphatidylinositol 3-phosphate (PI(3)P), phosphatidylinositol (3,4)-bisphosphate (PI(3,4)P2), and phosphatidylinositol (3,4,5)-trisphosphate (PI(3,4,5)P3). The last of these products, typically formed in the inner leaflet of the plasma membrane, serves as a docking site for phosphoinositide-binding protein domains such as the pleckstrin homology (PH) domain. There are two types of class I PI3K molecules that are activated by different stimuli, and these have the same heterodimeric structure comprising one 110 kD enzymatic subunit and a regulatory subunit that can vary in size. The class IA PI3K molecules, which include p110δ and p85α are activated by receptor tyrosine kinases to phosphorylate phosphatidylinositol (4,5)-bisphosphate (PI(4,5)P2) at the 3-position, generating the

PI(3,4,5)P (123). The generation of PI(3,4,5)P3 at the membrane initiates signaling by attracting the PH domain-containing proteins, including phosphoinositide-dependent kinase-1 (PDK1) and protein kinase B (PKB), also known as AKT. These stimulate further pathways downstream including the mammalian target of rapamycin (mTOR) kinase, which promotes cell growth, proliferation, and survival. Hence, activating mutations in class IA PI3K genes cause lymphoproliferation through the well-known proliferative mTOR pathway, but why this is associated with immunodeficiency loomed large as a key to understanding the pathogenesis and devising a new way to treat disease.

Investigation at the cellular level revealed that PI3K and mTOR hyperactivation in PASLI patients created imbalances in the subsets of T and B lymphocytes. On the T cell side, PASLI patients show a marked reduction in naïve T cells and a corresponding increase in effector-type T cell subsets that have become replicatively and functionally senescent with short telomeres and expression of the senescence marker CD57 on CD8+ T cells (28). Increased glucose uptake in the diseased T cells betrayed metabolic changes driven by PI3K/mTOR signaling that contributed to precocious effector cell maturation and led to most T cells converting into terminally differentiated cells that failed to function adequately (28, 40). On the B cell side, PASLI patients show defective B cell development with a preponderance of CD10+ transitional B cells and reduced frequency of CD27+ memory B cells. Functionally, patient B cells fail to properly secrete class switched immunoglobulins, particularly those specific to polysaccharide antigens. By histological analysis, PASLI patients have prominent germinal centers that lack a mantle zone (28). Thus, putting the cell surface signaling in overdrive caused the lymphocytes to differentiate abnormally.

The knowledge gained by the genetic and biochemical characterization of these patients immediately provided new therapeutic concepts. Fortunately, the mTOR kinase downstream of PI3K-AKT has been successfully targeted *in vivo* by the FDA-approved immunosuppressant rapamycin. The defects in PASLI disease provided a clear rationale for using PI3K/mTOR inhibitors to treat this disorder. Indeed, in an initial evaluation of the efficacy of rapamycin therapy in one patient, the data suggest that blocking this intrinsic drive toward effector differentiation at least partially restores the balance of naïve, memory, and effector CD8 T cells and markedly improves lymphadenopathy and splenomegaly (28). A larger cohort of patients treated with rapamycin are being studied (G. Uzel, personal communication). Thus, a counterintuitive therapeutic – treating an immunodeficiency with an immunosuppressant – becomes a rational targeted approach once the biochemical mechanism is clear.

Although rapamycin therapy holds great promise for an immediately useful treatment of PASLI disease, mTOR is only one pathway that is stimulated by activated AKT. The others, including NF-κB, GSK3, FOXO, and MDM2, will not be directly affected by rapamycin, so the effects of hyperactive PI3K would not be expected to be fully reversed. Consequently, the ideal therapy would be to directly and specifically target p110δ with an inhibitor to restrain activity to the normal range in PASLI patients. Specific targeting of PI3K isoforms has been the subject of intense research in the cancer field due to the high frequency with which this pathway is upregulated by direct mutation, gene amplification, or upstream receptor activation (124). Indeed, a p110δ-specific inhibitor called Idelalisib (formerly

CAL101 or GS1101) has shown promise as a therapy for chronic lymphocytic leukemia (CLL) and is now FDA approved for use in combination with rituximab for relapsed CLL (125, 126). We anticipate that directly inhibiting the source of augmented signaling in PASLI patients will maximize clinical benefit by enabling production of new B and T lymphocytes that escape the PI3K-driven development and differentiation defects.

Disease summary

The distinguishing characteristics of XMEN, PASLI, and CHAI/LATAIE disease are listed in Table 1 below.

Conclusions

With the dramatic drop in the cost of NGS, our data banks are bulging to the breaking point with DNA sequences. It is clear that bioinformatics, while enormously powerful, is not adequate on its own to accurately assess the medical significance of most new gene variants. This requires biochemical and molecular validation, which is slow, expensive, and labor intensive (Figure 2). However, our hope is that we have made the case in this essay that investing the time and resources for the appropriate biochemical investigation of gene variants yields important fundamental insights into immunology as well as new, targeted therapies for genetic diseases. We believe that technological advances provide an optimistic outlook for understanding the genetics of immune disorders. NGS and other molecular technologies have greatly expedited the identification of disease-causing variants. As of February 2015 in Online Mendelian Inheritance in Man, 2,937 genes have been associated with 4,163 Mendelian phenotypes, but the genes underlying ~50% of all known Mendelian phenotypes are still undiscovered and an untold number of additional Mendelian phenotypes have yet to be defined. We have found that an integrative approach to genomics with biochemical, molecular, and cellular examination of the phenotype defined in the clinic is most effective to understand pathways in the healthy and diseased immune system. We have provided examples of how causative gene identification in human disorders is already leading to unexpected new therapeutic interventions with less cost and morbidity than traditional HSCT. Another exciting frontier is the development of new technologies in genome engineering that could be used to directly correct the DNA lesion in patient cells. This will be more readily deployable in immunology and other hematopoietic disorders than for solid organ diseases. If the major effect of the gene defect is in cells of hematopoietic origin, then severe defects can be potentially corrected with genome editing of hematopoietic stem cells that could be used for autologous HSCT. If the defect is somewhat less severe, then alternative strategies such as correction of somatic cells, such as T cells, followed by re-infusion of the altered cells could be employed.

The ultimate success of clinical genomics will require further integration of knowledge at the molecular and clinical levels longitudinally. Also, well-defined genetic entities can help defining the prognosis of immunological diseases, the response to medical intervention, and appropriate genetic counseling. This will require organizational efforts on the clinical side including understanding the sociology of interaction with the medical delivery system, electronic medical records that can be integrated with genomics and medical research, and

incentivization of medical professionals to record phenotypic features according to the pathophysiology of the disease. At some point in the future, medical genetics may cease to be an independent specialty and evolve into an integrated part of medical practice.

Acknowledgments

References

1. McKusick VA. Genetics in medicine and medicine in genetics. Am J Med. 1963; 34:594–9. [PubMed: 13932120]

2. Bruton OC. Agammaglobulinemia. Pediatrics. 1952; 9:722–8. [PubMed: 14929630]

3. Tsukada S, Saffran DC, Rawlings DJ, Parolini O, Allen RC, Klisak I, Sparkes RS, Kubagawa H, Mohandas T, Quan S, et al. Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia. Cell. 1993; 72:279–90. [PubMed: 8425221]

4. Vetrie D, Vorechovsky I, Sideras P, Holland J, Davies A, Flinter F, Hammarstrom L, Kinnon C, Levinsky R, Bobrow M, et al. The gene involved in X-linked agammaglobulinaemia is a member of the src family of protein-tyrosine kinases. Nature. 1993; 361:226–33. [PubMed: 8380905]

5. Conley ME, Casanova JL. Discovery of single-gene inborn errors of immunity by next generation sequencing. Curr Opin Immunol. 2014; 30:17–23. [PubMed: 24886697]

6. Picard C, Fischer A. Contribution of high-throughput DNA sequencing to the study of primary immunodeficiencies. Eur J Immunol. 2014; 44:2854–61. [PubMed: 25154746]

7. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014; 346:1258096. [PubMed: 25430774]

8. Chun HJ, Zheng L, Ahmad M, Wang J, Speirs CK, Siegel RM, Dale JK, Puck J, Davis J, Hall CG, Skoda-Smith S, Atkinson TP, Straus SE, Lenardo MJ. Pleiotropic defects in lymphocyte activation caused by caspase-8 mutations lead to human immunodeficiency. Nature. 2002; 419:395–9. [PubMed: 12353035]

9. Kuehn HS, Ouyang W, Lo B, Deenick EK, Niemela JE, Avery DT, Schickel JN, Tran DQ, Stoddard J, Zhang Y, Frucht DM, Dumitriu B, Scheinberg P, Folio LR, Frein CA, Price S, Koh C, Heller T, Seroogy CM, Huttenlocher A, Rao VK, Su HC, Kleiner D, Notarangelo LD, Rampertaap Y, Olivier KN, McElwee J, Hughes J, Pittaluga S, Oliveira JB, Meffre E, Fleisher TA, Holland SM, Lenardo MJ, Tangye SG, Uzel G. Immune dysregulation in human subjects with heterozygous germline mutations in CTLA4. Science. 2014; 345:1623–7. [PubMed: 25213377]

10. Pennisi E. Human genome. Finally, the book of life and instructions for navigating it. Science. 2000; 288:2304–7. [PubMed: 10917823]

11. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. Science. 2002; 296:2225–9. [PubMed: 12029063]

12. Cho JH, Feldman M. Heterogeneity of autoimmune diseases: pathophysiologic insights from genetics and implications for new therapies. Nat Med. 2015; 21:730–8. [PubMed: 26121193]

13. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet. 2012; 90:7–24. [PubMed: 22243964]

14. Lupski JR, Belmont JW, Boerwinkle E, Gibbs RA. Clan genomics and the complex architecture of human disease. Cell. 2011; 147:32–43. [PubMed: 21962505]

15. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet. 2010; 11:415–25. [PubMed: 20479773]

16. Bittles AH, Black ML. Evolution in health and medicine Sackler colloquium: Consanguinity, human evolution, and complex diseases. Proc Natl Acad Sci U S A. 2010; 107(Suppl 1):1779–86. [PubMed: 19805052]

17. Cavalli-Sforza, LLMP., Piazza, A. The History and Geography of Human Genes. Princeton University Press; 1994.

18. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The next-generation sequencing revolution and its impact on genomics. Cell. 2013; 155:27–38. [PubMed: 24074859]

19. Weischenfeldt J, Symmons O, Spitz F, Korbel JO. Phenotypic impact of genomic structural variation: insights from and for human disease. Nat Rev Genet. 2013; 14:125–38. [PubMed: 23329113]

20. Michaelson JJ, Shi Y, Gujral M, Zheng H, Malhotra D, Jin X, Jian M, Liu G, Greer D, Bhandari A, Wu W, Corominas R, Peoples A, Koren A, Gore A, Kang S, Lin GN, Estabillo J, Gadomski T, Singh B, Zhang K, Akshoomoff N, Corsello C, McCarroll S, Iakoucheva LM, Li Y, Wang J, Sebat J. Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. Cell. 2012; 151:1431–42. [PubMed: 23260136]

21. Ponting CP. Loaded dice for human genome mutation. Cell. 2012; 151:1399–400. [PubMed: 23260131]

22. Veltman JA, Brunner HG. De novo mutations in human genetic disease. Nat Rev Genet. 2012; 13:565–75. [PubMed: 22805709]

23. Holzelova E, Vonarbourg C, Stolzenberg MC, Arkwright PD, Selz F, Prieur AM, Blanche S, Bartunkova J, Vilmer E, Fischer A, Le Deist F, Rieux-Laucat F. Autoimmune lymphoproliferative syndrome with somatic Fas mutations. N Engl J Med. 2004; 351:1409–18. [PubMed: 15459302]

24. Jing H, Zhang Q, Zhang Y, Hill BJ, Dove CG, Gelfand EW, Atkinson TP, Uzel G, Matthews HF, Mustillo PJ, Lewis DB, Kavadas FD, Hanson IC, Kumar AR, Geha RS, Douek DC, Holland SM, Freeman AF, Su HC. Somatic reversion in dedicator of cytokinesis 8 immunodeficiency modulates disease phenotype. J Allergy Clin Immunol. 2014; 133:1667–75. [PubMed: 24797421]

25. Al-Herz W, Bousfiha A, Casanova JL, Chatila T, Conley ME, Cunningham-Rundles C, Etzioni A, Franco JL, Gaspar HB, Holland SM, Klein C, Nonoyama S, Ochs HD, Oksenhendler E, Picard C, Puck JM, Sullivan K, Tang ML. Primary immunodeficiency diseases: an update on the classification from the international union of immunological societies expert committee for primary immunodeficiency. Front Immunol. 2014; 5:162. [PubMed: 24795713]

26. MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, Adams DR, Altman RB, Antonarakis SE, Ashley EA, Barrett JC, Biesecker LG, Conrad DF, Cooper GM, Cox NJ, Daly MJ, Gerstein MB, Goldstein DB, Hirschhorn JN, Leal SM, Pennacchio LA, Stamatoyannopoulos JA, Sunyaev SR, Valle D, Voight BF, Winckler W, Gunter C. Guidelines for investigating causality of sequence variants in human disease. Nature. 2014; 508:469–76. [PubMed: 24759409]

27. Angulo I, Vadas O, Garcon F, Banham-Hall E, Plagnol V, Leahy TR, Baxendale H, Coulter T, Curtis J, Wu C, Blake-Palmer K, Perisic O, Smyth D, Maes M, Fiddler C, Juss J, Cilliers D, Markelj G, Chandra A, Farmer G, Kielkowska A, Clark J, Kracker S, Debre M, Picard C, Pellier I, Jabado N, Morris JA, Barcenas-Morales G, Fischer A, Stephens L, Hawkins P, Barrett JC, Abinun M, Clatworthy M, Durandy A, Doffinger R, Chilvers ER, Cant AJ, Kumararatne D, Okkenhaug K, Williams RL, Condliffe A, Nejentsev S. Phosphoinositide 3-kinase delta gene mutation predisposes to respiratory infection and airway damage. Science. 2013; 342:866–71. [PubMed: 24136356]

28. Lucas CL, Kuehn HS, Zhao F, Niemela JE, Deenick EK, Palendira U, Avery DT, Moens L, Cannons JL, Biancalana M, Stoddard J, Ouyang W, Frucht DM, Rao VK, Atkinson TP, Agharahimi A, Hussey AA, Folio LR, Olivier KN, Fleisher TA, Pittaluga S, Holland SM, Cohen JI, Oliveira JB, Tangye SG, Schwartzberg PL, Lenardo MJ, Uzel G. Dominant-activating germline mutations in the gene encoding the PI(3)K catalytic subunit p110delta result in T cell senescence and human immunodeficiency. Nat Immunol. 2014; 15:88–97. [PubMed: 24165795]

29. MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, Jostins L, Habegger L, Pickrell JK, Montgomery SB, Albers CA, Zhang ZD, Conrad DF, Lunter G, Zheng H, Ayub Q, DePristo MA, Banks E, Hu M, Handsaker RE, Rosenfeld JA, Fromer M, Jin M, Mu XJ, Khurana E, Ye K, Kay M, Saunders GI, Suner MM, Hunt T, Barnes IH, Amid C, Carvalho-Silva

DR, Bignell AH, Snow C, Yngvadottir B, Bumpstead S, Cooper DN, Xue Y, Romero IG, Wang J, Li Y, Gibbs RA, McCarroll SA, Dermitzakis ET, Pritchard JK, Barrett JC, Harrow J, Hurles ME, Gerstein MB, Tyler-Smith C. Genomes Project C. A systematic survey of loss-of-function variants in human protein-coding genes. Science. 2012; 335:823–8. [PubMed: 22344438]

30. Xue Y, Chen Y, Ayub Q, Huang N, Ball EV, Mort M, Phillips AD, Shaw K, Stenson PD, Cooper DN, Tyler-Smith C. Genomes Project C. Deleterious- and disease-allele prevalence in healthy individuals: insights from current predictions, mutation databases, and population-scale resequencing. Am J Hum Genet. 2012; 91:1022–32. [PubMed: 23217326]

31. Ionita-Laza I, Makarov V, Yoon S, Raby B, Buxbaum J, Nicolae DL, Lin X. Finding disease variants in Mendelian disorders by using sequence data: methods and applications. Am J Hum Genet. 2011; 89:701–12. [PubMed: 22137099]

32. Li FY, Chaigne-Delalande B, Kanellopoulou C, Davis JC, Matthews HF, Douek DC, Cohen JI, Uzel G, Su HC, Lenardo MJ. Second messenger role for Mg2+ revealed by human T-cell immunodeficiency. Nature. 2011; 475:471–6. [PubMed: 21796205]

33. Chaigne-Delalande B, Li FY, O'Connor GM, Lukacs MJ, Jiang P, Zheng L, Shatzer A, Biancalana M, Pittaluga S, Matthews HF, Jancel TJ, Bleesing JJ, Marsh RA, Kuijpers TW, Nichols KE, Lucas CL, Nagpal S, Mehmet H, Su HC, Cohen JI, Uzel G, Lenardo MJ. Mg2+ regulates cytotoxic functions of NK and CD8 T cells in chronic EBV infection through NKG2D. Science. 2013; 341:186–91. [PubMed: 23846901]

34. Li FY, Chaigne-Delalande B, Su H, Uzel G, Matthews H, Lenardo MJ. XMEN disease: a new primary immunodeficiency affecting Mg2+ regulation of immunity against Epstein-Barr virus. Blood. 2014; 123:2148–52. [PubMed: 24550228]

35. Dobbs K, Dominguez Conde C, Zhang SY, Parolini S, Audry M, Chou J, Haapaniemi E, Keles S, Bilic I, Okada S, Massaad MJ, Rounioja S, Alwahadneh AM, Serwas NK, Capuder K, Ciftci E, Felgentreff K, Ohsumi TK, Pedergnana V, Boisson B, Haskologlu S, Ensari A, Schuster M, Moretta A, Itan Y, Patrizi O, Rozenberg F, Lebon P, Saarela J, Knip M, Petrovski S, Goldstein DB, Parrott RE, Savas B, Schambach A, Tabellini G, Bock C, Chatila TA, Comeau AM, Geha RS, Abel L, Buckley RH, Ikinciogullari A, Al-Herz W, Helminen M, Dogu F, Casanova JL, Boztug K, Notarangelo LD. Inherited DOCK2 Deficiency in Patients with Early-Onset Invasive Infections. N Engl J Med. 2015; 372:2409–22. [PubMed: 26083206]

36. Okada S, Markle JG, Deenick EK, Mele F, Averbuch D, Lagos M, Alzahrani M, Al-Muhsen S, Halwani R, Ma CS, Wong N, Soudais C, Henderson LA, Marzouqa H, Shamma J, Gonzalez M, Martinez-Barricarte R, Okada C, Avery DT, Latorre D, Deswarte C, Jabot-Hanin F, Torrado E, Fountain J, Belkadi A, Itan Y, Boisson B, Migaud M, Arlehamn CS, Sette A, Breton S, McCluskey J, Rossjohn J, de Villartay JP, Moshous D, Hambleton S, Latour S, Arkwright PD, Picard C, Lantz O, Engelhard D, Kobayashi M, Abel L, Cooper AM, Notarangelo LD, Boisson-Dupuis S, Puel A, Sallusto F, Bustamante J, Tangye SG, Casanova JL. IMMUNODEFICIENCIES. Impairment of immunity to Candida and Mycobacterium in humans with bi-allelic RORC mutations. Science. 2015; 349:606–13. [PubMed: 26160376]

37. Zhang Q, Davis JC, Lamborn IT, Freeman AF, Jing H, Favreau AJ, Matthews HF, Davis J, Turner ML, Uzel G, Holland SM, Su HC. Combined immunodeficiency associated with DOCK8 mutations. N Engl J Med. 2009; 361:2046–55. [PubMed: 19776401]

38. Zhang Y, Yu X, Ichikawa M, Lyons JJ, Datta S, Lamborn IT, Jing H, Kim ES, Biancalana M, Wolfe LA, DiMaggio T, Matthews HF, Kranick SM, Stone KD, Holland SM, Reich DS, Hughes JD, Mehmet H, McElwee J, Freeman AF, Freeze HH, Su HC, Milner JD. Autosomal recessive phosphoglucomutase 3 (PGM3) mutations link glycosylation defects to atopy, immune deficiency, autoimmunity, and neurocognitive impairment. J Allergy Clin Immunol. 2014; 133:1400–9. 9e1–5. [PubMed: 24589341]

39. Deau MC, Heurtier L, Frange P, Suarez F, Bole-Feysot C, Nitschke P, Cavazzana M, Picard C, Durandy A, Fischer A, Kracker S. A human immunodeficiency caused by mutations in the PIK3R1 gene. J Clin Invest. 2014; 124:3923–8. [PubMed: 25133428]

40. Lucas CL, Zhang Y, Venida A, Wang Y, Hughes J, McElwee J, Butrick M, Matthews H, Price S, Biancalana M, Wang X, Richards M, Pozos T, Barlan I, Ozen A, Rao VK, Su HC, Lenardo MJ. Heterozygous splice mutation in PIK3R1 causes human immunodeficiency with

lymphoproliferation due to dominant activation of PI3K. J Exp Med. 2014; 211:2537–47. [PubMed: 25488983]

41. Price S, Shaw PA, Seitz A, Joshi G, Davis J, Niemela JE, Perkins K, Hornung RL, Folio L, Rosenberg PS, Puck JM, Hsu AP, Lo B, Pittaluga S, Jaffe ES, Fleisher TA, Rao VK, Lenardo MJ. Natural history of autoimmune lymphoproliferative syndrome associated with FAS gene mutations. Blood. 2014; 123:1989–99. [PubMed: 24398331]

42. Siegel RM, Frederiksen JK, Zacharias DA, Chan FK, Johnson M, Lynch D, Tsien RY, Lenardo MJ. Fas preassociation required for apoptosis signaling and dominant inhibition by pathogenic mutations. Science. 2000; 288:2354–7. [PubMed: 10875918]

43. Oftedal BE, Hellesen A, Erichsen MM, Bratland E, Vardi A, Perheentupa J, Kemp EH, Fiskerstrand T, Viken MK, Weetman AP, Fleishman SJ, Banka S, Newman WG, Sewell WA, Sozaeva LS, Zayats T, Haugarvoll K, Orlova EM, Haavik J, Johansson S, Knappskog PM, Lovas K, Wolff AS, Abramson J, Husebye ES. Dominant Mutations in the Autoimmune Regulator AIRE Are Associated with Common Organ-Specific Autoimmune Diseases. Immunity. 2015; 42:1185–96. [PubMed: 26084028]

44. Boisson B, Quartier P, Casanova JL. Immunological loss-of-function due to genetic gain-of-function in humans: autosomal dominance of the third kind. Curr Opin Immunol. 2015; 32:90–105. [PubMed: 25645939]

45. Rieux-Laucat F, Casanova JL. Immunology. Autoimmunity by haploinsufficiency. Science. 2014; 345:1560–1. [PubMed: 25258064]

46. Schubert D, Bode C, Kenefeck R, Hou TZ, Wing JB, Kennedy A, Bulashevska A, Petersen BS, Schaffer AA, Gruning BA, Unger S, Frede N, Baumann U, Witte T, Schmidt RE, Dueckers G, Niehues T, Seneviratne S, Kanariou M, Speckmann C, Ehl S, Rensing-Ehl A, Warnatz K, Rakhmanov M, Thimme R, Hasselblatt P, Emmerich F, Cathomen T, Backofen R, Fisch P, Seidl M, May A, Schmitt-Graeff A, Ikemizu S, Salzer U, Franke A, Sakaguchi S, Walker LS, Sansom DM, Grimbacher B. Autosomal dominant immune dysregulation syndrome in humans with CTLA4 mutations. Nat Med. 2014; 20:1410–6. [PubMed: 25329329]

47. Bennett CL, Christie J, Ramsdell F, Brunkow ME, Ferguson PJ, Whitesell L, Kelly TE, Saulsbury FT, Chance PF, Ochs HD. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. Nat Genet. 2001; 27:20–1. [PubMed: 11137993]

48. Noguchi M, Yi H, Rosenblatt HM, Filipovich AH, Adelstein S, Modi WS, McBride OW, Leonard WJ. Interleukin-2 receptor gamma chain mutation results in X-linked severe combined immunodeficiency in humans. Cell. 1993; 73:147–57. [PubMed: 8462096]

49. Acuna-Hidalgo R, Bo T, Kwint MP, van de Vorst M, Pinelli M, Veltman JA, Hoischen A, Vissers LE, Gilissen C. Post-zygotic Point Mutations Are an Underrecognized Source of De Novo Genomic Variation. Am J Hum Genet. 2015; 97:67–74. [PubMed: 26054435]

50. de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T, Vulto-van Silfhout AT, Koolen DA, de Vries P, Gilissen C, del Rosario M, Hoischen A, Scheffer H, de Vries BB, Brunner HG, Veltman JA, Vissers LE. Diagnostic exome sequencing in persons with severe intellectual disability. N Engl J Med. 2012; 367:1921–9. [PubMed: 23033978]

51. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet. 2011; 12:745–55. [PubMed: 21946919]

52. Cooper DN, Krawczak M, Polychronakos C, Tyler-Smith C, Kehrer-Sawatzki H. Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease. Hum Genet. 2013; 132:1077–130. [PubMed: 23820649]

53. Shawky R. Reduced penetrance in human inherited disease. Egypt J Med Hum Genet. 2014; 15:103–11.

54. Vu V, Verster AJ, Schertzberg M, Chuluunbaatar T, Spensley M, Pajkic D, Hart GT, Moffat J, Fraser AG. Natural Variation in Gene Expression Modulates the Severity of Mutant Phenotypes. Cell. 2015; 162:391–402. [PubMed: 26186192]

55. Zlotogora J. Penetrance and expressivity in the molecular age. Genet Med. 2003; 5:347–52. [PubMed: 14501829]

56. Bustamante J, Boisson-Dupuis S, Abel L, Casanova JL. Mendelian susceptibility to mycobacterial disease: genetic, immunological, and clinical features of inborn errors of IFN-gamma immunity. Semin Immunol. 2014; 26:454–70. [PubMed: 25453225]

57. Ciancanelli MJ, Huang SX, Luthra P, Garner H, Itan Y, Volpi S, Lafaille FG, Trouillet C, Schmolke M, Albrecht RA, Israelsson E, Lim HK, Casadio M, Hermesh T, Lorenzo L, Leung LW, Pedergnana V, Boisson B, Okada S, Picard C, Ringuier B, Troussier F, Chaussabel D, Abel L, Pellier I, Notarangelo LD, Garcia-Sastre A, Basler CF, Geissmann F, Zhang SY, Snoeck HW, Casanova JL. Infectious disease. Life-threatening influenza and impaired interferon amplification in human IRF7 deficiency. Science. 2015; 348:448–53. [PubMed: 25814066]

58. Gilissen C, Hoischen A, Brunner HG, Veltman JA. Disease gene identification strategies for exome sequencing. Eur J Hum Genet. 2012; 20:490–7. [PubMed: 22258526]

59. Knecht C, Krawczak M. Molecular genetic epidemiology of human diseases: from patterns to predictions. Hum Genet. 2014; 133:425–30. [PubMed: 24241280]

60. Ng SB, Nickerson DA, Bamshad MJ, Shendure J. Massively parallel sequencing and rare disease. Hum Mol Genet. 2010; 19:R119–24. [PubMed: 20846941]

61. Hsu AP, Johnson KD, Falcone EL, Sanalkumar R, Sanchez L, Hickstein DD, Cuellar-Rodriguez J, Lemieux JE, Zerbe CS, Bresnick EH, Holland SM. GATA2 haploinsufficiency caused by mutations in a conserved intronic element leads to MonoMAC syndrome. Blood. 2013; 121:3830–7. S1–7. [PubMed: 23502222]

62. Liadaki K, Sun J, Hammarstrom L, Pan-Hammarstrom Q. New facets of antibody deficiencies. Curr Opin Immunol. 2013; 25:629–38. [PubMed: 24012250]

63. Lougaris VTG, Baronio M, Vitali M, Plebani A. The genetic heterogeneity of common variable immunodeficiency (CVID): an update. J Vaccines Vaccin. 2014; 5

64. Frousios K, Iliopoulos CS, Schlitt T, Simpson MA. Predicting the functional consequences of non-synonymous DNA sequence variants--evaluation of bioinformatics tools and development of a consensus strategy. Genomics. 2013; 102:223–8. [PubMed: 23831115]

65. Sengupta M, Sarkar D, Ganguly K, Sengupta D, Bhaskar S, Ray K. In silico analyses of missense mutations in coagulation factor VIII: identification of severity determinants of haemophilia A. Haemophilia. 2015

66. van der Velde KJ, Kuiper J, Thompson BA, Plazzer JP, van Valkenhoef G, de Haan M, Jongbloed JD, Wijmenga C, de Koning TJ, Abbott KM, Sinke R, Spurdle AB, Macrae F, Genuardi M, Sijmons RH, Swertz MA, InSi GHTG. Evaluation of CADD Scores in Curated Mismatch Repair Gene Variants Yields a Model for Clinical Validation and Prioritization. Hum Mutat. 2015; 36:712–9. [PubMed: 25871441]

67. Wu J, Li Y, Jiang R. Integrating multiple genomic data to predict disease-causing nonsynonymous single nucleotide variants in exome sequencing studies. PLoS Genet. 2014; 10:e1004237. [PubMed: 24651380]

68. Casanova JL, Conley ME, Seligman SJ, Abel L, Notarangelo LD. Guidelines for genetic studies in single patients: lessons from primary immunodeficiencies. J Exp Med. 2014; 211:2137–49. [PubMed: 25311508]

69. Zhang YSH, Lenardo M. Genomics is rapidly advancing precision medicine treatments for immune disorders. Nat Immunol. 2015 In Press.

70. Mohr SE, Smith JA, Shamu CE, Neumuller RA, Perrimon N. RNAi screening comes of age: improved techniques and complementary approaches. Nat Rev Mol Cell Biol. 2014; 15:591–600. [PubMed: 25145850]

71. Oliveira JB, Bidere N, Niemela JE, Zheng L, Sakai K, Nix CP, Danner RL, Barb J, Munson PJ, Puck JM, Dale J, Straus SE, Fleisher TA, Lenardo MJ. NRAS mutation causes a human autoimmune lymphoproliferative syndrome. Proc Natl Acad Sci U S A. 2007; 104:8953–8. [PubMed: 17517660]

72. Plotkin JB. Transcriptional regulation is only half the story. Mol Syst Biol. 2010; 6:406. [PubMed: 20739928]

73. Chakravarti A, Clark AG, Mootha VK. Distilling pathophysiology from complex disease genetics. Cell. 2013; 155:21–6. [PubMed: 24074858]

74. Candotti F, O'Shea JJ, Villa A. Severe combined immune deficiencies due to defects of the common gamma chain-JAK3 signaling pathway. Springer Semin Immunopathol. 1998; 19:401–15. [PubMed: 9618765]

75. Miller JN, Pearce DA. Nonsense-mediated decay in genetic disease: friend or foe? Mutat Res Rev Mutat Res. 2014; 762:52–64. [PubMed: 25485595]

76. Kousi M, Katsanis N. Genetic modifiers and oligogenic inheritance. Cold Spring Harb Perspect Med. 2015:5.

77. Girard SL, Dion PA, Bourassa CV, Geoffroy S, Lachance-Touchette P, Barhdadi A, Langlois M, Joober R, Krebs MO, Dube MP, Rouleau GA. Mutation burden of rare variants in schizophrenia candidate genes. PLoS One. 2015; 10:e0128988. [PubMed: 26039597]

78. Parikshak NN, Gandal MJ, Geschwind DH. Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. Nat Rev Genet. 2015; 16:441–58. [PubMed: 26149713]

79. Miller, JHPSE. Complex adaptive systems: An introduction to computational models of social life. Princeton University Press; 2007.

80. Holland, JH. Signals and boundaries: building blocks for complex adaptive systems. MIT Press; 2014.

81. Gursky VS, SY, Samsonova MG. Mechanisms of developmental robustness. Biosystems. 2012; 109:329–35. [PubMed: 22687821]

82. Naujokat C, Fuchs D, Berges C. Adaptive modification and flexibility of the proteasome system in response to proteasome inhibition. Biochim Biophys Acta. 2007; 1773:1389–97. [PubMed: 17582523]

83. Rossi A, Kontarakis Z, Gerri C, Nolte H, Holper S, Kruger M, Stainier DY. Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature. 2015; 524:230–3. [PubMed: 26168398]

84. Milner JD, Brenchley JM, Laurence A, Freeman AF, Hill BJ, Elias KM, Kanno Y, Spalding C, Elloumi HZ, Paulson ML, Davis J, Hsu A, Asher AI, O'Shea J, Holland SM, Paul WE, Douek DC. Impaired T(H)17 cell differentiation in subjects with autosomal dominant hyper-IgE syndrome. Nature. 2008; 452:773–6. [PubMed: 18337720]

85. Flanagan SE, Haapaniemi E, Russell MA, Caswell R, Lango Allen H, De Franco E, McDonald TJ, Rajala H, Ramelius A, Barton J, Heiskanen K, Heiskanen-Kosma T, Kajosaari M, Murphy NP, Milenkovic T, Seppanen M, Lernmark A, Mustjoki S, Otonkoski T, Kere J, Morgan NG, Ellard S, Hattersley AT. Activating germline mutations in STAT3 cause early-onset multi-organ autoimmune disease. Nat Genet. 2014; 46:812–4. [PubMed: 25038750]

86. Haapaniemi EM, Kaustio M, Rajala HL, van Adrichem AJ, Kainulainen L, Glumoff V, Doffinger R, Kuusanmaki H, Heiskanen-Kosma T, Trotta L, Chiang S, Kulmala P, Eldfors S, Katainen R, Siitonen S, Karjalainen-Lindsberg ML, Kovanen PE, Otonkoski T, Porkka K, Heiskanen K, Hanninen A, Bryceson YT, Uusitalo-Seppala R, Saarela J, Seppanen M, Mustjoki S, Kere J. Autoimmunity, hypogammaglobulinemia, lymphoproliferation, and mycobacterial disease in patients with activating mutations in STAT3. Blood. 2015; 125:639–48. [PubMed: 25349174]

87. Milner JD, Vogel TP, Forbes L, Ma CA, Stray-Pedersen A, Niemela JE, Lyons JJ, Engelhardt KR, Zhang Y, Topcagic N, Roberson ED, Matthews H, Verbsky JW, Dasu T, Vargas-Hernandez A, Varghese N, McClain KL, Karam LB, Nahmod K, Makedonas G, Mace EM, Sorte HS, Perminow G, Rao VK, O'Connell MP, Price S, Su HC, Butrick M, McElwee J, Hughes JD, Willet J, Swan D, Xu Y, Santibanez-Koref M, Slowik V, Dinwiddie DL, Ciaccio CE, Saunders CJ, Septer S, Kingsmore SF, White AJ, Cant AJ, Hambleton S, Cooper MA. Early-onset lymphoproliferation and autoimmunity caused by germline STAT3 gain-of-function mutations. Blood. 2015; 125:591–9. [PubMed: 25359994]

88. Boisson-Dupuis S, Kong XF, Okada S, Cypowyj S, Puel A, Abel L, Casanova JL. Inborn errors of human STAT1: allelic heterogeneity governs the diversity of immunological and infectious phenotypes. Curr Opin Immunol. 2012; 24:364–78. [PubMed: 22651901]

89. Emison ES, Garcia-Barcelo M, Grice EA, Lantieri F, Amiel J, Burzynski G, Fernandez RM, Hao L, Kashuk C, West K, Miao X, Tam PK, Griseri P, Ceccherini I, Pelet A, Jannot AS, de Pontual L, Henrion-Caude A, Lyonnet S, Verheij JB, Hofstra RM, Antinolo G, Borrego S, McCallion AS, Chakravarti A. Differential contributions of rare and common, coding and noncoding Ret

mutations to multifactorial Hirschsprung disease liability. Am J Hum Genet. 2010; 87:60–74. [PubMed: 20598273]

90. Brown MS, Goldstein JL. Koch's postulates for cholesterol. Cell. 1992; 71:187–8. [PubMed: 1423585]

91. Katsanis SH, Katsanis N. Molecular genetic testing and the future of clinical genomics. Nat Rev Genet. 2013; 14:415–26. [PubMed: 23681062]

92. Rizzo JM, Buck MJ. Key principles and clinical applications of "next-generation" DNA sequencing. Cancer Prev Res (Phila). 2012; 5:887–900. [PubMed: 22617168]

93. Javitt GH, Carner KS. Regulation of next generation sequencing. J Law Med Ethics. 2014; 42(Suppl 1):9–21.

94. Tabor HK, Auer PL, Jamal SM, Chong JX, Yu JH, Gordon AS, Graubert TA, O'Donnell CJ, Rich SS, Nickerson DA, Project NES, Bamshad MJ. Pathogenic variants for Mendelian and complex traits in exomes of 6,517 European and African Americans: implications for the return of incidental results. Am J Hum Genet. 2014; 95:183–93. [PubMed: 25087612]

95. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, McGuire AL, Nussbaum RL, O'Daniel JM, Ormond KE, Rehm HL, Watson MS, Williams MS, Biesecker LG. American College of Medical G, Genomics. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med. 2013; 15:565–74. [PubMed: 23788249]

96. Burke W, Antommaria AH, Bennett R, Botkin J, Clayton EW, Henderson GE, Holm IA, Jarvik GP, Khoury MJ, Knoppers BM, Press NA, Ross LF, Rothstein MA, Saal H, Uhlmann WR, Wilfond B, Wolf SM, Zimmern R. Recommendations for returning genomic incidental findings? We need to talk! Genet Med. 2013; 15:854–9. [PubMed: 23907645]

97. Gargis AS, Kalman L, Bick DP, da Silva C, Dimmock DP, Funke BH, Gowrisankar S, Hegde MR, Kulkarni S, Mason CE, Nagarajan R, Voelkerding KV, Worthey EA, Aziz N, Barnes J, Bennett SF, Bisht H, Church DM, Dimitrova Z, Gargis SR, Hafez N, Hambuch T, Hyland FC, Luna RA, MacCannell D, Mann T, McCluskey MR, McDaniel TK, Ganova-Raeva LM, Rehm HL, Reid J, Campo DS, Resnick RB, Ridge PG, Salit ML, Skums P, Wong LJ, Zehnbauer BA, Zook JM, Lubin IM. Good laboratory practice for clinical next-generation sequencing informatics pipelines. Nat Biotechnol. 2015; 33:689–93. [PubMed: 26154004]

98. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med. 2015; 372:793–5. [PubMed: 25635347]

99. Sassi A, Lazaroski S, Wu G, Haslam SM, Fliegauf M, Mellouli F, Patiroglu T, Unal E, Ozdemir MA, Jouhadi Z, Khadir K, Ben-Khemis L, Ben-Ali M, Ben-Mustapha I, Borchani L, Pfeifer D, Jakob T, Khemiri M, Asplund AC, Gustafsson MO, Lundin KE, Falk-Sorqvist E, Moens LN, Gungor HE, Engelhardt KR, Dziadzio M, Stauss H, Fleckenstein B, Meier R, Prayitno K, Maul-Pavicic A, Schaffer S, Rakhmanov M, Henneke P, Kraus H, Eibel H, Kolsch U, Nadifi S, Nilsson M, Bejaoui M, Schaffer AA, Smith CI, Dell A, Barbouche MR, Grimbacher B. Hypomorphic homozygous mutations in phosphoglucomutase 3 (PGM3) impair immunity and increase serum IgE levels. J Allergy Clin Immunol. 2014; 133:1410–9. 9e1–13. [PubMed: 24698316]

100. Stray-Pedersen A, Backe PH, Sorte HS, Morkrid L, Chokshi NY, Erichsen HC, Gambin T, Elgstoen KB, Bjoras M, Wlodarski MW, Kruger M, Jhangiani SN, Muzny DM, Patel A, Raymond KM, Sasa GS, Krance RA, Martinez CA, Abraham SM, Speckmann C, Ehl S, Hall P, Forbes LR, Merckoll E, Westvik J, Nishimura G, Rustad CF, Abrahamsen TG, Ronnestad A, Osnes LT, Egeland T, Rodningen OK, Beck CR, Boerwinkle EA, Gibbs RA, Lupski JR, Orange JS, Lausch E, Hanson IC. Baylor-Johns Hopkins Center for Mendelian G. PGM3 mutations cause a congenital disorder of glycosylation with severe immunodeficiency and skeletal dysplasia. Am J Hum Genet. 2014; 95:96–107. [PubMed: 24931394]

101. Maguire ME, Cowan JA. Magnesium chemistry and biochemistry. Biometals. 2002; 15:203–10. [PubMed: 12206387]

102. Alangari A, Alsultan A, Adly N, Massaad MJ, Kiani IS, Aljebreen A, Raddaoui E, Almomen AK, Al-Muhsen S, Geha RS, Alkuraya FS. LPS-responsive beige-like anchor (LRBA) gene mutation in a family with inflammatory bowel disease and combined immunodeficiency. J Allergy Clin Immunol. 2012; 130:481–8. e2. [PubMed: 22721650]

103. Burns SO, Zenner HL, Plagnol V, Curtis J, Mok K, Eisenhut M, Kumararatne D, Doffinger R, Thrasher AJ, Nejentsev S. LRBA gene deletion in a patient presenting with autoimmunity

without hypogammaglobulinemia. J Allergy Clin Immunol. 2012; 130:1428–32. [PubMed: 22981790]

104. Lopez-Herrera G, Tampella G, Pan-Hammarstrom Q, Herholz P, Trujillo-Vargas CM, Phadwal K, Simon AK, Moutschen M, Etzioni A, Mory A, Srugo I, Melamed D, Hultenby K, Liu C, Baronio M, Vitali M, Philippet P, Dideberg V, Aghamohammadi A, Rezaei N, Enright V, Du L, Salzer U, Eibel H, Pfeifer D, Veelken H, Stauss H, Lougaris V, Plebani A, Gertz EM, Schaffer AA, Hammarstrom L, Grimbacher B. Deleterious mutations in LRBA are associated with a syndrome of immune deficiency and autoimmunity. Am J Hum Genet. 2012; 90:986–1001. [PubMed: 22608502]

105. Tivol EA, Borriello F, Schweitzer AN, Lynch WP, Bluestone JA, Sharpe AH. Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA-4. Immunity. 1995; 3:541–7. [PubMed: 7584144]

106. Waterhouse P, Penninger JM, Timms E, Wakeham A, Shahinian A, Lee KP, Thompson CB, Griesser H, Mak TW. Lymphoproliferative disorders with early lethality in mice deficient in Ctla-4. Science. 1995; 270:985–8. [PubMed: 7481803]

107. Walker LS, Sansom DM. The emerging role of CTLA4 as a cell-extrinsic regulator of T cell responses. Nat Rev Immunol. 2011; 11:852–63. [PubMed: 22116087]

108. Wing K, Yamaguchi T, Sakaguchi S. Cell-autonomous and -non-autonomous roles of CTLA-4 in immune regulation. Trends Immunol. 2011; 32:428–33. [PubMed: 21723783]

109. Warnatz K, Denz A, Drager R, Braun M, Groth C, Wolff-Vorbeck G, Eibel H, Schlesier M, Peter HH. Severe deficiency of switched memory B cells (CD27(+)IgM(−)IgD(−)) in subgroups of patients with common variable immunodeficiency: a new approach to classify a heterogeneous disease. Blood. 2002; 99:1544–51. [PubMed: 11861266]

110. Isnardi I, Ng YS, Menard L, Meyers G, Saadoun D, Srdanovic I, Samuels J, Berman J, Buckner JH, Cunningham-Rundles C, Meffre E. Complement receptor 2/CD21− human naive B cells contain mostly autoreactive unresponsive clones. Blood. 2010; 115:5026–36. [PubMed: 20231422]

111. Lo B, Zhang K, Lu W, Zheng L, Zhang Q, Kanellopoulou C, Zhang Y, Liu Z, Fritz JM, Marsh R, Husami A, Kissell D, Nortman S, Chaturvedi V, Haines H, Young LR, Mo J, Filipovich AH, Bleesing JJ, Mustillo P, Stephens M, Rueda CM, Chougnet CA, Hoebe K, McElwee J, Hughes JD, Karakoc-Aydiner E, Matthews HF, Price S, Su HC, Rao VK, Lenardo MJ, Jordan MB. Patients with LRBA deficiency show CTLA4 loss and immune dysregulation responsive to abatacept therapy. Science. 2015; 349:436–40. [PubMed: 26206937]

112. Wang JW, Howson J, Haller E, Kerr WG. Identification of a novel lipopolysaccharide-inducible gene with key features of both A kinase anchor proteins and chs1/beige proteins. J Immunol. 2001; 166:4586–95. [PubMed: 11254716]

113. Gebauer D, Li J, Jogl G, Shen Y, Myszka DG, Tong L. Crystal structure of the PH-BEACH domains of human LRBA/BGL. Biochemistry. 2004; 43:14873–80. [PubMed: 15554694]

114. de Souza N, Vallier LG, Fares H, Greenwald I. SEL-2, the C. elegans neurobeachin/LRBA homolog, is a negative regulator of lin-12/Notch activity and affects endosomal traffic in polarized epithelial cells. Development. 2007; 134:691–702. [PubMed: 17215302]

115. Qureshi OS, Kaur S, Hou TZ, Jeffery LE, Poulter NS, Briggs Z, Kenefeck R, Willox AK, Royle SJ, Rappoport JZ, Sansom DM. Constitutive clathrin-mediated endocytosis of CTLA-4 persists during T cell activation. J Biol Chem. 2012; 287:9429–40. [PubMed: 22262842]

116. Mead KI, Zheng Y, Manzotti CN, Perry LC, Liu MK, Burke F, Powner DJ, Wakelam MJ, Sansom DM. Exocytosis of CTLA-4 is dependent on phospholipase D and ADP ribosylation factor-1 and stimulated during activation of regulatory T cells. J Immunol. 2005; 174:4803–11. [PubMed: 15814706]

117. A randomized trial of hydroxychloroquine in early rheumatoid arthritis: the HERA Study. Am J Med. 1995; 98:156–68. [PubMed: 7847432]

118. Clegg DO, Dietz F, Duffy J, Willkens RF, Hurd E, Germain BF, Wall B, Wallace DJ, Bell CL, Sleckman J. Safety and efficacy of hydroxychloroquine as maintenance therapy for rheumatoid arthritis after combination therapy with methotrexate and hydroxychloroquine. J Rheumatol. 1997; 24:1896–902. [PubMed: 9330929]

119. Costedoat-Chalumeau N, Dunogue B, Morel N, Le Guern V, Guettrot-Imbert G. Hydroxychloroquine: a multifaceted treatment in lupus. Presse Med. 2014; 43:e167–80. [PubMed: 24855048]

120. Crank MC, Grossman JK, Moir S, Pittaluga S, Buckner CM, Kardava L, Agharahimi A, Meuwissen H, Stoddard J, Niemela J, Kuehn H, Rosenzweig SD. Mutations in PIK3CD can cause hyper IgM syndrome (HIGM) associated with increased cancer susceptibility. J Clin Immunol. 2014; 34:272–6. [PubMed: 24610295]

121. Hartman HN, Niemela J, Hintermeyer MK, Garofalo M, Stoddard J, Verbsky JW, Rosenzweig SD, Routes JM. Gain of Function Mutations of PIK3CD as a Cause of Primary Sclerosing Cholangitis. J Clin Immunol. 2014

122. Kracker S, Curtis J, Ibrahim MA, Sediva A, Salisbury J, Campr V, Debre M, Edgar JD, Imai K, Picard C, Casanova JL, Fischer A, Nejentsev S, Durandy A. Occurrence of B-cell lymphomas in patients with activated phosphoinositide 3-kinase delta syndrome. J Allergy Clin Immunol. 2014; 134:233–6. [PubMed: 24698326]

123. Cantley LC. The phosphoinositide 3-kinase pathway. Science. 2002; 296:1655–7. [PubMed: 12040186]

124. Yuan TL, Cantley LC. PI3K pathway alterations in cancer: variations on a theme. Oncogene. 2008; 27:5497–510. [PubMed: 18794884]

125. Furman RR, Sharman JP, Coutre SE, Cheson BD, Pagel JM, Hillmen P, Barrientos JC, Zelenetz AD, Kipps TJ, Flinn I, Ghia P, Eradat H, Ervin T, Lamanna N, Coiffier B, Pettitt AR, Ma S, Stilgenbauer S, Cramer P, Aiello M, Johnson DM, Miller LL, Li D, Jahn TM, Dansey RD, Hallek M, O'Brien SM. Idelalisib and rituximab in relapsed chronic lymphocytic leukemia. N Engl J Med. 2014; 370:997–1007. [PubMed: 24450857]

126. Shah A, Mangaonkar A. Idelalisib: A Novel PI3Kdelta Inhibitor for Chronic Lymphocytic Leukemia. Ann Pharmacother. 2015
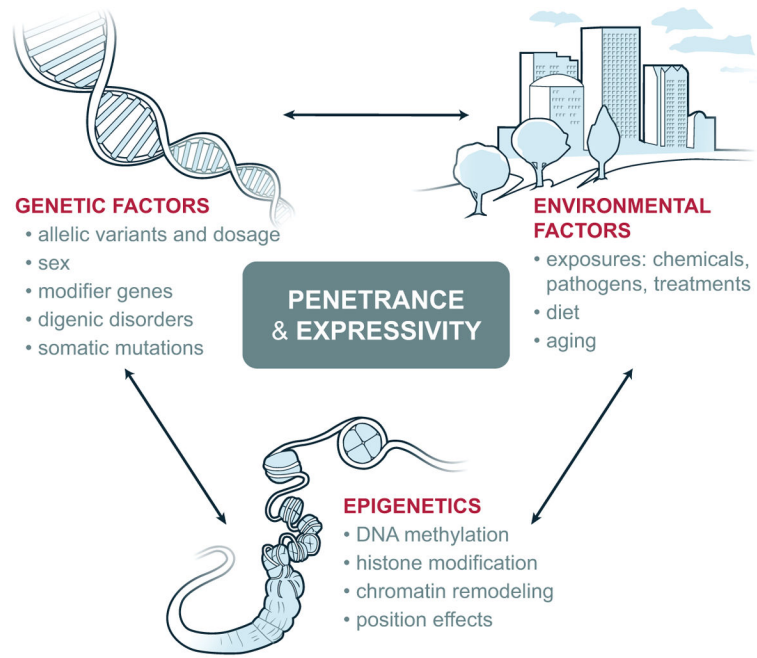
Figure 1.
The causes of variable penetrance and expressivity of disease gene mutations.
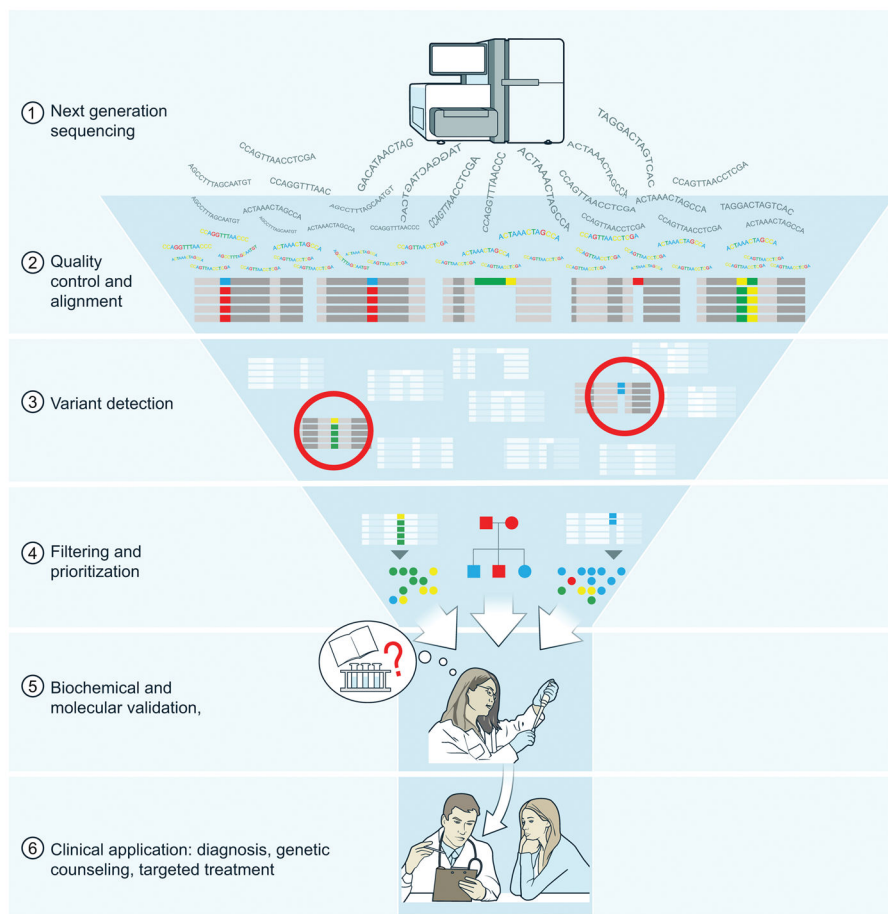
Figure 2.
The bottleneck to genetic disease discovery is the biochemical and molecular validation of candidate disease variants obtained through next-generation DNA sequencing. Although computational methods can reduce and prioritize variants, further validation through direct experiments is required. The benefits of clinical application can only pertain once validation has been carried out.

Table 1

	CVID	ALPS	CAEBV	XMEN	PASLI	CHAI/LATAIE
Sinopulmonary infections	Yes				Yes, sometimes with bronchiectasis	Yes, if hypogamm
Ig defects	Yes				High IgM	Often hypogamm
Lymphoproliferative disease		Yes			Yes	Yes
Viremia	Variable	No	EBV	EBV	Often CMV and/or EBV	
EBV+ B cell lymphoma risk	No	No	Yes	Yes	Yes	
Lymphocyte tissue infiltration	No	No	No	No	**Yes** Mucosal nodules, sometimes obstructive	Tissue infiltration
Expanded cell population	Variable	CD3+CD4/8−(B220+)	CD3+CD8+	CD3+CD8+ (low CD4:CD8) CD8+NKG2D−	CD3+CD8+ **CD3+CCR7− CD8+CD57+ CD20+CD10+**	More memory T cells and CD21^{lo} B cells
Disease gene product	Variable	FAS, FASL, caspase 10, N-Ras	Variable	MAGT1	p110δ or p85α	CTLA4 or LRBA