Mini Review

A review on the impact of genetics and genome wide association studies in autoimmunity

Abstract

Disorders in immune system cause auto immune diseases in human and it affects almost 10% of population. Mendelian syndromes of auto immunity have an impact on monogenic mutations and it results in lack of tolerance in central and peripheral system. Most of the auto immune disorders arise from the complex interaction between polygenic risk component and an environmental component. Risk factor can contribute to individual changes in nucleotides. Genome wide association studies (GWAS) can provide the list of biological pathways that drives the auto immune diseases in humans. Interpretation of GWAS requires the integration of multiple datasets in genomics. Improvised level of understanding in auto immunity with respect to genetics will lead to the development of novel diagnostics and targeted therapies.

Keywords: autoimmune diseases, monogenic mutations, genome wide association studies

Harishchander Anandaram
Department of Bioinformatics, Sathyabama University, India

Correspondence: Department of Bioinformatics, Sathyabama University, Chennai, India, Email harishchander.a@gmail.com

Abbreviations: AID, autoimmune diseases, GWAS, genome wide association studies

Introduction

Autoimmune diseases (AIDs) in humans are considered as major issues in public health. AIDs affect almost 10% of human population.[1] AIDs in human are disorders with complexity and they develop from the interactions between polygenic risk and environmental factors.[2] Investigations of genetics in AIDs contain the potential to have an unbiased view of etiologies in common conditions to identify novel targets for therapy. In the era of pre genomics, the understanding of heritability in disease were derived based on the high prevalence rate of autoimmune disease in twins[3] or family members[4] compared with non-biological relatives in a shared environment.[5] Apart from the recognition of risk factors influenced by genetics in autoimmune disease, there exists a challenge to identify the causal nucleotide variants and correlate their functional effects. The sequencing of human genome and rapidly emerging technologies in genomics helps us to retrieve the genetic variants that contribute to the risk factors of autoimmune disease. Understandings of genetics in human AIDs were expanded with a high rate of prevalence in the last 15years. In this review, we analyze the biological lessons learnt from the genetic studies of AIDs in humans.

In rare cases, monogenic syndromes in autoimmune disease revealed a high rate of penetrance in mutations to disrupt the mechanisms which are essential for the tolerance of central and peripheral systems in immunity. Genome-wide association studies (GWAS) provided an insight to analyze the dysregulation in immune system caused by the genetic variants that contribute risk of autoimmunity. The challenge here involves providing frameworks for the interpretation of GWAS with the development of new technologies in genome sequencing. Studies in human genetics of autoimmune diseases involves the integration of multiple dataset in genomics, it includes genotyping, epigenomic annotation and quantitative studies of gene expression in the relevant conditions.

Mongenic syndrome (Immune dysregulation)

Syndromes in AIDs provided the insight into biological pathways

for the maintenance of homeostasis in immune.[6] The prevalence of polygenic autoimmune diseases is increased in patients with Mendelian immune dysregulation syndromes caused by monogenic mutations were identified. The linkage analysis and positional cloning were successful and they were combined with the mouse genetic models to identify causal mutations. Patients with polyendocrinopathy, immune dysregulation, enteropathy and X-linked (IPEX) syndrome highlighted the vital role of FOXP3 in Treg.[7] Patients with autoimmune polyendocrinopathy, candidiasis-ectodermal dystrophy (APECED) also known as autoimmune polyendocrine syndrome type 1 revealed the vital role of autoimmune regulator (AIRE) in central tolerance and thymic selection.[8]

Mutations in the encoding gene *CTLA4* have recently been discovered in families with Mendelian multi-organ autoimmune disease syndromes.[9] These rare mutations (including missense, nonsense, and splice variants) cause severe disease in heterozygous patients. The mutations were associated with the impairment of Treg suppressive function. Genetic variants in *CTLA4* were associated with more the risk of autoimmune diseases including type 1 diabetes (T1D) and Graves' disease.[10] The functional consequences of a risk allele (SNPrs3087243 located in the 3′-UTR of the *CTLA4*; A/G) were investigated using phosphorylated mAbs to target the components of TCR signaling in naive and memory T cells.

The relative responsiveness to TCR stimulation (assessed by phosphorylation levels of downstream signaling molecules) was altered in naïve (CD4+CD45RAhi) and memory (CD4+CD45RAlo) T cells obtained from individuals with an allele with disease susceptibility at *CTLA4*. In the earlier reports that allelic variation associated with autoimmune disease can alter the signaling threshold of CD4+ T cells. These findings were consistent with a spectrum of allelic variants at the same locus were involved in causing the differing degrees of dysregulation in immune system and contributes to a risk factor in AIDs.

GWAS

Identification and characterization of variants in autoimmune diseases is challenging because of they are relatively small to contribute to risk factors. Variants are not purified by negative

selection, as a result of the modest biological effects were compared with the disruption of genes in Mendelian diseases. A few loci, like *MHC* on chromosome 6p21[11] and the *NOD2* locus[12] were initially linked to autoimmunity due to their relatively high odds ratios. However, identification of a larger set of loci that contribute to the complex nature of risk factors in autoimmune disease requires the development of new tools to address genetics.

GWAS studies were designed to detect variants that confer a modest risk factor in common diseases as opposed to rare diseases caused by a catalog of highly penetrant mutations.[13] The approach here is to identify the genetic variants that are preferentially associated with patients with a disease or trait which are related to the healthy individuals. GWAS using SNPs from the International HapMap Project (http://hapmap.ncbi.nlm.nih.gov/) have allowed us to view an unbiased approach in scanning the whole genome and identifying a disease-associated region, with a success point in identifying the association of a large number of loci with human AIDs.

Interpretation of GWAS

Though the SNPs were inherited in germ line, they could affect genes that are expressed in a particular cell type. Indeed, the GWAS loci that were associated with autoimmune diseases were enriched in genes that are preferentially expressed in a particular subset of immune system.[14] A comparison of disease loci with a gene expression in 223 cell types of the ImmGen dataset revealed the fact that SLE loci encodes genes that are preferentially expressed in the subset of B cell, whereas the loci that are RA-associated was enriched in genes and they are preferentially expressed in the CD4+ effector memory in T cells. Further, the knowledge of these pathogenic cell types may suggest the biologically relevant genes that were affected by the GWAS variants that do not meet strict thresholds in genome-wide significance. Subsequent analysis of human RNAseq data had confirmed their enrichment of genes in loci that are associated with the immune-mediated diseases that are preferentially expressed in the subset of CD4+ T cells.[15]

GWAS loci can also encode proteins that participate in a disproportionate number of physical interactions that form the biological pathways.[14] In addition, integration of multiple tools for annotation may help us in extracting enriched set of genes from loci that linked to the disease phenotypes. Collectively, a triage of candidate genes with linkage disequilibrium blocks pathways were affected by the disease variants. The discovery of such pathways will highlight the biology of the complex nature in autoimmune diseases to direct the process and efforts of drug discovery. Protein interaction on the basis of pathway analysis can suggest the candidate genes within GWAS loci. The implication of key pathogenic pathways includes the annotation of pathways for JAK/STAT signaling and TCR signaling.[15]

Methodology

In case of Genome Wide Association studies, the initial step is to identify the case-control to compare the two large groups of individuals with one healthy control group and one affected by a disease. All individuals in each group are genotyped with known polymorphisms of single nucleotide (SNPs). Then the odds ratio of two groups was calculated with the allele frequency. Additionally, a P-value is calculated to determine the significance of the odds ratio by chi-squared test.[16]

Conclusion

Genetic studies on syndromes of monogenic immune dysregulation had informed us about the understanding of essential mechanisms involved in immune tolerance. Novel technologies were in genomics developed over the past decade can enable systematic studies on risk factors in complex diseases of autoimmune nature. GWAS implicated a hundreds of relevant loci to explore autoimmune disease. One approach was to understand the pathological consequences of variation in GWAS involve the phenotypic analysis of cells from a cohort of a genotyped individual. Apart from individual approaches, systems-scale approaches allowed the progress from GWAS to pathogenic pathways, transcriptional circuits and types of immune cell in specific cellular conditions.

Future prospective

Improved level of analytic and experimental techniques raised the hope for clinical applications in genetics of autoimmune diseases. A refined level of understanding of variants contributing to autoimmune disease pathology could aid in the development of novel diagnostic tests for the identification of disease biomarkers. Insights in genetics will lead to the identification of novel pathogenic circuits to focus for the process of drug discovery in future by tailoring the specific epigenetic and biochemical pathways that are dysregulated in human autoimmune diseases.

Acknowledgements

None.

Conflict of interest

The author declares that there is no conflict of interest regarding the publication of this article.

References

1. Cheng MH, Anderson MS. Monogenic autoimmunity. *Annu Rev Immunol*. 2012;30:393–427.

2. Schubert D, Bode C, Kenefeck R, et al. Autosomal dominant immune dysregulation syndrome in humans with CTLA4 mutations. *Nat Med*. 2014;20(12):1410–1416.

3. Zeissig S, Petersen BS, Tomczak M, et al. Early–onset Crohn's disease autoimmunity associated with a variant in CTLA–4. *Gut*. 2015;64(12):1889–1897.

4. Kuehn HS, Ouyang W, Lo B, et al. Immune dysregulation in human subjects with heterozygous germline mutations in CTLA4. *Science*. 2014;345(6204):1623–1627.

5. Chu X, Pan CM, Zhao SX, et al. A genome–wide association study identifies two new risk loci for Graves' disease. *Nat Genet*. 2011;43(9):897–901.

6. Onengut–Gumuscu S, Chen WM, Burren O, et al. Fine mapping of type 1 diabetes susceptibility loci evidence for colocalization of causal variants with lymphoid gene enhancers. *Nat Genet*. 2015;47(4):381–386.

7. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP–trait associations. *Nucleic Acids Res*. 2014;42(Database issue):D1001–D1006.

8. Cotsapas C, Voight BF, Rossin E, et al. Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet*. 2011;7(8):e1002254.

9. Cotsapas C, Hafler DA. Immune–mediated disease genetics: the shared basis of pathogenesis. *Trends Immunol*. 2013;34(1):22–26.

10. Farh KK, Marson A, Zhu J, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature*. 2015;518(7539):337–343.

11. Hu X, Kim H, Stahl E, et al. Integrating autoimmune risk loci with gene–expression data identifies specific pathogenic immune cell subsets. *Am J Hum Genet*. 2011;89(4):496–506.

12. Zhang W, Ferguson J, Ng SM, et al. Effector CD4+ T cell expression signatures and immune–mediated disease associated genes. *PLoS One*. 2012;7(6):e38510.

13. Rossin EJ, Lage K, Raychaudhuri S, et al. Proteins encoded in genomic regions associated with immune–mediated disease physically interact and suggest underlying biology. *PLoS Genet*. 2011;7(1):e1001273.

14. Segre AV, DIAGRAM Consortium, MAGIC investigators, et al. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet*. 2010;6(8):e1001058.

15. International Multiple Sclerosis Genetics Consortium. Network–based multiple sclerosis pathway analysis with GWAS data from 15,000 cases and 30,000 controls. *Am J Hum Genet*. 2013;92(6):854–865.

16. Clarke GM, Anderson CA, Pettersson FH, et al. Basic statistical analysis in genetic case–control studies. *Nat Protoc*. 2011;6(2):121–133.