Genomic solutions for cell biology and complex disease research

An overview of recent publications featuring Illumina® technology
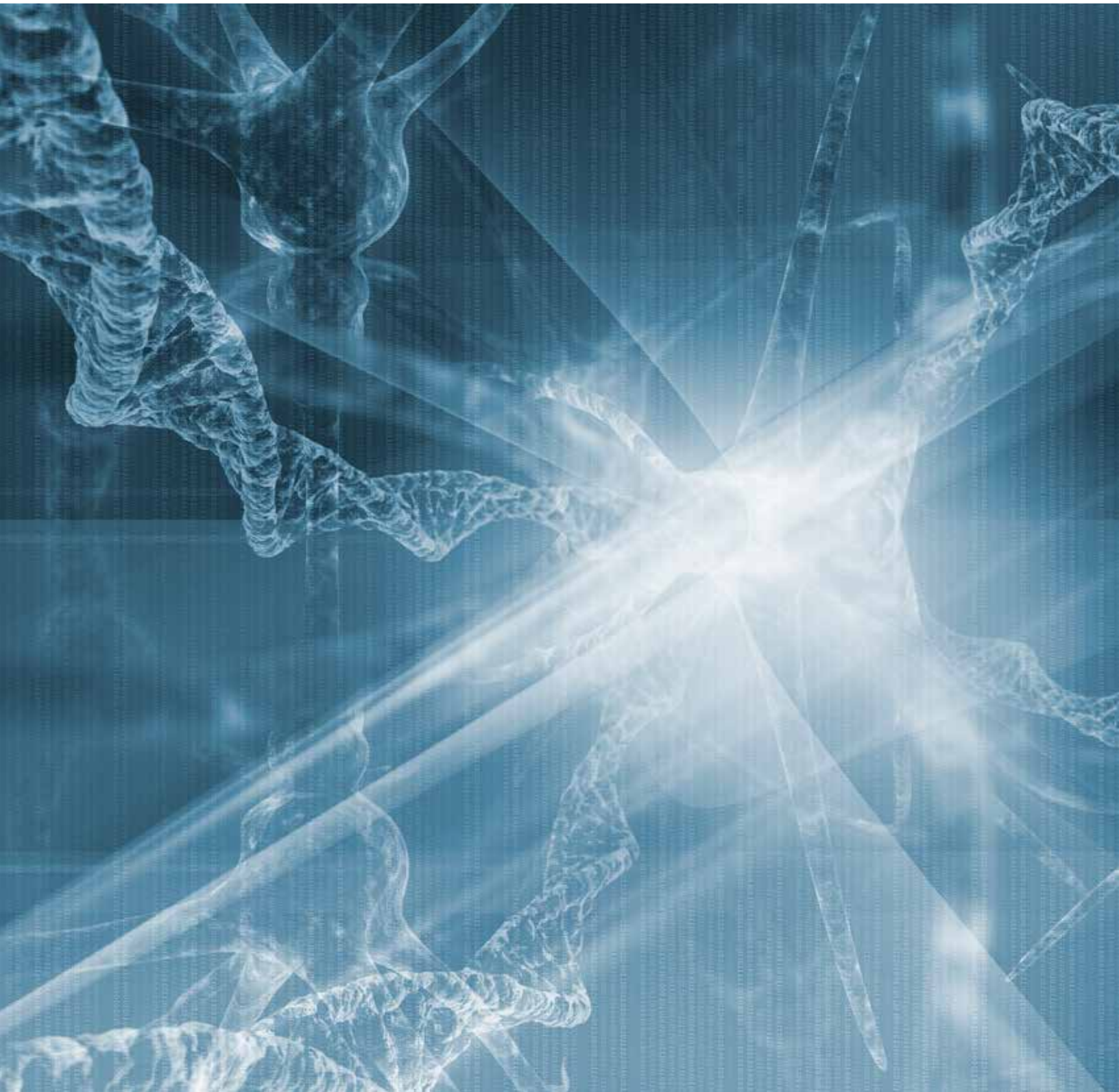
TABLE OF CONTENTS

This document highlights recent publications that demonstrate the use of Illumina technologies in immunology research. To learn more about the platforms and assays cited, visit www.illumina.com.

INTRODUCTION

Complex diseases are the result of additive or epistatic interactions between genetic and environmental factors. They are distinguished from Mendelian traits (or simple traits) as they do not follow a specific model of inheritance and are not the result of a single mutated gene.[1]

Some of these diseases are highly heritable; however, the currently known genetic variants can explain only some of the estimated heritability.[2] In the past, various hypotheses have been proposed for the genomic basis of complex traits. Many studies assumed that multiple genetic factors influence the phenotype.[3, 4] Others consider each complex disease as heterogeneous samples where disease is caused by rare variants.[5]

Compared to Mendelian diseases, complex diseases are usually more frequent in the population.[6] Examples include autoimmune and rheumatic diseases, atherosclerosis and many forms of heart disease, neurological and psychiatric disorders, and cancer. Table 1 lists examples of complex diseases, with a particular focus on autoimmune, neurological, and psychiatric diseases and disorders.

1. Lander E. S. and Schork N. J. (1994) Genetic dissection of complex traits. Science 265: 2037-2048
2. Manolio T. A., Collins F. S., Cox N. J., Goldstein D. B., Hindorff L. A., et al. (2009) Finding the missing heritability of complex diseases. Nature 461: 747-753
3. Schizophrenia Psychiatric Genome-Wide Association Study C. (2011) Genome-wide association study identifies five new schizophrenia loci. Nat Genet 43: 969-976
4. Schizophrenia Working Group of the Psychiatric Genomics C. (2014) Biological insights from 108 schizophrenia-associated genetic loci. Nature 511: 421-427
5. consortium C. (2015) Sparse whole-genome sequencing identifies two loci for major depressive disorder. Nature 523: 588-591
6. Revill P. (2007) Genomics of common diseases. Drug News Perspect 20: 475-479

Table 1. Complex Phenotypes of an Autoimmune, Neurological, or Psychiatric Nature

Disease	Nature
Rheumatoid arthritis	Autoimmune
Systemic lupus erythematosus	Autoimmune
Crohn's disease	Autoimmune
Ulcerative colitis	Autoimmune
Psoriasis	Autoimmune
Celiac disease	Autoimmune
Type I diabetes	Autoimmune
Allergy	Autoimmune
Asthma	Autoimmune
Hay fever	Autoimmune
Sjogren's syndrome	Autoimmune
Graves' disease	Autoimmune
Hashimoto's disease	Autoimmune
Vitiligo	Autoimmune
Scleroderma	Autoimmune
Pernicious anemia	Autoimmune
Multiple sclerosis	Autoimmune/Neurological
Frontotemporal dementia	Neurological
Parkinson's disease	Neurological
Alzheimer's disease	Neurological
Amyotrophic lateral sclerosis	Neurological
Autism-spectrum disorders	Neurological
Dementia	Neurological
Dementia with Lewy bodies	Neurological
Narcolepsy	Neurological

DNA Polymorphisms and Genetic Variation

It is estimated that humans share 99.5% of their DNA sequence with one another.[7] Given the estimated human genome size of 3200 Mb, this means that about 320 Mb of genes are polymorphic.[8] These polymorphisms, or DNA variations, can involve base substitutions, insertions, deletions, or repeats. Generally, they are distinguished by sequence and length polymorphisms (Figure 1). Single-nucleotide polymorphisms (SNPs) and copy-number variants (CNVs) are important sources of individual variation, and they play an important role in the susceptibility to complex diseases.[9, 10, 11, 12]

Most polymorphisms are neutral and have no phenotypic effect, but some are at the basis of what differentiates individuals.[13] The functional effect of any given DNA polymorphism on a phenotype can range from full penetrance (as in Mendelian disorders) to none (neutrality). For complex or multifactorial traits, the penetrance may be intermediate to low.[14]

Technological advances now allow us to isolate DNA (or RNA) from multiple sample types, amplify and sequence regions of the genome, and sequence whole genomes. These methods enable the analysis of multiple polymorphisms in large samples of individuals.
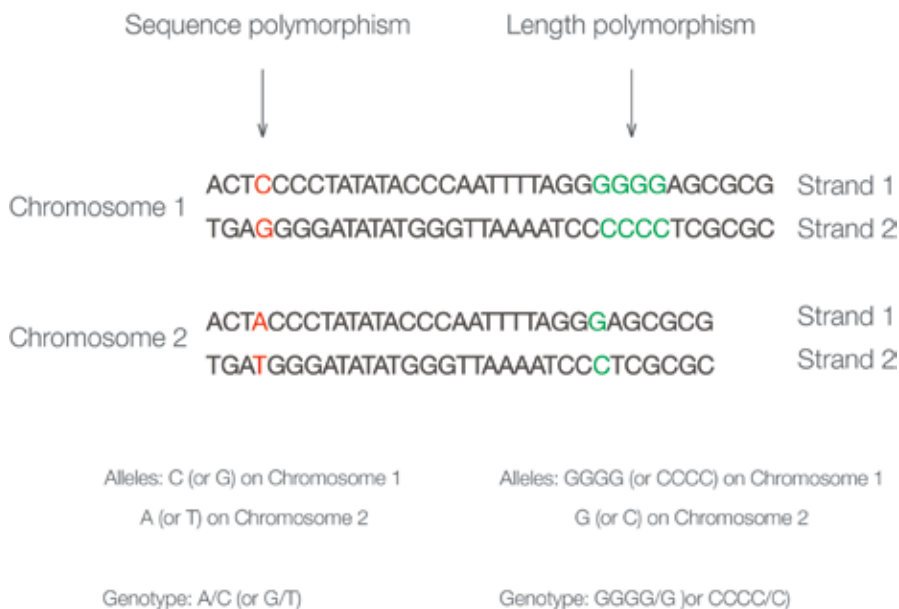
7.   FAQ about Genetic and Genomic Science. National Human Genome Research Institute website. http://www.genome.gov/19016904. Updated March 2, 2016. Accessed August 15, 2016.

8.   Human Genome Project Completion: Frequently Asked Questions. National Human Genome Research Institute website. https://www.genome.gov/11006943. Updated October 30, 2010. Accessed August 15, 2016

9.   Girirajan S., Campbell C. D. and Eichler E. E. (2011) Human copy number variation and complex genetic disease. Annu Rev Genet 45: 203-226

10.   Need A. C., Ge D., Weale M. E., Maia J., Feng S., et al. (2009) A genome-wide investigation of SNPs and CNVs in schizophrenia. PLoS Genet 5: e1000373

11.   Do C. B., Hinds D. A., Francke U. and Eriksson N. (2012) Comparison of family history and SNPs for predicting risk of complex disease. PLoS Genet 8: e1002973

12.   McCarroll S. A. and Altshuler D. M. (2007) Copy-number variation and association studies of human disease. Nat Genet 39: S37-42

13.   Kimura M. (1991) The neutral theory of molecular evolution: a review of recent evidence. Jpn J Genet 66: 367-386

14.   Cooper D. N., Krawczak M., Polychronakos C., Tyler-Smith C. and Kehrer-Sawatzki H. (2013) Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease. Hum Genet 132: 1077-1130

Figure 1: Genetic sequences, strands, and genetic variability.

Genetic Structure and Linkage Disequilibrium

In the study of monogenic diseases, data are collected from affected families to identify regions of the genome that cosegregate with the disease in many independent families and/or over multiple generations. These methods use linkage mapping sets, with interspersed polymorphisms across the genome as representative markers. This method is called linkage analysis. It has identified mutations successfully for several Mendelian diseases, such as Huntington's chorea.[15, 16]

Given the sporadic nature of common complex diseases (influence of the environment, lack of a transmission model, and low penetrance), they are not suitable for linkage analysis. Rather than looking at single affected families, the study of these traits is usually done on populations, comparing affected cases with individual controls (population-based case-control association studies).

The design of association studies is based on the concept of linkage disequilibrium (LD). LD is the nonrandom association between 2 alleles, and it is determined by 2 factors:

- The distance between the 2 loci considered
- The recombination rate in the region included between the 2 loci

15. Bates G. P. (2005) History of genetic disease: the molecular genetics of Huntington disease - a history. Nat Rev Genet 6: 766-773
16. Gusella J. F., Wexler N. S., Conneally P. M., Naylor S. L., Anderson M. A., et al. (1983) A polymorphic DNA marker genetically linked to Huntington's disease. Nature 306: 234-238
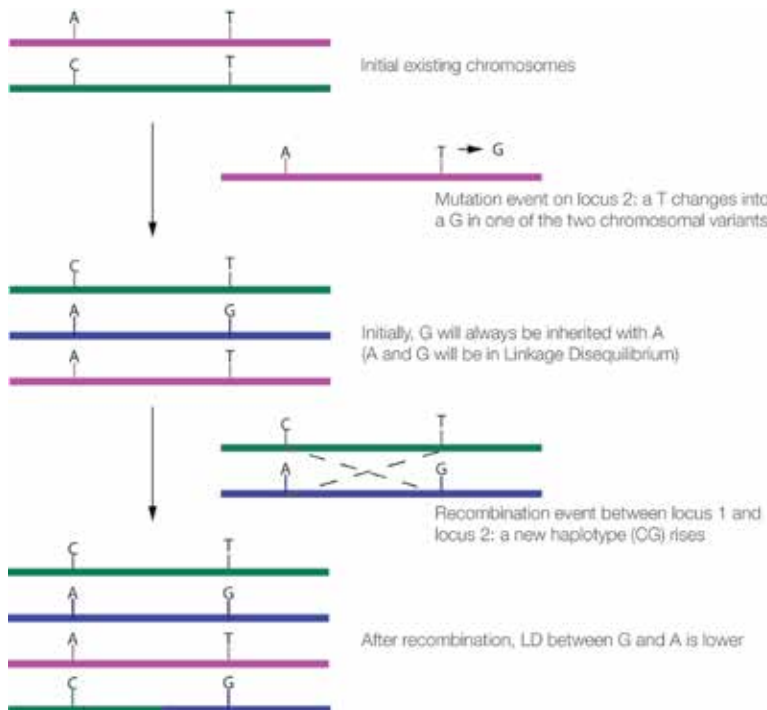
Figure 2. The concept of LD and the rise of new haplotypes.

Sekar A., Bialas A. R., de Rivera H., Davis A., Hammond T. R., et al. (2016) Schizophrenia risk from complex variation of complement component 4. Nature.

Schizophrenia is a heritable illness that involves loss of gray matter. Although highly studied, the pathogenic mechanisms that give rise to this disease are still unknown. Previous studies have demonstrated that the strongest genetic association with schizophrenia is shown with the major histocompatibility complex (MHC) locus. This study investigated the genes and mechanisms involved in this association. The authors performed a combination of genotyping and expression analyses on DNA and RNA extracted from brains of schizophrenic patients and controls. They demonstrated that the association between the MHC locus and schizophrenia is partly due to structurally diverse alleles of the component 4 (*C4*) genes. These alleles were associated with levels of expression of *C4A* and *C4B* isoforms in the brain. All of the alleles associated with schizophrenia were associated with a higher expression of *C4A*. They observed that *C4A* in mouse mediated postnatal synapse elimination. These results are suggestive of an excessive complement activity in the development of schizophrenia.

Illumina Technology: Infinium®

Franke B., Stein J. L., Ripke S., Anttila V., Hibar D. P., et al. (2016) Genetic influences on schizophrenia and subcortical brain volumes: large-scale proof of concept. Nat Neurosci.

Schizophrenia is a heritable and devastating disease. Schizophrenia patients differ in brain structure and function from normal individuals. As genetic studies provide information on genetic variants associated with schizophrenia and brain imaging phenotypes, the authors used these data to check for genetic overlap. They integrated the results from common variant studies of schizophrenia (33,636 cases and 43,008 controls) and volumes of brain structures (mainly subcortical). For the available data, the authors were not able to find a genetic overlap between schizophrenia risk and subcortical volume measurements.

Ryu S., Atzmon G., Barzilai N., Raghavachari N. and Suh Y. (2016) Genetic landscape of APOE in human longevity revealed by high-throughput sequencing. Mech Ageing Dev.

The apolipoprotein E (*APOE*) gene has been associated consistently with several phenotypes, such as longevity. In long-lived individuals, 2 common haplotypes are either significantly depleted (ε4 allele) or enriched (ε2 allele) compared to controls. In this study, the authors applied a high-throughput sequencing approach called Pool-Seq on the exons and the 2 kb proximal promoter of *APOE* in 450 centenarians and 500 controls of Ashkenazi Jewish descent. They found common regulatory variants (rs405509, P = 0.006; rs769449, P = 0.036) that were significantly depleted in centenarians. Additionally, they observed variants that showed significant enrichment of the ε2 allele (rs7412, P = 0.003), and significant depletion of ε3/ε4 (rs429358, P = 0.005) in centenarians.

Illumina Technology: Illumina® GA$_{IIx}$

For Research Use Only. Not for use in diagnostic procedures.

GENETIC MECHANISMS

From Genotype to Phenotype: General Considerations

Genomic technologies have allowed sequencing the complete genome of many organisms. However, the connection between DNA polymorphisms and phenotypic effects is still challenging to characterize.[17] This issue is particularly true for complex traits, partly due to the modifications that can occur between a DNA sequence and the manifestation of a phenotype. For example, splice variants and RNA editing can occur during gene expression, and additional modifications can be added by posttranslational modification (PTM) (Figure 3).

The complex mapping of genotypes to phenotypes is called the genotype-to-phenotype problem[18] and cannot be addressed by a single approach alone. This section and the next will review the different approaches that are applied to the study of genotype-phenotype conversions. These approaches include:

- Genetic mechanisms and sources of variation
- The use of human, animal, and cellular models

Systems genetics is an approach that uses a combination of experimental and statistical methods to integrate data from multiple studies focusing on different molecular levels. This method aims at investigating how variants identified in genome-wide association studies (GWAS) affect each phase of the etiology of the disease. It can also be used to describe the architecture of a phenotype and identifying genes, gene networks, and pathways that underlie a complex trait.[19, 20]

17. Benfey P. N. and Mitchell-Olds T. (2008) From genotype to phenotype: systems biology meets natural variation. Science 320: 495-497

18. Dowell R. D., Ryan O., Jansen A., Cheung D., Agarwala S., et al. (2010) Genotype to phenotype: a complex problem. Science 328: 469

19. van der Sijde M. R., Ng A. and Fu J. (2014) Systems genetics: From GWAS to disease pathways. Biochim Biophys Acta 1842: 1903-1909

20. Civelek M. and Lusis A. J. (2014) Systems genetics approaches to understand complex traits. Nat Rev Genet 15: 34-48
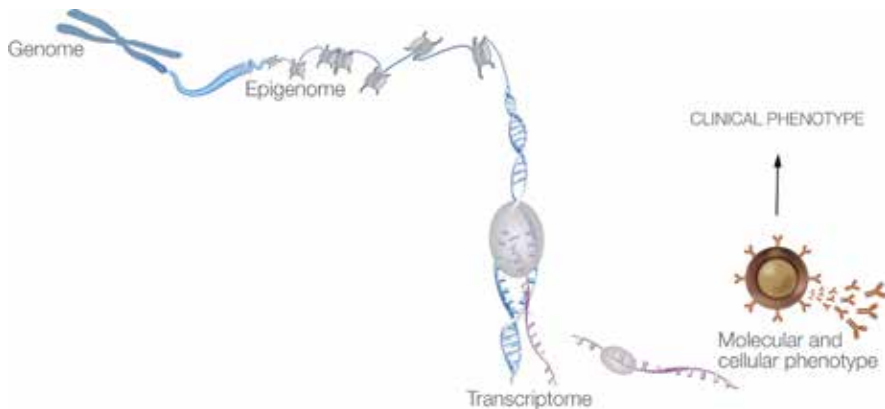
Figure 3. Multiple steps separate a genomic sequence and the manifestation of a clinical phenotype. There are different levels of possible sources of variability.

For Research Use Only. Not for use in diagnostic procedures.

8   Genomic solutions for cell biology and complex disease research

Genetic Associations

Genetic association studies are designed to test the association between a genetic factor and a phenotype. Phenotypes can either be:

- Quantitative, or continuous, when they can be measured. Examples include height, weight, age, expression levels, plaque size, survival length, and blood pressure. Most complex diseases are quantitative traits.[21, 22]
- Discrete, or discontinuous, when only their presence or absence can be determined. Examples include the presence of a disease or risk factor. Discontinuous traits are often associated with Mendelian diseases.[23]

Association studies usually consist of a statistically significant sample size of individuals. They measure the phenotype of interest as well as the genetic variants being tested. Risk factors and potential confounders are taken into account in the analysis. For quantitative traits, the strength of the association between the genetic factor and the phenotype is measured through linear models.[24] For discrete binary traits, cases carrying the trait are compared to a matching sample of controls not carrying the trait. This design is called a case-control association study. It uses methods such as a chi-squared test for a bivariate analysis or logistic regression for a multivariate analysis[25] (Figure 4).

Genetic association studies used to consider a single candidate gene in the past (candidate gene association study). Today, most studies are performed at the genomic level (GWAS). GWAS can rapidly scan markers that are representative of the genome in large datasets to find genetic variations associated with a disease or phenotype. Since the mid-2000s, GWAS have reported considerable success in the identification of risk variants for multiple phenotypes.[26, 27, 28, 29]

21. Plomin R., Haworth C. M. and Davis O. S. (2009) Common disorders are quantitative traits. Nat Rev Genet 10: 872-878

22. McCarthy M. I., Abecasis G. R., Cardon L. R., Goldstein D. B., Little J., et al. (2008) Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 9: 356-369

23. Kendler K. S. (2015) A joint history of the nature of genetic variation and the nature of schizophrenia. Mol Psychiatry 20: 77-83

24. Fan R., Wang Y., Mills J. L., Wilson A. F., Bailey-Wilson J. E., et al. (2013) Functional linear models for association analysis of quantitative traits. Genet Epidemiol 37: 726-742

25. Thomas D. C. (2009) Genome-wide association studies for discrete traits. Genet Epidemiol 33 Suppl 1: S8-12

26. Genome-Wide Association Studies Fact Sheet. National Human Genome Research Institute website. https://www.genome.gov/20019523. Updated August 27, 2015. Accessed August 15, 2016.

27. Wellcome Trust Case Control C. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661-678

28. Johnson A. D. and O'Donnell C. J. (2009) An open access database of genome-wide association results. BMC Med Genet 10: 6

29. Hibar D. P., Stein J. L., Renteria M. E., Arias-Vasquez A., Desrivieres S., et al. (2015) Common genetic variants influence human subcortical brain structures. Nature 520: 224-229
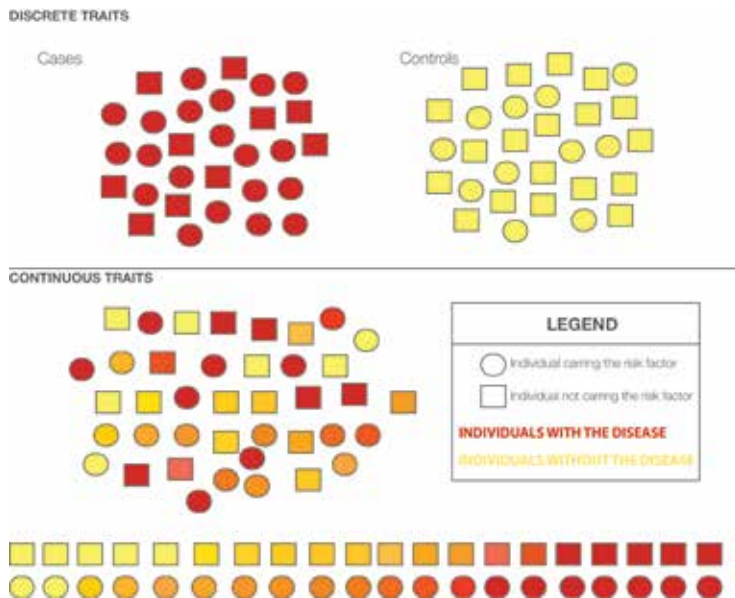
Figure 4. Different approaches in the study of discrete vs continuous traits. For discrete traits, samples are divided into cases and controls. For continuous traits, the use of linear models evaluates the correlation between the risk factor and the phenotype.

For Research Use Only. Not for use in diagnostic procedures.

Association Studies: DNA Arrays

Even though DNA sequencing costs have dropped dramatically in the past decade, it is not cost-effective to perform association studies sequencing whole genomes. Array technology leverages oligonucleotide probes to allow for the genotyping (as well as expression analysis) of up to hundreds of thousands of genomic markers at the same time.

Arrays are best designed to select representative SNPs that cover all genetic variation in a population. This result is achieved by sequencing a representative diverse set of individuals to identify haplotypes, defined as a collection of alleles that are likely to be inherited together.[30, 31] The use of DNA arrays and whole-genome SNP imputation is instrumental in association studies.

Imputation relies on a reference database of fully sequenced genomes to predict genotypes that are not assayed in a larger sample of individuals. The approach consists of first reconstructing haplotypes for the samples of interest using the haplotypes from the reference set (haplotype phasing), and then estimating genotypes [32] (Figures 5 and 6).

30. International HapMap C. (2003) The International HapMap Project. Nature 426: 789-796
31. Tanaka T. (2005) [International HapMap project]. Nihon Rinsho 63 Suppl 12: 29-34
32. Neale B. M. (2010) Introduction to linkage disequilibrium, the HapMap, and imputation. Cold Spring Harb Protoc 2010: pdb top74

Sequencing of reference population (sequence length = 51 bases)

Identification of haplotypes, and design array that uses a few SNPs to represent the region
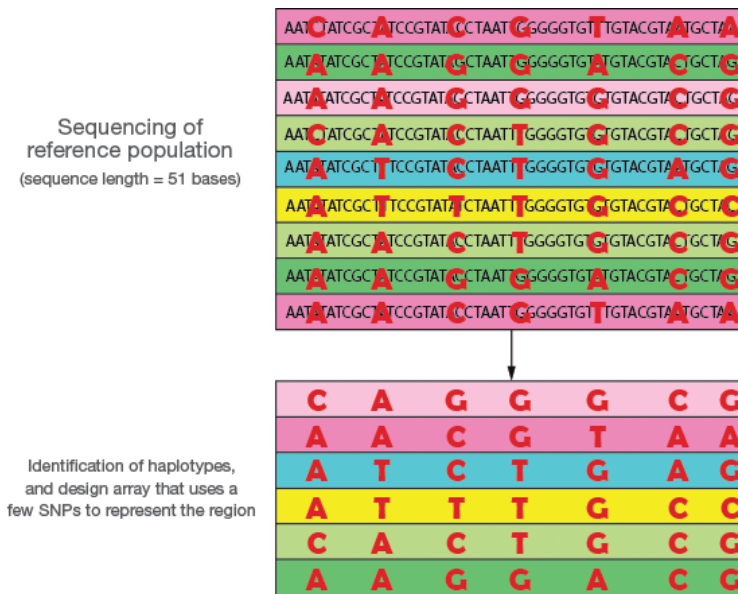
Figure 5. Haplotypes are groups of multiple alleles that are likely to be inherited together because of linkage disequilibrium. They can be identified through sequencing of multiple individuals and can provide representative (tag) SNPs to implement on arrays.
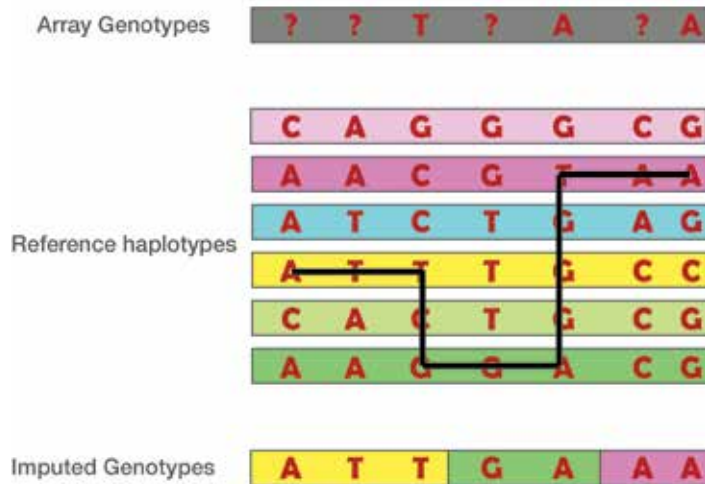
Figure 6. The process of imputation estimates the missing genotypes of an individual by first estimating the haplotypes (a process known as haplotype phasing) from the available genotype data.

Association Studies: Sequencing

Arrays are often used to study the associations between common variants and complex traits. This approach has yielded beneficial results; however, it is now widely accepted that rare variants also play an important role in the study of complex diseases.[33, 34]

The identification of rare variants associated with diseases can be accomplished by sequencing the full genome (whole-genome sequencing) or full exome (whole-exome sequencing) of affected individuals. These approaches have been successful in both family-based and cohort studies eg, in the study of autism,[35, 36] and lipid and adiponectin levels.[37]

Note that the identification of variants associated with a disease does not prove causality automatically. Once a risk variant is identified, many more studies are required to prove causality and provide a mechanism of how the variant affects the pathogenesis of disease. [38, 39]

Figure 7. Family-based exome or whole-genome sequencing studies have provided remarkable results in understanding the association between rare variants and complex diseases.

33. Cirulli E. T. and Goldstein D. B. (2010) Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet 11: 415-425

34. Steinberg S., Stefansson H., Jonsson T., Johannsdottir H., Ingason A., et al. (2015) Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease. Nat Genet 47: 445-447

35. Yuen R. K., Thiruvahindrapuram B., Merico D., Walker S., Tammimies K., et al. (2015) Whole-genome sequencing of quartet families with autism spectrum disorder. Nat Med 21: 185-191

36. Shi L., Zhang X., Golhar R., Otieno F. G., He M., et al. (2013) Whole-genome sequencing in an autism multiplex family. Mol Autism 4: 8

37. Consortium U. K., Walter K., Min J. L., Huang J., Crooks L., et al. (2015) The UK10K project identifies rare variants in health and disease. Nature 526: 82-90

38. Hrdlickova B., Westra H. J., Franke L. and Wijmenga C. (2011) Celiac disease: moving from genetic associations to causal variants. Clin Genet 80: 203-313

39. Tasan M., Musso G., Hao T., Vidal M., MacRae C. A., et al. (2015) Selecting causal genes from genome-wide association studies via functionally coherent subnetworks. Nat Methods 12: 154-159

Hu Y., Shmygelska A., Tran D., Eriksson N., Tung J. Y., et al. (2016) GWAS of 89,283 individuals identifies genetic variants associated with self-reporting of being a morning person. Nat Commun 7: 10448.

Chronobiology studies circadian rhythms in living organisms. The aim of this study was to investigate the genetic factors that predispose an individual to be a "morning person" as compared to a "night person." The authors conducted a GWAS on 38,937 self-reported morning persons and 50,346 self-reported night persons across a total of 8 million genotyped or imputed SNPs. They identified 15 significantly associated SNPs, 7 of which were located near established circadian genes (*RGS16*, *VIP*, *PER2*, *HCRTR2*, *RASD1*, *PER3*, and *FBXL3*). They also performed pathway analysis and found enrichment of both circadian and phototransduction factors.

Illumina Technology: BeadChip®

Davies G., Armstrong N., Bis J. C., Bressler J., Chouraki V., et al. (2015) Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). Mol Psychiatry 20: 183-192.

Cognitive function is a substantially heritable trait. The authors investigated the genetic contribution to cognitive function variation in middle-aged and older adults. They conducted a meta-analysis of GWAS of 31 cohorts for a total of 53,949 individuals who had taken several cognitive tests. They identified 13 genome-wide significant SNP associations in 3 genomic locations: 6q16.2 (rs10457441, *MIR2113*, $P = 3.93 \times 10$-9), 14q12 (rs17522122, *AKAP6*, $P = 2.55 \times 10$-8), and 19q13.32 (rs10119, *APOE/TOMM40*, $P = 5.67 \times 10$-9), and 1 gene-based association (HMGN1, $P = 10$-6). These genomic regions were all associated with neuropsychiatric phenotypes previously. They then estimated heritability by using a genome-wide complex trait analysis procedure on 2 larger cohorts, the Atherosclerosis Risk in Communities (N = 6,617) and the Health and Retirement Study (5,976). The proportion of phenotypic variance accounted for by the genotyped common SNPs was (29 ± 5)% and (28 ± 7)%, respectively.

Illumina Technology: BeadChip, Infinium

Mistry V., Bockett N. A., Levine A. P., Mirza M. M., Hunt K. A., et al. (2015) Exome sequencing of 75 individuals from multiply affected coeliac families and large scale resequencing follow up. PLoS One 10: e0116845.

Celiac disease (CeD) is a highly heritable autoimmune disease. It involves chronic inflammation of the small intestine in response to dietary wheat consumption. GWAS have identified 40 loci that, together with information on the human leukocyte antigen region, explain up to ~40% of the heritability of the disease. This study used exome sequencing data from 75 individuals belonging to 55 families to identify rare variants associated with CeD with a larger effect size (odds ratio ~ 2–5). Assessing shared variants between affected individuals, applying a model free linkage test, as well as gene burden tests for multiple potentially causal variants, the authors selected interesting variants and genes to follow up. They resequenced these genes in 2248 cases and 2230 controls, and identified 939 variants in coding regions (92% rare and 60% novel). However, gene burden tests performed on such variants identified no statistical association with the phenotype.

Illumina Technology: Illumina GA$_{IIx}$

Gudbjartsson D. F., Helgason H., Gudjonsson S. A., Zink F., Oddson A., et al. (2015) Large-scale whole-genome sequencing of the Icelandic population. Nat Genet 47: 435-444.

The authors performed whole-genome sequencing with a median depth of 20´ on samples from 2636 Icelanders. They found 20 million SNPs and 1.5 million insertion-deletions (indels). They described the density and frequency spectra of sequence variants in relation to their functional annotation, gene position, pathway, and conservation score. The results also demonstrated an excess of homozygosity and rare protein-coding variants in their populations, which were imputed down to a minor allele frequency of 0.1% in 104,220 individuals previously genotyped with several arrays. They also observed a recessive frameshift mutation in *MYL4* associated with early-onset atrial fibrillation, several mutations in *ABCB4* that increase risk of liver disease, and an intronic variant in *GNAS* associated with increased thyroid-stimulating hormone levels if maternally inherited.

Illumina Technology: Illumina GA$_{IIx}$, HiSeq®, cBot®, BeadChip, Infinium

For Research Use Only. Not for use in diagnostic procedures.

12   Genomic solutions for cell biology and complex disease research

Consortium U. K., Walter K., Min J. L., Huang J., Crooks L., et al. (2015) The UK10K project identifies rare variants in health and disease. Nature 526: 82-90.
The aim of this project was to explore the contribution that rare and low-frequency genetic variants have on several human traits. The study analyzed sequencing data from whole genomes or exomes of almost 10,000 individuals belonging to different population-based or disease cohorts. The authors identified and characterized over 24 million novel variants and generated an imputation panel of high accuracy. Using a single marker and rare variant association test, they also identified novel alleles associated with triglyceride (*APOB*), adiponectin (*ADIPOQ*), and low-density lipoprotein levels (*LDLR* and *RGAG1*). They collected data about population structure and functional annotation of rare and low-frequency variants and estimated the benefits of sequencing in association studies, in a disease-specific fashion. The authors also developed a web-based tool reporting individual genetic and phenotypic information, aimed at facilitating the exploration of results from future association studies.

Illumina Technology: HiSeq, various arrays

40. Albert F. W. and Kruglyak L. (2015) The role of regulatory variation in complex traits and disease. Nat Rev Genet 16: 197-212

41. Wang Z., Gerstein M. and Snyder M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57-63

42. McGettigan P. A. (2013) Transcriptomics in the RNA-seq era. Curr Opin Chem Biol 17: 4-11

43. Feng H., Qin Z. and Zhang X. (2013) Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. Cancer Lett 340: 179-191

Gene Expression

Gene expression is the mechanism through which the information contained in a genetic sequence is transcribed into RNA, resulting in a final product that is either a protein or noncoding RNA. The regulation of mRNA transcription and processing directly affects protein synthesis. The proteins, in turn, mediate cellular function to establish the phenotype of the cell.[40]

Gene expression studies measure the mRNA levels that are expressed by a cell under certain conditions or after a determinate stimulation. The expression level of a gene is itself a quantitative trait, and the study of how certain DNA polymorphisms influence gene expression is called expression quantitative loci (eQTL) mapping.

Gene Expression Studies: cDNA Microarrays

cDNA microarrays are microchips containing oligonucleotide probes that enable expression analysis of up to hundreds of thousands of genomic markers at the same time.

Gene Expression Studies: RNA Sequencing

Although arrays were the first technology applied to large-scale expression studies, they have been replaced, in large part, by RNA sequencing (RNA-Seq).[41] RNA-Seq provides information not only about gene expression but also polymorphisms, eQTLs, splice variants, epigenetic modifications, RNA editing, and gene fusion detection.

RNA-Seq uses next-generation sequencing (NGS) to quantify the abundance of RNA for expression analysis, as well to determine the underlying gene sequence. Careful analysis of the results, along with adaptation of the sample preparation protocols, can provide remarkable insight into aspects of RNA processing and the control of transcription. Examples include PTMs, RNA splicing, RNA bound to RNA-binding proteins (RBP), RNA expressed at various developmental stages, unique RNA isoforms, RNA degradation, and regulation of other RNA species.[42, 43] Further studies of RNA transcription and translation will enable a better understanding of RNA regulation and gene expression.

For Research Use Only. Not for use in diagnostic procedures.

Figure 8. Scientists have discovered a link between long-term memory and protein synthesis in the brain.[44, 45]

44. Davis H. P. and Squire L. R. (1984) Protein synthesis and memory: a review. Psychol Bull 96: 518-559
45. Holt C. E. and Schuman E. M. (2013) The central dogma decentralized: new perspectives on RNA function and local translation in neurons. Neuron 80: 648-657

Kim S., Becker J., Bechheim M., Kaiser V., Noursadeghi M., et al. (2014) Characterizing the genetic basis of innate immune response in TLR4-activated human monocytes. Nat Commun 5: 5236.
Toll-like receptors (TLRs) play a fundamental role in immunity, and dysregulation in TLR signaling can lead to autoimmune diseases, sepsis, or cancer. In this study, the authors extracted monocytes from a total of 185 male healthy volunteers aged 18–35. For each individual, cells were cultured overnight either untreated or treated with lipopolysaccharides to stimulate TLR4 expression. RNA and DNA were then extracted and analyzed through the use of arrays. The results revealed 1471 eQTLs unique to TLR4 stimulation. Among these, the authors found functional SNPs for the expression of *NEU4*, *CCL14*, *CBX3*, and *IRF5* on TLR4 activation. They also found that SNPs conferring risk to primary biliary cirrhosis (PBC), inflammatory bowel disease (IBD), and celiac disease were immune-response eQTLs for *PDGFB* and *IL18R1*, making these 2 genes plausible candidates for studying the pathophysiology of these disorders in the context of TLR4 activation.

Illumina Technology: BeadChip, Infinium

Fava V. M., Manry J., Cobat A., Orlova M., Van Thuc N., et al. (2016) A Missense LRRK2 Variant Is a Risk Factor for Excessive Inflammatory Responses in Leprosy. PLoS Negl Trop Dis 10: e0004412.
Leprosy is a chronic dermatoneurological disease, caused by infection from the etiological agent *Mycobacterium leprae*. Depending on the epidemiological settings, some leprosy patients suffer from excessive proinflammatory responses termed type-1 reactions (T1R). This study evaluated the role of the *LRRK2* gene and T1R in different sample settings. This gene encodes a multifunctional protein that modulates proinflammatory responses. LRRK2 variants have been associated with leprosy in some studies but results have been inconsistent in replications. This study selected a total of 1372 individuals and divided them into 2 groups: leprosy-affected families with and without T1R. The authors tested association on 156 SNPs using a genotyping array. *LRRK2* expression levels were also analyzed using an expression array. They identified a total of 18 T1R-specific variants, of which the core SNP capturing the T1R association was the missense variant M239T (rs3761863), which affects the turnover of the LRRK2 protein. Moreover, a bin of 9 SNPs associated with T1R were eQTLs for LRRK2 in unstimulated whole blood cells, but not after exposure to the *Mycobacterium leprae* antigen.

Illumina Technology: BeadChip

Bentham J., Morris D. L., Cunninghame Graham D. S., Pinder C. L., Tombleson P., et al. (2015) Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. Nat Genet 47: 1457-1464.
Systemic lupus erythematosus (SLE) is a complex autoimmune disease, characterized by loss of immune tolerance to nuclear and cell-surface antigens. The authors genotyped 4946 SLE cases and 1286 healthy controls using an array and integrated the data with other studies, obtaining a total of 7219 cases and 15,991 controls of European ancestry. They used these samples for a meta-analysis and replication study, from which they mapped 43 susceptibility loci, including 10 new associations. Imputation provided evidence for missense variants associated with the disease, covering 8 genes. The authors identified other putatively causal genes by the analysis of associated alleles for cis-acting eQTL effects in a range of *ex vivo* immune cells. They found an over-representation of transcription factors among SLE susceptibility genes.

Illumina Technology: BeadChip, Infinium

For Research Use Only. Not for use in diagnostic procedures.

14 Genomic solutions for cell biology and complex disease research

Johnson M. R., Shkura K., Langley S. R., Delahaye-Duriez A., Srivastava P., et al. (2016) Systems genetics identifies a convergent gene network for cognition and neurodevelopmental disease. Nat Neurosci 19: 223-232.
This study investigated the genetic determinants of cognition. The authors performed a system-level analysis of expression data from healthy individuals and patients with temporal-lobe epilepsy (TLE). They also analyzed a murine model. They used genome-wide expression analysis to infer gene-regulatory networks conserved across species and brain regions. Two of the identified networks showed enrichment for common variants affecting cognitive abilities. Exome sequencing results from 6871 trios showed that M3 genes were also enriched for mutations from patients with neurodevelopmental diseases.

Illumina Technology: TruSeq® RNA kit, HiSeq® 2000, BeadChip

Sareddy G. R., Zhang Q., Wang R., Scott E., Zou Y., et al. (2015) Proline-, glutamic acid-, and leucine-rich protein 1 mediates estrogen rapid signaling and neuroprotection in the brain. Proc Natl Acad Sci U S A 112: E6673-6682.
17-β-estradiol (E2) has been implicated as neuroprotective in several neurodegenerative studies. In previous studies, the authors identified the proline-, glutamic acid–, and leucine-rich protein 1 (PELP1) as a coregulator of E2. The present study evaluated the role of PELP1 in mediating the neuroprotective and cognitive effects of E2 in the brain, using a forebrain-specific *PELP1* knockout (FBKO) mouse model. Ovariectomized placebo- and E2-implanted FBKO and control mice were subjected to global cerebral ischemia (GCI), and the hippocampi were collected at various time points after reperfusion (10 min, 30 min, and 3 h). RNA-Seq demonstrated that several genes related to inflammation, metabolism, and survival were altered in *PELP1* FBKO mice. It also showed a significant reduction in the activation of the Wnt/β-catenin signaling pathway. *PELP1* FBKO studies also revealed that PELP1 is required for E2-mediated neuroprotection and preservation of cognitive function after GCI.

Illumina Technology: TruSeq® RNA Sample Prep Kit, HiSeq 2000

Mittal A., Pachter L., Nelson J. L., Kjaergaard H., Smed M. K., et al. (2015) Pregnancy-Induced Changes in Systemic Gene Expression among Healthy Women and Women with Rheumatoid Arthritis. PLoS One 10: e0145204.
Pregnancy induces systemic biological changes, with beneficial effects on certain autoimmune conditions, such as rheumatoid arthritis (RA). This study characterized, in depth, the differences in pregnancy-induced changes between healthy women and RA patients. The authors performed RNA-Seq in 5 healthy women and 20 RA patients. A total of 4710 genes were significantly associated to pregnancy status, regardless of the presence of RA. Only 98 genes exhibited expression patterns that were differentially associated with pregnancy in RA patients vs controls.

Illumina Technology: TruSeq RNA Sample Prep Kit, HiSeq 2500

46. Holt C. E. and Schuman E. M. (2013) The central dogma decentralized: new perspectives on RNA function and local translation in neurons. Neuron 80: 648-657
47. Egger G., Liang G., Aparicio A. and Jones P. A. (2004) Epigenetics in human disease and prospects for epigenetic therapy. Nature 429: 457-463
48. Waddington C. H. (2012) The epigenotype. 1942. Int J Epidemiol 41: 10-13
49. Holliday R. (2006) Epigenetics: a historical overview. Epigenetics 1: 76-80
50. Epigenetics: Fundamentals. What is Epigenetics webpage. http://www.whatisepigenetics.com/fundamentals/. Accessed August 15, 2016.
51. Lockett G. A., Patil V. K., Soto-Ramirez N., Ziyab A. H., Holloway J. W., et al. (2013) Epigenomics and allergic disease. Epigenomics 5: 685-699
52. Qureshi I. A. and Mehler M. F. (2014) Epigenetic mechanisms underlying the pathogenesis of neurogenetic diseases. Neurotherapeutics 11: 708-720

Epigenetics

With a few exceptions, cells originating from one organism contain the same gene sequences in their DNA. Some of these genes are ubiquitously expressed, while others are only expressed in certain cell types,[46] or exhibit different patterns of expression in different cell types. These differences are at the basis of cellular differentiation.[47, 48, 49] Epigenetics is the science that studies the dynamic alterations in the transcriptional potential of a cell.

Epigenetic changes can be heritable changes in the expression of a gene (a phenotype) that are not determined by changes in germline DNA.[50] The environment can also play a role in influencing epigenetic changes. Studies have demonstrated that factors such as age, environment and lifestyle, and disease status can have a significant effect on the epigenome.[51, 52]

DNA Methylation

DNA methylation and hydroxymethylation are involved in development, X-chromosome inactivation, cell differentiation, tissue-specific gene expression, epigenetic variation, imprinting, and diseases.[53, 54, 55, 56, 57] Methylation usually occurs at the 5' position of cytosines. It plays a crucial role in gene regulation and chromatin remodeling.

Figure 9. The active agouti gene in mice codes for yellow coat color. Pregnant mice with the active gene fed with a diet rich in methyl donors will give birth to offspring whose agouti gene will be switched off.[58, 59]

Most cytosine methylation occurs in cytosines located near guanines, called CpG sites. These CpG sites are often located upstream of promoters, or within the gene body. CpG islands are genomic segments with high GC content and CpG ratios that are statistically higher than expected by chance.

While cytosine methylation (5mC) is known as a silencing mark that represses genes, cytosine hydroxymethylation (5hmC) is shown to be an activating mark that promotes gene expression and is a proposed intermediate in the DNA demethylation pathway.[60, 61, 62] Similar to 5mC, 5hmC is involved in development, cancers, cell differentiation, and diseases.[63]

5mC and/or 5hmC can be a diagnostic tool to help identify the effects of nutrition, carcinogens,[64] and environmental factors in relation to diseases. The impact of these base modifications on gene regulation depends on their locations within the genome. Therefore, it is important to determine the exact position of the modified bases.

DNA-Protein Interactions and Histone Modifications

Chromatin remodeling is a dynamic process driven by epigenetic factors that change DNA-protein interactions. These factors can involve protein modifications, such as histone methylation, acetylation, phosphorylation, and ubiquitination.[65] Histone modifications determine gene activation by recruiting regulatory factors and maintaining an open or closed chromatin state.

53. Kraiczy J., Nayak K., Ross A., Raine T., Mak T. N., et al. (2015) Assessing DNA methylation in the developing human intestinal epithelium: potential link to inflammatory bowel disease. Mucosal Immunol

54. Kaut O., Ramirez A., Pieper H., Schmitt I., Jessen F., et al. (2014) DNA methylation of the TNF-alpha promoter region in peripheral blood monocytes and the cortex of human Alzheimer's disease patients. Dement Geriatr Cogn Disord 38: 10-15

55. Coppieters N., Dieriks B. V., Lill C., Faull R. L., Curtis M. A., et al. (2014) Global changes in DNA methylation and hydroxymethylation in Alzheimer's disease human brain. Neurobiol Aging 35: 1334-1344

56. Sanchez-Mut J. V., Aso E., Panayotis N., Lott I., Dierssen M., et al. (2013) DNA methylation map of mouse and human brain identifies target genes in Alzheimer's disease. Brain 136: 3018-3027

57. Bakulski K. M., Dolinoy D. C., Sartor M. A., Paulson H. L., Konen J. R., et al. (2012) Genome-wide DNA methylation differences between late-onset Alzheimer's disease and cognitively normal controls in human frontal cortex. J Alzheimers Dis 29: 571-588

58. Dolinoy D. C., Weidman J. R., Waterland R. A. and Jirtle R. L. (2006) Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. Environ Health Perspect 114: 567-572

59. Dolinoy D. C. (2008) The agouti mouse model: an epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome. Nutr Rev 66 Suppl 1: S7-11

60. Kloosterman W. P. and Plasterk R. H. (2006) The diverse functions of microRNAs in animal development and disease. Dev Cell 11: 441-450

61. Feng H., Qin Z. and Zhang X. (2013) Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. Cancer Lett 340: 179-191

62. Holt C. E. and Schuman E. M. (2013) The central dogma decentralized: new perspectives on RNA function and local translation in neurons. Neuron 80: 648-657

63. Pfeifer G. P., Kadam S. and Jin S. G. (2013) 5-hydroxymethylcytosine and its potential roles in development and cancer. Epigenetics Chromatin 6: 10

64. Thomson J. P., Lempiainen H., Hackett J. A., Nestor C. E., Muller A., et al. (2012) Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. Genome Biol 13: R93

65. Rivera C. M. and Ren B. (2013) Mapping human epigenomes. Cell 155: 39-55

For Research Use Only. Not for use in diagnostic procedures.

16   Genomic solutions for cell biology and complex disease research

Epigenetic factors play roles in tissue development,[66] embryogenesis, cell fate, immune response, and diseases.[67] Bacterial pathogens can elicit transcriptional repression of immune genes by chromatin remodeling. The study of protein-DNA interactions has also demonstrated that chromatin remodeling[68] can respond to external factors, such as alcohol abuse,[69] cigarette smoking,[70] and drug use.

Noncoding RNA–Associated Gene Silencing

The central dogma of biology states that genetic information coded by DNA is transcribed into individual transcripts that are composed of mRNA. Each transcript contains the information required to synthesize a particular protein (or a small number of proteins).[71] Noncoding RNAs (ncRNAs) represent an exception to the central dogma.[72]

It is estimated that, while 70%–90% of DNA is transcribed into RNA, only 3% is translated into proteins. Transfer RNA (tRNA) and ribosomal RNA (rRNA) have structural and regulatory roles in translation. Many other classes of ncRNAs, originally thought to be remnants of the transcription process, are now known to have biological functions.[73, 74, 75]

ncRNAs are divided into 2 groups: short ncRNAs (microRNAs, siRNAs, snRNAs, exRNAs, piRNAs, and scaRNAs) and long ncRNAs (such as Xist and HOTAIR).[76] These classes of molecules are the subject of many current studies and are involved in a variety of functions, including gene expression, regulation, replication, and silencing. Past studies have found associations between altered genetic expression due to ncRNAs and diseases such as Alzheimer's and autism.[77, 78]

Gjoneska E., Pfenning A. R., Mathys H., Quon G., Kundaje A., et al. (2015) Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. Nature 518: 365-369.
Alzheimer's disease (AD) is characterized by the accumulation of β-amyloid plaques and neuronal loss that leads to cognitive decline. This study investigated chromatin state alterations during neurodegeneration. The authors used chromatin-immunoprecipitation sequencing (ChIP-Seq) and RNA-Seq to profile transcriptional and chromatin state dynamics across early and late pathology in the hippocampus of an adult inducible-mouse model of AD-like neurodegeneration. They found a coordinated downregulation of synaptic plasticity genes and regulatory regions, as well as upregulation of immune response genes and regulatory regions, targeted by factors belonging to the ETS family of transcriptional regulators. Orthologous human regions to increasing-level enhancers showed immune cell–specific enhancer signatures, as well as immune cell expression of quantitative loci. Decreasing-level enhancer orthologs showed fetal brain–specific enhancer activity. Of note, AD-associated genetic variants were specifically enriched in increasing-level enhancer orthologs, implicating immune-system involvement in AD predisposition.

Illumina Technology: TruSeq, HiSeq 2000

Scobie K. N., Damez-Werno D., Sun H., Shao N., Gancarz A., et al. (2014) Essential role of poly(ADP-ribosyl)ation in cocaine action. Proc Natl Acad Sci U S A 111: 2005-2010.
Poly(ADP-ribose) polymerase 1 (PARP-1) is a ubiquitous and abundant nuclear protein that catalyzes the synthesis of a polymer called poly(ADP-ribose) on histones and other substrate proteins, forming transcriptional regulatory complexes with other chromatin proteins. This study analyzed the role of PARP-1 in cocaine-induced molecular, neural, and behavioral plasticity, using ChIP-Seq on cocaine-exposed mice. The results demonstrated a global, genome-wide enrichment of PARP-1 in nucleus accumbens (NAc) and identified several PARP-1 targeted genes that could be implicated in the long-lasting effects of cocaine. In particular, the authors identified sidekick-1, a gene important for synaptic connections during development, as a critical target of PARP-1. This gene is involved in cocaine's behavioral effects as well as its ability to induce dendritic spines on NAc neurons.

Illumina Technology: HiSeq 2000

66. Pinello L., Xu J., Orkin S. H. and Yuan G. C. (2014) Analysis of chromatin-state plasticity identifies cell-type-specific regulators of H3K27me3 patterns. Proc Natl Acad Sci U S A 111: E344-353

67. Bonomi R., Mukhopadhyay U., Shavrin A., Yeh H. H., Majhi A., et al. (2015) Novel Histone Deacetylase Class IIa Selective Substrate Radiotracers for PET Imaging of Epigenetic Regulation in the Brain. PLoS One 10: e0133512

68. Kahrstrom C. T. (2013) Epigenetics: Legionella makes its mark on histones. Nat Rev Genet 14: 370

69. Warnault V. and Ron D. (2013) Chromatin remodeling: a new landscape to treat harmful alcohol-use disorders. Future Med Chem 5: 2011-2013

70. Marwick J. A., Kirkham P. A., Stevenson C. S., Danahay H., Giddings J., et al. (2004) Cigarette smoke alters chromatin remodeling and induces proinflammatory genes in rat lungs. Am J Respir Cell Mol Biol 31: 633-642

71. Crick F. (1970) Central dogma of molecular biology. Nature 227: 561-563

72. Holt C. E. and Schuman E. M. (2013) The central dogma decentralized: new perspectives on RNA function and local translation in neurons. Neuron 80: 648-657

73. Zaratiegui M., Irvine D. V. and Martienssen R. A. (2007) Noncoding RNAs and gene silencing. Cell 128: 763-776

74. Malecova B. and Morris K. V. (2010) Transcriptional gene silencing through epigenetic changes mediated by non-coding RNAs. Curr Opin Mol Ther 12: 214-222

75. Holoch D. and Moazed D. (2015) RNA-mediated epigenetic regulation of gene expression. Nat Rev Genet 16: 71-84

76. Kashi K., Henderson L., Bonetti A. and Carninci P. (2016) Discovery and functional analysis of lncRNAs: Methodologies to investigate an uncharacterized transcriptome. Biochim Biophys Acta 1859: 3-15

77. Simchovitz A., Soreq L. and Soreq H. (2016) Transcriptome profiling in Parkinson's leukocytes: from early diagnostics to neuroimmune therapeutic prospects. Curr Opin Pharmacol 26: 102-109

78. Zhang J., Yuan L., Zhang X., Hamblin M. H., Zhu T., et al. (2016) Altered long non-coding RNA transcriptomic profiles in brain microvascular endothelium after cerebral ischemia. Exp Neurol 277: 162-170

Elboudwarej E., Cole M., Briggs F. B., Fouts A., Fain P. R., et al. (2016) Hypomethylation within gene promoter regions and type 1 diabetes in discordant monozygotic twins. J Autoimmun.

Epidemiological studies have provided evidence of a genetic basis for type 1 diabetes (T1D), yet the currently known genetic variants associated with diabetes cannot explain entirely disease risk. This study investigated whether epigenetic modifications of DNA contribute to T1D. The authors used the Infinium HumanMethylation450 BeadChip array on 7 long-term disease-discordant monozygotic (MZ) twin pairs and 5 pairs of HLA-identical disease-discordant non–twin siblings (NTS) to test for DNA methylation (DNAm) and T1D associations. They found strong evidence for hypomethylation of CpG sites within promoter regions in MZ twins with T1D as compared to their healthy control. The authors divided DNA methylation data into 3 categories for further analysis, including: 1) the MHC region; 2) non-MHC genes with reported T1D association through GWAS; and 3) the epigenome, or remainder of sites that did not include MHC- and T1D-associated genes. The results showed modest methylation differences between discordant MZ in the first 2 groups (*BACH2*, *INS-IGF2*, and *CLEC16A*, DNAm difference range 2.2%–5%), and greater differences in the third (*MAGI2*, *FANCC*, and *PCDHB16*, DNAm difference range 6.9%–16.1%). These results, however, were not replicated in 6 additional MZ twin discordant pairs.

Illumina Technology: Infinium

79. Post-translational modifications – Latest research and news. Nature website. http://www.nature.com/subjects/post-translational-modifications. Accessed August 15, 2016

80. Karve T. M. and Cheema A. K. (2011) Small changes huge impact: the role of protein posttranslational modifications in cellular homeostasis and disease. J Amino Acids 2011: 207691

81. Stamnaes J. and Sollid L. M. (2015) Celiac disease: Autoimmunity in response to food antigen. Semin Immunol 27: 343-352

82. Marre M. L., James E. A. and Piganelli J. D. (2015) beta cell ER stress and the implications for immunogenicity in type 1 diabetes. Front Cell Dev Biol 3: 67

83. Mattson M. P. (2010) Acetylation unleashes protein demons of dementia. Neuron 67: 900-902

Posttranslational Modifications

Translation is the process through which ribosomes synthesize proteins by decoding the mRNA. In eukaryotic cells, this process happens in the cytosol or across the membrane of the endoplasmic reticulum (ER). Once translation is completed, several enzymes catalyze PTMs.

These modifications are generally covalent additions of a functional group to a protein (such as phosphorylation, glycosylation and glycation, ubiquitination, and hydroxylation), or the proteolytic processing and processes that are necessary for a newly synthesized protein to mature and reach the functional fold.[79, 80]

Certain PTMs have been linked to complex phenotypes and diseases. Some examples are celiac disease (CeD), type 1 Diabetes (T1D), neurodegenerative disorders, and drug response.[81, 82, 83, 84]

For Research Use Only. Not for use in diagnostic procedures.

18 Genomic solutions for cell biology and complex disease research

Ingolia N. T., Ghaemmaghami S., Newman J. R. and Weissman J. S. (2009) Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. Science 324: 218-223.
Active mRNA translation sequencing (ARTSeq), also called ribosome profiling (Ribo-Seq), isolates RNA that is being processed by the ribosome in order to monitor the translation process. In this method, ribosome-bound RNA first undergoes digestion. The RNA is then extracted, and the rRNA is depleted. Extracted RNA is reverse-transcribed to cDNA. Deep sequencing of the cDNA provides the sequences of RNAs bound by ribosomes during translation. This method has been refined to improve the quality and quantitative nature of the results. Careful attention should be paid to: 1) generation of cell extracts in which ribosomes have been faithfully halted along the mRNA they are translating *in vivo*; 2) nuclease digestion of RNAs that are not protected by the ribosome, followed by recovery of the ribosome-protected mRNA fragments; and 3) quantitative conversion of the protected RNA fragments into a DNA library that can be analyzed by deep sequencing. The addition of harringtonine (an alkaloid that inhibits protein biosynthesis) causes ribosomes to accumulate precisely at initiation codons and assists in their detection.

Illumina Technology: Illumina GA$_{II}$

Posttranslational Modifications: Featured Method—Ribosome Profiling Sequencing (Ribo-Seq)/ARTSeq™

Figure 10. Overview of the Ribo-Seq/ARTSeq method.

Schafer S., Adami E., Heinig M., Rodrigues K. E., Kreuchwig F., et al. (2015) Translational regulation shapes the molecular landscape of complex disease phenotypes. Nat Commun 6: 7200.
RNA expression phenotypes are heritable; however, knowledge of the extent to which translational regulation influences phenotypic traits is still lacking. To investigate and quantify strain-specific translational regulation, the authors performed genome-wide RNA-Seq and Ribo-Seq in heart and liver tissue from 2 inbred rat strains (n = 5 per strain). They observed that most of the enriched pathways in both heart and liver could be detected through Ribo-Seq but not mRNA expression analysis, suggesting that protein variation might affect the penetrance of disease-causing variants. By studying candidate genes associated with cholesterol, heart rate, and metabolic traits from previous GWAS in humans, they also identified a multitude of examples of strain-specific transcriptional and/or translational regulation genes in rat. These results demonstrate that the knowledge of interindividual variability in the translated genome will be useful in unraveling the role of genes and regulatory pathways in the expression of complex phenotypes.

Illumina Technology: HiSeq 2000

MODEL SYSTEMS

As most research cannot be conducted on humans directly, the use of model systems is important to the understanding of complex diseases.

Animal Models

Animals share many physiological and anatomical similarities with humans,[85] and animal models have informed much of the current knowledge of cell biology and physiology.[86] Although species-to-species differences may represent a limit in the translation of results from animal model to human,[87] animal models still represent an indispensable tool for molecular biologists.[88] Several animal models have provided important results in the study of complex diseases (Table 2).

Table 2. Some Animal Model Systems Used in the Study of Complex Diseases

Common Name	Species Name	References
Roundworm	*Caenorhabditis elegans*	89, 90, 91
Fruit fly	*Drosophila melanogaster*	92, 93, 94
Zebrafish	*Danio rerio*	95, 96, 97
African clawed frog	*Xenopus laevis*	98, 99, 100
Mouse	*Mus musculus*	101, 102, 103
Rat	*Rattus norvegicus*	104, 105, 106
Nonhuman primates	Mostly Macaca species	107, 108, 109

Valenci I., Yonai L., Bar-Yaacov D., Mishmar D. and Ben-Zvi A. (2015) Parkin modulates heteroplasmy of truncated mtDNA in Caenorhabditis elegans. Mitochondrion 20: 64-70.
The Parkin protein is a key player in the removal of damaged mitochondria via mitophagy, and it is mutated in many forms of Parkinsonism. In the mitochondrial DNA (mtDNA) form, damaged mitochondria carry mutations, creating a mixture of mtDNA populations (heteroplasmy). Previous studies demonstrated that higher expression levels of Parkin are associated with reduced heteroplasmy in human cytoplasmic hybrids. This study investigated whether Parkin has the same effect in vivo. The authors performed massive parallel sequencing of mtDNA from 3 strains of *Caenorhabditis elegans* expressing different levels of the homologous gene for Parkin, pdr-1. Their results showed that pdr-1 modulates the levels of heteroplasmy in C. elegans, as higher pdr-1 expression levels were associated with reduced heteroplasmy. These results provide evidence of pdr-1 functionality *in vivo* and its evolutionary conservation.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, MiSeq

Duff M. O., Olson S., Wei X., Garrett S. C., Osman A., et al. (2015) Genome-wide identification of zero nucleotide recursive splicing in Drosophila. Nature 521: 376-379.
Recursive splicing is a process through which large introns are removed by multiple steps of splicing. These steps happen at ratchet points—5' sites recreated after splicing. Originally, only 4 genes in Drosophila were known to undergo recursive splicing. In this study, the authors isolated RNA from *D. melanogaster*, *D. sechellia*, *D. makuba*, *D. pseudoobscura*, and *D. virilis*, as well as from 20 human tissues, and performed RNA-Seq. They identified 130 introns in 115 *Drosophila* genes that undergo recursive splicing, demonstrating that it is a common mechanism in these species. They also identified 4 recursively spliced genes in humans, of which 1 is also recursively spliced in Drosophila. Through RNA interference experiments in Drosophila cells, they also demonstrated that depletion of U2AF inhibits recurrent splicing.

Illumina Technology: TruSeq Stranded Total RNA Sample Prep Kit, HiSeq 2500, NextSeq 500

84. Peleg S., Sananbenesi F., Zovoilis A., Burkhardt S., Bahari-Javan S., et al. (2010) Altered histone acetylation is associated with age-dependent memory impairment in mice. Science 328: 753-756

85. The Animal Model. Speaking of Research website. http://speakingofresearch.com/facts/the-animal-model/. Accessed August 15, 2016

86. Fields S. and Johnston M. (2005) Cell biology. Whither model organism research? Science 307: 1885-1886

87. Rhrissorrakrai K., Belcastro V., Bilal E., Norel R., Poussin C., et al. (2015) Understanding the limits of animal models as predictors of human biology: lessons learned from the sbv IMPROVER Species Translation Challenge. Bioinformatics 31: 471-483

88. Roelfsema P. R. and Treue S. (2014) Basic neuroscience research with nonhuman primates: a small but indispensable component of biomedical research. Neuron 82: 1200-1204

89. Xiao Y., Liu F., Zhang Z., Tang J., Zou C. G., et al. (2016) Gut-Colonizing Bacteria Promote C. elegans Innate Immunity by Producing Nitric Oxide. Cell Rep 14: 1301-1307

90. Valenci I., Yonai L., Bar-Yaacov D., Mishmar D. and Ben-Zvi A. (2015) Parkin modulates heteroplasmy of truncated mtDNA in Caenorhabditis elegans. Mitochondrion 20: 64-70

91. Safka Brozkova D., Deconinck T., Griffin L. B., Ferbert A., Haberlova J., et al. (2015) Loss of function mutations in HARS cause a spectrum of inherited peripheral neuropathies. Brain 138: 2161-2172

92. Duff M. O., Olson S., Wei X., Garrett S. C., Osman A., et al. (2015) Genome-wide identification of zero nucleotide recursive splicing in Drosophila. Nature 521: 376-379

93. Zhang K. X., Tan L., Pellegrini M., Zipursky S. L. and McEwen J. M. (2016) Rapid Changes in the Translatome during the Conversion of Growth Cones to Synaptic Terminals. Cell Rep 14: 1258-1271

94. Yang C. P., Fu C. C., Sugino K., Liu Z., Ren Q., et al. (2016) Transcriptomes of lineage-specific Drosophila neuroblasts profiled by genetic targeting and robotic sorting. Development 143: 411-421

95. Don E. K., de Jong-Curtain T. A., Doggett K., Hall T. E., Heng B., et al. (2016) Genetic basis of hindlimb loss in a naturally occurring vertebrate model. Biol Open

96. Wu C. C., Kruse F., Vasudevarao M. D., Junker J. P., Zebrowski D. C., et al. (2016) Spatially Resolved Genome-wide Transcriptional Profiling Identifies BMP Signaling as Essential Regulator of Zebrafish Cardiomyocyte Regeneration. Dev Cell 36: 36-49

97. de la Calle Mustienes E., Gomez-Skarmeta J. L. and Bogdanovic O. (2015) Genome-wide epigenetic cross-talk between DNA methylation and H3K27me3 in zebrafish embryos. Genom Data 6: 7-9

For Research Use Only. Not for use in diagnostic procedures.

20   Genomic solutions for cell biology and complex disease research

Yang C. P., Fu C. C., Sugino K., Liu Z., Ren Q., et al. (2016) Transcriptomes of lineage-specific Drosophila neuroblasts profiled by genetic targeting and robotic sorting. Development 143: 411-421.
The brain consists of numerous distinct neurons that originate from a limited number of progenitors. These progenitors in *Drosophila* are called neuroblasts, and each of these produces a specific neuronal lineage. This study identified the transcriptional networks that underlie the development of distinct neuroblast lineages. The authors marked and isolated lineage-specific neuroblasts for RNA-Seq. Through the activation of a conditional neuroblast driver in specific lineages, they labeled particular neuroblasts throughout neurogenesis. The authors built a custom single-cell picking device, through which they recovered the targeted neurons. They performed RNA-Seq of mushroom body, antennal lobe, and type II neuroblasts compared with nonselective neuroblasts, neurons, and glia. The results revealed a vast repertoire of transcription factors expressed among neuroblasts in diverse patterns. Besides transcription factors that are likely to be pan-neuroblasts, there are many transcription factors that are selectively enriched or expressed in certain neuroblasts. The unique combinations of transcription factors present in different neuroblasts may govern different, lineage-specific neuron fates.

Illumina Technology: HiSeq 2500

de la Calle Mustienes E., Gomez-Skarmeta J. L. and Bogdanovic O. (2015) Genome-wide epigenetic cross-talk between DNA methylation and H3K27me3 in zebrafish embryos. Genom Data 6: 7-9.
DNA methylation and histone modifications are epigenetic changes that play a role in the complex regulation of vertebrate embryogenesis. Various studies have described the interaction between methylation and the Polycomb-dependent H3K27me3 histone mark in different organisms, and both these marks are necessary for proper development. In this study, the authors collected data from zebrafish embryos at 3 different stages: blastula (dome), pharyngula (24 hpf), and hatching (48 hpf). They harvested 500 dome, 100 24-hpf, and 100 48-hpf samples for DNA affinity capture (MethylCap-Seq) and 10 times as many for each group for ChIP-Seq. They provided genome-wide DNA methylation and H3K27me3 maps for the 3 stages, observing a strong antagonism between the 2 epigenetic marks present in CpG islands and their compatibility throughout the bulk of the genome. These results are in accord with what has been observed previously in mammalian embryonic stem cell (ESC) lines.

Illumina Technology: HiSeq 2000

Liu C., Lou C. H., Shah V., Ritter R., Talley J., et al. (2016) Identification of microRNAs and microRNA targets in Xenopus gastrulae: The role of miR-26 in the regulation of Smad1. Dev Biol 409: 26-38.
miRNAs play diverse roles in development. This study characterized miRNA–target mRNA relationships in embryonic development. The authors isolated RNA from 6 independent sets of *Xenopus laevis* samples from which they obtained small RNAs for sequencing. They identified 180 miRNAs, and located their precursor sequences in the *Xenopus* genome. Of these, 141 had been previously identified in *Xenopus tropicalis*, while 39 had not been previously described. Through a biochemical approach, they also isolated miRNAs associated with the RNA-induced silencing complex (RISC) in the early gastrula and demonstrated that some of these complexes may be under regulation by miRNAs. One of these is Smad1, for which target prediction algorithms predicted a miRNA binding site for miR-26. Disruption of Smad1-miR26 interactions through target protector morpholino oligonucleotide (TPMO) resulted in a 2-fold accumulation of Smad1 protein, an increase in its target genes, and a partially ventralized phenotype in 25% of the embryos. Overexpression of miR-26 resulted in a moderate decrease in Smad-1 dependent genes.

Illumina Technology: Illumina GA$_{IIx}$

Dichmann D. S., Walentek P. and Harland R. M. (2015) The alternative splicing regulator Tra2b is required for somitogenesis and regulates splicing of an inhibitory Wnt11b isoform. Cell Rep 10: 527-536.
Alternative mRNA splicing is a common mechanism in vertebrates. This study investigated the role of Transformer-2b (Tra2b), whose homolog promotes a cascade of alternative splicing events resulting in sex determination in Drosophila. The authors knocked down the *tra2b* gene in *Xenopus laevis* and *X. tropicalis* using morpholino-oligonucleotides. This resulted in embryos carrying multiple defects, including the failure to form somites. They used RNA-Seq to identify splice changes in transcripts from *tra2b* morphants. They identified 142 splice changes, mostly intron retention and exon skipping, of which 89% were not previously annotated—in particular, a previously unknown isoform of *wnt11b* that retains the last intron. The induction of the retention of this intron resulted in defective somitogenesis but not in the other defects observed in *tra2b* morphants.

Illumina Technology: TruSeq RNA Sample Preparation Kit, HiSeq 2000

98. Dichmann D. S., Walentek P. and Harland R. M. (2015) The alternative splicing regulator Tra2b is required for somitogenesis and regulates splicing of an inhibitory Wnt11b isoform. Cell Rep 10: 527-536

99. Liu C., Lou C. H., Shah V., Ritter R., Talley J., et al. (2016) Identification of microRNAs and microRNA targets in Xenopus gastrulae: The role of miR-26 in the regulation of Smad1. Dev Biol 409: 26-38

100. Tsujioka H., Kunieda T., Katou Y., Shirahige K. and Kubo T. (2015) Unique gene expression profile of the proliferating Xenopus tadpole tail blastema cells deciphered by RNA-sequencing analysis. PLoS One 10: e0111655

101. Angermueller C., Clark S. J., Lee H. J., Macaulay I. C., Teng M. J., et al. (2016) Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. Nat Methods

102. Castro-Mejia J., Jakesevic M., Krych L., Nielsen D. S., Hansen L. H., et al. (2016) Treatment with a Monoclonal Anti-IL-12p40 Antibody Induces Substantial Gut Microbiota Changes in an Experimental Colitis Model. Gastroenterol Res Pract 2016: 4953120

103. Kao C. Y., He Z., Zannas A. S., Hahn O., Kuhne C., et al. (2016) Fluoxetine treatment prevents the inflammatory response in a mouse model of posttraumatic stress disorder. J Psychiatr Res 76: 74-83

104. Lee Y. S., Lewis J. A., Ippolito D. L., Hussainzada N., Lein P. J., et al. (2016) Repeated exposure to neurotoxic levels of chlorpyrifos alters hippocampal expression of neurotrophins and neuropeptides. Toxicology 340: 53-62

105. Hannon E., Chand A. N., Evans M. D., Wong C. C., Grubb M. S., et al. (2016) A role for Ca1 and calcineurin signaling in depolarization-induced changes in neuronal DNA methylation. Neuroepigenetics 3: 1-6

106. Imperio C. G., McFalls A. J., Colechio E. M., Masser D. R., Vrana K. E., et al. (2015) Assessment of individual differences in the rat nucleus accumbens transcriptome following taste-heroin extended access. Brain Res Bull

107. Vermunt M. W., Tan S. C., Castelijns B., Geeven G., Reinink P., et al. (2016) Epigenomic annotation of gene regulatory alterations during evolution of the primate brain. Nat Neurosci 19: 494-503

108. Liu Z., Li X., Zhang J. T., Cai Y. J., Cheng T. L., et al. (2016) Autism-like behaviours and germline transmission in transgenic monkeys overexpressing MeCP2. Nature 530: 98-102

109. Lun M. P., Johnson M. B., Broadbelt K. G., Watanabe M., Kang Y. J., et al. (2015) Spatially heterogeneous choroid plexus transcriptomes encode positional identity and contribute to regional CSF production. J Neurosci 35: 4903-4916

Kao C. Y., He Z., Zannas A. S., Hahn O., Kuhne C., et al. (2016) Fluoxetine treatment prevents the inflammatory response in a mouse model of posttraumatic stress disorder. J Psychiatr Res 76: 74-83.
The molecular mechanisms that affect posttraumatic stress disorder (PTSD) are still elusive. In this study, the authors performed RNA-Seq on samples from prelimbic cortex, anterior cingulate cortex (ACC), basolateral amygdala, central nucleus of amygdala, nucleus accumbens (NAc), and the CA1 area of the dorsal hippocampus between a mouse model of PTSD (induced by foot shock) and controls. They identified differentially expressed genes that were then clustered for *in silico* pathway analysis; genes from 11 clusters included increased inflammatory response in ACC from shocked mice. Microglial activation was also higher in ACC of shocked mice. Chronic fluoxetine treatment initiated in the aftermath of the trauma prevented inflammatory gene expression alterations in ACC and mitigated PTSD-like symptoms. These results suggest a role for inflammation in PTSD pathobiology, as well as therapeutic applications for anti-inflammatory agents.

Illumina Technology: HiSeq 2000

Angermueller C., Clark S. J., Lee H. J., Macaulay I. C., Teng M. J., et al. (2016) Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. Nat Methods Teng M. J., et al. (2016.
scM&T is a method for parallel genome-wide methylome and transcriptome sequencing. It allows for the discovery of associations between transcriptional and epigenetic variation. The authors cultured E14 mouse ESCs and collected single cells by flow cytometry. They used genome and transcriptome sequencing (G&T-Seq) to purify DNA in order to prepare single-cell bisulfite libraries (scBS-Seq), and prepared RNA-Seq libraries from the single-cell cDNA libraries using the Nextera XT kit. Multiplexed scBS-Seq and RNA-Seq libraries were then sequenced separately. Profiling of 61 mouse ESCs confirmed previously known links between DNA methylation and transcription. Additionally, it showed previously unrecognized associations between heterogeneously methylated distal regulatory elements and transcription of key pluripotency genes.

Illumina Technology: Nextera XT Sample Prep Kit, HiSeq 2000

Castro-Mejia J., Jakesevic M., Krych L., Nielsen D. S., Hansen L. H., et al. (2016) Treatment with a Monoclonal Anti-IL-12p40 Antibody Induces Substantial Gut Microbiota Changes in an Experimental Colitis Model. Gastroenterol Res Pract 2016: 4953120.
Crohn's disease (CD) is associated with an imbalance in gut microbiota. The use of the anti-IL-12p40 monoclonal antibody (12p40-mAb) has therapeutic effects in patients with CD. This study investigated whether 12p40-mAb administration induces changes in the gut microbiome of mice with adoptive transfer colitis (AdTr-colitis). The authors treated AdTr-colitis mice with either 12p40-mAb, rat-IgG2a, or NaCl from days 21 to 47 from induction of colitis and monitored disease progression. The mice were sacrificed on day 47 and the gut microbiome was characterized by gel electrophoresis and 16S rRNA gene-amplicon high-throughput sequencing on DNA extracted from cecal, colonic, and fecal material. Treatment with 12p40-mAb reduced or eliminated most pathological parameters associated with colitis, and the gut microbiota had a higher Firmicutes/Bacteroides ratio compared to mice treated with rat-IgG2a. The authors also found significant correlations between 17 bacterial genera and biological markers.

Illumina Technology: Nextera Index Kit, MiSeq

Imperio C. G., McFalls A. J., Colechio E. M., Masser D. R., Vrana K. E., et al. (2015) Assessment of individual differences in the rat nucleus accumbens transcriptome following taste-heroin extended access. Brain Res Bull.
The factors that distinguish heroin users who fall into addiction from those who do not are still unknown. In this study, the authors performed whole genome RNA-Seq in cells from the nucleus accumbens (NAc) of rats classified into 3 groups based on addiction-like behaviors. The results revealed a number of differentially expressed genes involved in immunity, neuronal activity, and behavior.

Illumina Technology: TruSeq Stranded HT Library Preparation Kit, HiSeq 2500

Liu Z., Li X., Zhang J. T., Cai Y. J., Cheng T. L., et al. (2016) Autism-like behaviours and germline transmission in transgenic monkeys overexpressing MeCP2. Nature 530: 98-102.
Methyl-CpG binding protein 2 (MeCP2) is crucial in transcriptional regulation and miRNA processing. Mutations in its encoding gene, *MECP2*, are associated with autism-spectrum disorders. The mouse model of MeCP2 expression has met with challenges in identifying autistic behaviors. In this study, the authors used lentivirus-based transgenic cynomolgus monkeys (*Macaca fascicularis*) to study the effects of MeCP2 expression in the brain. They used Western blotting and immunostaining of brain tissues to verify transgenic *MECP2* expression, and deep sequencing to identify the genomic integration sites of the transgene. Intracytoplasmic sperm injection with sperm from F_0 transgenic monkeys demonstrated germline transmission and Mendelian segregation of several *MECP2* transgenes in the F_1 progeny, where transgenic monkeys showed autistic-like behaviors compared to wild type (WT). These results established the utility of genetically engineered nonhuman primates in the study of brain disorders.

Illumina Technology: MiSeq

Camp J. G., Badsha F., Florio M., Kanton S., Gerber T., et al. (2015) Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. Proc Natl Acad Sci U S A 112: 15672-15677.

Cerebral organoids are 3D cultures of human cerebral tissue developed from pluripotent stem cells. Although they are emerging as models of cortical development, the extent to which organoids recapitulate neural development *in vivo* is still unclear. To clarify this mechanism, the authors used single-cell RNA-Seq to dissect and compare cell composition and progenitor-to-neuron lineage relationships in human cerebral organoids and fetal neocortex. Covariation network analysis on the fetal neocortex data revealed several interactions among genes central to neuroprogenitor proliferation and neural differentiation, some of which were previously known. In the organoid, they identified diverse progenitors and differentiated cell types of neuronal and mesenchymal lineages, as well as cells that derived from regions resembling the fetal neocortex. The identified organoid cortical cells use gene expression programs significantly similar to those of fetal tissue to organize into cerebral cortex–like regions.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, HiSeq 2500

Xiao Y., Liu F., Zhang Z., Tang J., Zou C. G., et al. (2016) Gut-Colonizing Bacteria Promote *C. elegans* Innate Immunity by Producing Nitric Oxide. Cell Rep 14: 1301-1307.

Several scientific results are demonstrating the importance of commensal gut bacteria to immunity and metabolism. The mechanisms through which these microorganisms influence the host's health, however, are unknown. The authors investigated bacterial diversity in the gut of *Caenorhabditis elegans*. They used culture-independent Illumina MiSeq sequencing of the bacterial 16S rRNA gene amplicons and demonstrated that the *C. elegans* gut microbiota is different from the biota in the soil. *Bacillus subtilis* is the most frequent species in the worm's gut. By producing nitric oxide (NO), this bacterium confers resistance to pathogenic bacteria such as *Pseudomonas aeruginosa*, *Salmonella enterica*, and *Enterococcus faecalis*. Deletion of the gene that produces NO synthetase, *nos*, reduces the protective effect promoting immune responses to *P. aeruginosa*.

Illumina Technology: MiSeq

Dooley J., Tian L., Schonefeldt S., Delghingaro-Augusto V., Garcia-Perez J. E., et al. (2016) Genetic predisposition for beta cell fragility underlies type 1 and type 2 diabetes. Nat Genet.

It is known that T1D and type 2 diabetes (T2D) share pathophysiological characteristics, yet links at the mechanistic level are still elusive. T1D is caused by autoimmunity against pancreatic beta cells, resulting in insulin deficiency, while T2D is initiated by gradual metabolic changes that render target tissues resistant to insulin. In this study, the authors used the properties of the insHEL transgene in nonobese diabetic (NOD) mice as a sensitizer for beta cell failure. They found that its diabetes-inducing properties result from unfolded protein stress, rather than immunological effects. Through genotyping, they identified genetic variation in *Xrcc4* and *Glis3* genes that altered the response of NOD beta cells to unfolded proteins and enhanced the apoptotic and senescent fates. RNA-Seq of isolated islets revealed that the same transcriptional relationship was observed in human islets, demonstrating the role of beta cell fragility in genetic predisposition to diabetes.

Illumina Technology: BeadChip, HiSeq 2000

Stem Cells

Stem cells are undifferentiated cells that are able to divide and differentiate into specialized cells. Each stem cell chooses a specific lineage determined by its genetic code and response to the environment, giving rise to an array of unique, heterogeneous populations of cells.[110] Given their regenerative potential, stem cells are promising candidate treatments for certain diseases where tissue is lost, such as diabetes and heart disease.[111]

110. Voet T., Kumar P., Van Loo P., Cooke S. L., Marshall J., et al. (2013) Single-cell paired-end genome sequencing reveals structural variation per cell cycle. Nucleic Acids Res 41: 6119-6138

111. Stem Cell Basics – Introduction. National Institutes of Health website. http://stemcells.nih.gov/info/basics/pages/basics1.aspx. Updated March 5, 2015. Accessed August 15, 2016

ESCs are found in the inner mass of the blastocyst during embryonal development. ESCs are defined as pluripotent, as they can differentiate into any cell type, and they derive from eggs that have been fertilized *in vitro* [112, 113, 114, 115] (Figure 11). Adult stem cells are relatively undifferentiated cells found in certain tissues, such as adipose tissue, bone marrow, and blood.[116, 117] Here, they can divide and self-regenerate, and replace cells that are lost either physiologically or pathologically. Adult stem cells are multipotent, as they have the ability to differentiate into different cell types but are often limited to certain cell types.[118]

A cell's fate is not unidirectional, but the appropriate stimuli can reprogram a differentiated cell and induce it to regain pluripotency. Induced pluripotent stem cells (iPSCs) are adult cells that have been reprogrammed to an embryonic pluripotent stage through the use of an appropriate transcription factor cocktail.[119] Human iPSCs are capable of generating cells belonging to all 3 germ layers,[120] and they are currently used in drug development [121, 122, 123, 124, 125] and disease modeling.[126, 127, 128, 129, 130]

Although the use of stem cells in therapeutic applications holds great hope for the future, additional research is needed in this field. The National Institute of Health (NIH) has a stem cell unit for this purpose and provides resources such as stem cell libraries and projects,[131] and the NIH Human Embryonic Stem Cell Registry.[132] The Regenerative Medicine Program aims at accelerating the development of new medical applications for cell-based approaches.[133]
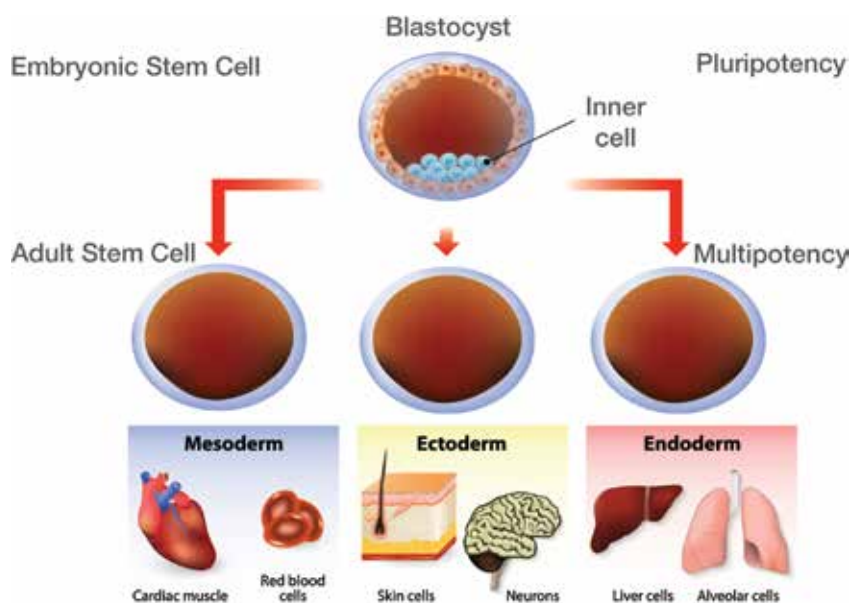

Figure 11. Stem cells and cell potency.

112. Thomson J. A., Itskovitz-Eldor J., Shapiro S. S., Waknitz M. A., Swiergiel J. J., et al. (1998) Embryonic stem cell lines derived from human blastocysts. Science 282: 1145-1147

113. De Coppi P., Bartsch G., Jr., Siddiqui M. M., Xu T., Santos C. C., et al. (2007) Isolation of amniotic stem cell lines with potential for therapy. Nat Biotechnol 25: 100-106

114. International Stem Cell I., Adewumi O., Aflatoonian B., Ahrlund-Richter L., Amit M., et al. (2007) Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. Nat Biotechnol 25: 803-816

115. Boyer L. A., Lee T. I., Cole M. F., Johnstone S. E., Levine S. S., et al. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122: 947-956

116. Becker A. J., Mc C. E. and Till J. E. (1963) Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. Nature 197: 452-454

117. Jiang Y., Jahagirdar B. N., Reinhardt R. L., Schwartz R. E., Keene C. D., et al. (2002) Pluripotency of mesenchymal stem cells derived from adult marrow. Nature 418: 41-49

118. Adult Stem Cell. Science Daily website. http://www.sciencedaily.com/terms/adult_stem_cell.htm. Accessed August 15, 2016

119. Takahashi K. and Yamanaka S. (2016) A decade of transcription factor-mediated reprogramming to pluripotency. Nat Rev Mol Cell Biol 17: 183-193

120. What are Induced Pluripotent Stem Cells? National Institutes of Health website. http://stemcells.nih.gov/info/basics/pages/basics10.aspx. Updated March 4, 2015. Accessed August 15, 2016

121. Inoue H. and Yamanaka S. (2011) The use of induced pluripotent stem cells in drug development. Clin Pharmacol Ther 89: 655-661

122. Avior Y., Sagi I. and Benvenisty N. (2016) Pluripotent stem cells in disease modelling and drug discovery. Nat Rev Mol Cell Biol 17: 170-182

123. Ooi L., Sidhu K., Poljak A., Sutherland G., O'Connor M. D., et al. (2013) Induced pluripotent stem cells as tools for disease modelling and drug discovery in Alzheimer's disease. J Neural Transm (Vienna) 120: 103-111

124. Grskovic M., Javaherian A., Strulovici B. and Daley G. Q. (2011) Induced pluripotent stem cells--opportunities for disease modelling and drug discovery. Nat Rev Drug Discov 10: 915-929

125. Deng W. (2010) Induced pluripotent stem cells: paths to new medicines. A catalyst for disease modelling, drug discovery and regenerative therapy. EMBO Rep 11: 161-165

126. Ebert A. D., Liang P. and Wu J. C. (2012) Induced pluripotent stem cells as a disease modeling and drug screening platform. J Cardiovasc Pharmacol 60: 408-416

127. Soldner F. and Jaenisch R. (2012) Medicine. iPSC disease modeling. Science 338: 1155-1156

128. Wu S. M. and Hochedlinger K. (2011) Harnessing the potential of induced pluripotent stem cells for regenerative medicine. Nat Cell Biol 13: 497-505

Dong X., Chen K., Cuevas-Diaz Duran R., You Y., Sloan S. A., et al. (2015) Comprehensive Identification of Long Non-coding RNAs in Purified Cell Types from the Brain Reveals Functional LncRNA in OPC Fate Determination. PLoS Genet 11: e1005669.

This study examined the functions of long noncoding RNAs (lncRNAs) in the development of the central nervous system. The authors cultured mouse neural stem cells (NSCs) and differentiated them into oligodendrocyte precursor cells (OPCs). After 3 days of differentiation, they extracted RNA and performed RNA Seq as well as ChIP-Seq. By integrating transcription factor–binding and cell-type-specific transcriptomic data with the results from this analysis, they constructed a novel framework, useful for systematically identifying lncRNAs regulated during the development of OPCs from NSCs. Through this framework, they identified several candidates, the first of which is lnc-OPC.

Illumina Technology: HiSeq2000

Muraro P. A., Robins H., Malhotra S., Howell M., Phippard D., et al. (2014) T cell repertoire following autologous stem cell transplantation for multiple sclerosis. J Clin Invest 124: 1168-1172.

Autologous hematopoietic stem cell transplantation (HSCT) is a common application in cancer. In this study, the authors assessed the feasibility of this application in an autoimmune context by applying high-throughput deep sequencing of the T-cell receptor β chain (TCRβ) in a sample of 25 poor-prognosis multiple sclerosis patients. They found that HSCT has different effects in CD4+ and CD8+ T-cell repertoires. In the former, dominant TCRs present before treatment were undetectable after reconstitution. In contrast, dominant CD8+ clones were not efficiently removed, and the reconstituted repertoire originated from clonal expansion of cells that were present before the treatment. Importantly, TCR diversity was lower in patients that failed to respond to treatment.

Illumina Technology: MiSeq

Huang K., Shen Y., Xue Z., Bibikova M., April C., et al. (2014) A panel of CpG methylation sites distinguishes human embryonic stem cells and induced pluripotent stem cells. Stem Cell Reports 2: 36-43.

Currently, there is debate about whether human iPSCs are epigenetically identical to human ESCs. To investigate this issue, the authors analyzed methylation patterns in 114 human iPSCs and compared them to the methylation patterns from 155 human ESCs. They found 82 CpG methylation sites that can distinguish iPSCs from ESCs. Of these, 12 were subject to hypermethylation, in part by the protein DNA (cytosine-5-)-methyltransferase 3 beta, DNMT3B. The authors conclude that DNMT3B is involved in *de novo* methylation during reprogramming, partly contributing to the unique iPSC signature in human cells.

Illumina Technology: Infinium, BeadChip

Maza I., Caspi I., Zviran A., Chomsky E., Rais Y., et al. (2015) Transient acquisition of pluripotency during somatic cell transdifferentiation with iPSC reprogramming factors. Nat Biotechnol 33: 769-774.

Recent publications suggest a transdifferentiation method to differentiate fibroblasts into various mature somatic cell types by brief expression of the iPSC reprogramming factors Oct4, Sox2, Klf4, and c-Myc (OSKM) and cell expansion in media that promotes differentiation. In this study, the authors tested this method using RNA-Seq and bisulfite sequencing on murine cells. They performed genetic lineage tracing for expression of endogenous Nanog and Oct4 and for chromosome X reactivation, as these events mark the acquisition of pluripotency. They found that the vast majority of reprogrammed NSCs or cardiomyocytes pass through a transient pluripotent state, and their derivation is coupled to iPSC formation mechanisms at the molecular level.

Illumina Technology: TruSeq RNA Sample Preparation Kit v2, HiSeq 1500

129. Sterneckert J. L., Reinhardt P. and Scholer H. R. (2014) Investigating human disease using stem cell models. Nat Rev Genet 15: 625-639

130. Kim C. (2014) Disease modeling and cell based therapy with iPSC: future therapeutic option with fast and safe application. Blood Res 49: 7-14

131. NIH Gene Libraries and Projects. National Institutes of Health website. http://stemcells.nih.gov/research/pages/genes.aspx. Updated April 3, 2015. Accessed August 15, 2016

132. NIH Human Embryonic Stem Cell Registry. National Institutes of Health website. http://grants.nih.gov/stem_cells/registry/current.htm. Accessed August 15, 2016

133. NIH Common Fund. National Institutes of Health website. http://commonfund.nih.gov/stemcells/index. Updated July 12, 2016. Accessed August 15, 2016

For Research Use Only. Not for use in diagnostic procedures.

Single-Cell Sequencing

The existence of diverse cell types has been known from early studies on cell morphology.[134] However, recent studies are revealing the remarkably high diversity of individual states, even within cells belonging to the same type.[135] Although individual cells of the same phenotype are commonly seen as identical functional units of a tissue or an organ, deep sequencing from single cells is uncovering a more complex heterogeneity of cell states.[136]

Single-cell technologies are instrumental in characterizing cell type and state,[137] detecting correlations between genetic traits and phenotype at the single-cell level, analyzing a cell's response to environmental stimuli, detecting underrepresented cell types in disease, and studying cellular interactions in the immune response.[138]

For example, studies on single stem cells in different phases of differentiation have helped in elucidating how a stem cell responds to its environment, and how it chooses a specific lineage.[139, 140, 141, 142, 143] Also, as individual neurons have mosaic genomes that exhibit CNVs even between cells from the same region,[144] the analysis of single neurons is of interest not only to understand brain function but also neurological and psychological disorders.[145, 146, 147, 148] Similarly, single-cell sequencing can be instrumental in studying the molecular mechanisms of immunity as well as in identifying disease correlates and immunological interventions.[149]

Macosko E. Z., Basu A., Satija R., Nemesh J., Shekhar K., et al. (2015) Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 161: 1202-1214.

Drop-Seq analyzes mRNA transcripts from droplets of individual cells in a highly parallel fashion. This single-cell sequencing method utilizes a microfluidic device to compartmentalize droplets containing a single cell, lysis buffer, and a micro-bead covered with barcoded primers. Each primer contains: 1) a 30 bp oligo(dT) sequence to bind mRNAs; 2) an 8 bp molecular index to identify each mRNA strand uniquely; 3) a 12 bp barcode unique to each cell; and 4) a universal sequence identical across all beads. Following compartmentalization, cells in the droplets are lysed and the released mRNA hybridize to the oligo(dT) tract of the primer beads. All droplets are then pooled and broken to release the beads. After the beads are isolated, they are reverse-transcribed with template-switching. This reaction generates the first cDNA strand with a PCR primer sequence in place of the universal sequence. cDNAs are PCR-amplified and sequencing adapters prepared using the Nextera XT Library Prep Kit. Barcoded mRNA samples are ready for sequencing.

Illumina Technology: HiSeq 2000

134. Alberts B., Johnson A., Lewis J., Raff M., Roberts K., et al. (2015) Molecular Biology of the Cell.

135. Eberwine J., Sul J. Y., Bartfai T. and Kim J. (2014) The promise of single-cell sequencing. Nat Methods 11: 25-27

136. Dueck H., Eberwine J. and Kim J. (2016) Variation is function: Are single cell differences functionally important?: Testing the hypothesis that single cell variation is required for aggregate function. Bioessays 38: 172-180

137. Trapnell C. (2015) Defining cell types and states with single-cell genomics. Genome Res 25: 1491-1498

138. Dura B., Dougan S. K., Barisa M., Hoehl M. M., Lo C. T., et al. (2015) Profiling lymphocyte interactions at the single-cell level by microfluidic cell pairing. Nat Commun 6: 5940

139. Macaulay I. C. and Voet T. (2014) Single cell genomics: advances and future perspectives. PLoS Genet 10: e1004126

140. Treutlein B., Brownfield D. G., Wu A. R., Neff N. F., Mantalas G. L., et al. (2014) Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature 509: 371-375

141. Streets A. M., Zhang X., Cao C., Pang Y., Wu X., et al. (2014) Microfluidic single-cell whole-transcriptome sequencing. Proc Natl Acad Sci U S A 111: 7048-7053

142. Guo H., Zhu P., Wu X., Li X., Wen L., et al. (2013) Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. Genome Res 23: 2126-2135

143. Xue Z., Huang K., Cai C., Cai L., Jiang C. Y., et al. (2013) Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature 500: 593-597

144. McConnell M. J., Lindberg M. R., Brennand K. J., Piper J. C., Voet T., et al. (2013) Mosaic copy number variation in human neurons. Science 342: 632-637

Figure 12. Overview of the Drop-Seq method.

Bjorklund A. K., Forkel M., Picelli S., Konya V., Theorell J., et al. (2016) The heterogeneity of human CD127 innate lymphoid cells revealed by single-cell RNA sequencing. Nat Immunol.

Innate lymphoid cells (ILCs) are important participants in homeostasis and inflammation. Heterogeneity and inflammation are characteristics of ILCs. In this study, the authors performed RNA-Seq on hundreds of individual tonsil CD127+ and natural killer (NK) cells. Through unbiased transcriptional clustering, they identified 4 distinct populations: ILC1, ILC2, ILC3, and NK. Furthermore, the high resolution of single-cell transcriptome sequencing allowed them to identify 3 transcriptionally and functionally diverse subpopulations for the ILC3 group. These results provide new insights into ILC biology in homeostasis and also have implications for the dysregulation of the immune system.

Illumina Technology: HiSeq 2000

Brennecke P., Reyes A., Pinto S., Rattay K., Nguyen M., et al. (2015) Single-cell transcriptome analysis reveals coordinated ectopic gene-expression patterns in medullary thymic epithelial cells. Nat Immunol 16: 933-941.

Expression of tissue-restricted self-antigens (TRAs) in medullary thymic epithelial cells (mTECs) plays an important role in the induction of self-tolerance and in the prevention of autoimmunity. Each TRA is expressed in only a few mTECs; however, the regulation of this process in single cells and its coordination at the population level is poorly understood. To identify recurrent TRAs, the authors performed single-cell population sequencing in a cohort of C57BL/6 mice, from which they isolated medullary thymic epithelial cells and found multiple TRA coexpression patterns. They also demonstrated chromatin accessibility in coexpressed genes clustered in the genome with an assay for transposase-accessible chromatin (ATAC-Seq) on thymic tissue from children undergoing cardiac surgery. These results suggest that TRA expression in mTECs is a coordinated process, possibly involving chromatin remodeling.

Illumina Technology: HiSeq 2500

Tipton C. M., Fucile C. F., Darce J., Chida A., Ichikawa T., et al. (2015) Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. Nat Immunol 16: 755-765.

The pathogenesis of acute SLE is mediated by antibody-secreting cells (ASCs). However, the origin of ASCs, as well as their diversity and contribution to autoantibodies, remain unknown. The authors performed deep sequencing, proteomic profiling of autoantibodies, and single-cell analysis. The results demonstrated highly diversified ASCs, punctuated by clones expressing the variable heavy-chain regionVH4-34 that produced serum autoantibodies. A fraction of ASC clones contained autoantibodies without mutations, suggesting that the differentiation might happen outside of the germinal centers. A substantial ASC segment was derived from a distinct subset of newly activated naïve cells of considerable clonality that persisted in the circulation for several months. The authors conclude, from these results, that selection of SLE autoreactivities occurs during polyclonal activation, with prolonged recruitment of recently activated naïve B cells.

Illumina Technology: MiSeq, Nextera XT DNA Sample Prep Kit

Stubbington M. J., Lonnberg T., Proserpio V., Clare S., Speak A. O., et al. (2016) T cell fate and clonality inference from single-cell transcriptomes. Nat Methods .

The authors developed a computational method that allows the reconstruction of full-length, paired TCR sequences from T lymphocyte single-cell RNA-Seq data. The analysis links T cell specificity with functional response by revealing clonal relationships among cells, along with their transcriptional profiles. The authors applied this method to a mouse Salmonella infection model and found that T cell clonotypes span early activated CD4+ T cells, as well as mature effector and memory cells.

Illumina Technology: Nextera XT DNA Sample Prep Kit, HiSeq 2500, MiSeq

Lossius A., Johansen J. N., Vartdal F. and Holmoy T. (2016) High-throughput sequencing of immune repertoires in multiple sclerosis. Annals of CLinical and Translational Neurology .

T cells and B cells play crucial roles in the initiation and development of multiple sclerosis (MS). The activation of these cells is likely mediated through recognition of antigens by T- and B-cell receptors, which are highly polymorphic due to recombination and somatic mutation. In this review, the authors summarize studies that used earlier methods to explore T- and B-receptor repertoires. They describe how NGS has provided new knowledge of MS by analyzing these repertoires in higher detail and depth.

145. Iourov I. Y., Vorsanova S. G. and Yurov Y. B. (2012) Single cell genomics of the brain: focus on neuronal diversity and neuropsychiatric diseases. Curr Genomics 13: 477-488

146. Evrony G. D., Cai X., Lee E., Hills L. B., Elhosary P. C., et al. (2012) Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. Cell 151: 483-496

147. Poduri A., Evrony G. D., Cai X. and Walsh C. A. (2013) Somatic mutation, genomic variation, and neurological disease. Science 341: 1237758

148. Eberwine J. and Bartfai T. (2011) Single cell transcriptomics of hypothalamic warm sensi-tive neurons that control core body tempera-ture and fever response Signaling asymmetry and an extension of chemical neuroanatomy. Pharmacol Ther 129: 241-259

149. Chattopadhyay P. K., Gierahn T. M., Roederer M. and Love J. C. (2014) Single-cell technol-ogies for monitoring immune systems. Nat Immunol 15: 128-135

DeKosky B. J., Kojima T., Rodin A., Charab W., Ippolito G. C., et al. (2015) In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. Nat Med 21: 86-91.
Immune repertoire sequencing is a critical step in understanding adaptive responses, with important applications in the fields of infectious and autoimmune diseases. However, the determination of native antibody variable heavy-light pairs (VH-VL) is a challenging task. In this study, the authors developed a low-cost, single-cell, emulsion-based technology for the sequencing of antibody VH-VL repertoires from > 2 × 106 B cells per experiment, with demonstrated pairing precision > 97%. The application of this analysis to 3 human volunteers provided new insights, including: 1) the identity, frequency, and pairing propensity of shared VL genes; 2) the detection of allelic inclusion in healthy individuals (an autoimmune mechanism); and 3) the occurrence of features associated with broadly neutralizing antibodies to rapidly evolving viruses.

Illumina Technology: MiSeq

Eugster A., Lindner A., Catani M., Heninger A. K., Dahl A., et al. (2015) High diversity in the TCR repertoire of GAD65 autoantigen-specific human CD4+ T cells. J Immunol 194: 2531-2538.
This study offers insights into TCR diversity in T1D. The authors collected samples from 6 patients with T1D and 10 islet autoantibody–positive children to apply single-cell TCR α- and β-chain sequencing. For 1650 GAD65-specific CD4+ cells isolated from GAD65 proliferation assays and/or GAD 5571 tetramer staining, they identified 1003 TCRs, demonstrating a high diversity. There was a limited overlap (< 5%) between TCRs of GAD65- and GAD65 5571 tetramer+ and CD4+ T cells. Few TCRs were repeatedly found in GAD65-specific cells at different time points for an individual patient, and no TCR was observed in more than 1 patient. These results fail to provide strong support for TCR-targeted therapies in T1D.

Illumina Technology: MiSeq, HiSeq 2500

Hoh R. A., Joshi S. A., Liu Y., Wang C., Roskin K. M., et al. (2016) Single B-cell deconvolution of peanut-specific antibody responses in allergic patients. J Allergy Clin Immunol 137: 157-167.
The cellular bases of food allergies are not fully understood. Frequencies, cellular phenotypes, epitope specificity, and clonal diversity of allergen-specific B cells are of major pathogenic and therapeutic significance. In this study, the authors extracted B cells binding fluorescently labeled Ara h1 or Ara h2 allergens from 18 allergic patients at baseline, 13 patients undergoing therapy, and 9 healthy controls and isolated them with flow-cytometric sorting. They used deep sequencing of the B cell repertoires, as well as protein quantification techniques, to identify members of the allergen-specific clones from B cells. Median allergen-binding B-cell frequencies were 0.0097% (Ara h1) and 0.029% (Ara h2) for B cells in blood from baseline patients and were 3-fold higher during therapy. Five of 57 allergen-specific cells belonged to clones containing IgE-expressing members. Almost all allergen-specific antibodies were mutated, and binding to both conformational and linear allergen epitopes was detected. Increasing somatic mutations of IgG4 members of a clone was observed in immunotherapy, while IgE mutation levels in the clone did not increase. The authors conclude that most peanut allergen–binding B cells express mutated and isotype-switched antibodies. Further, immunotherapy increases their frequency in the blood, and even narrowly defined allergen epitopes are recognized by numerous distinct B cell clones in a patient. They also suggest that oral immunotherapy can stimulate somatic mutation of allergen-specific IgG4.

Illumina Technology: MiSeq

Linnemann C., Heemskerk B., Kvistborg P., Kluin R. J., Bolotin D. A., et al. (2013) High-throughput identification of antigen-specific TCRs by TCR gene capture. Nat Med 19: 1534-1541.
TCR transfer into a patient's T cells is a promising therapy for both viral infections and cancer. In this study, the authors developed a high-throughput DNA-based strategy to identify TCR sequences by capturing and sequencing genomic DNA fragments encoding TCR genes. They tested this approach by assembling a large library of cancer germline tumor antigen–reactive TCRs. They also demonstrated the feasibility of identifying antigen-specific TCRs in oligoclonal T cell populations, in both human and TCR-humanized murine cells. Finally, they demonstrated the ability to identify tumor-reactive TCRs within intratumoral T cell subsets without knowledge of antigen specificities.

Illumina Technology: TruSeq Sample Preparation Kit, HiSeq 2000

Evrony G. D., Lee E., Mehta B. K., Benjamini Y., Johnson R. M., et al. (2015) Cell lineage analysis in human brain using endogenous retroelements. Neuron 85: 49-59.

Somatic mutations in neural cells occur during brain development and are increasingly implicated as causes of neurogenetic diseases. This study examined spontaneous somatic mutations as clonal marks to track cell lineages in human brain. The authors used high-coverage whole-genome sequencing of postmortem single neurons collected from a healthy 17 year old individual. Somatic mutation analyses in more than 30 locations throughout the nervous system identified multiple lineages and sublineages of cells marked by different LINE-1 (L1) retrotransposition events and subsequent mutation of poly(A) microsatellites within L1. One clone contained thousands of cells limited to the left middle frontal gyrus, while a second clone contained millions of cells distributed over the entire left hemisphere. These patterns mirror those observed for somatic mutation disorders of brain development and suggest that focally distributed mutations are also prevalent in normal brains.

Illumina Technology: HiSeq 2000

Darmanis S., Sloan S. A., Zhang Y., Enge M., Caneda C., et al. (2015) A survey of human brain transcriptome diversity at the single cell level. Proc Natl Acad Sci U S A 112: 7285-7290.

The complexity of the human brain, in terms of comprised cell types, is vast. To explore it, the authors used single-cell RNA-Seq on 466 cells from adult and fetal brain. They extracted adult tissue from the temporal lobe of 8 patients undergoing epilepsy surgery and obtained fetal brain tissue during 4 elective abortions. Through RNA-Seq, the authors were able to classify single cells into the major neuronal, glial, and vascular cell types in the brain. They then classified neurons into individual communities and showed that these neurons preserve the categorization of interneuron subtypes typically observed with the use of classic markers. They applied RNA-Seq to fetal cortical neurons to compare gene expression profiles between adult and fetal neurons. Additionally, they identified those expression gradients that reflect the transition between replicating and quiescent neuronal populations. Finally, they observed the expression of MHC type I genes in a subset of adult neurons but not in fetal neurons.

Illumina Technology: Nextera XT DNA Sample Prep Kit, NextSeq

Fuzik J., Zeisel A., Mate Z., Calvigioni D., Yanagawa Y., et al. (2016) Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. Nat Biotechnol 34: 175-183.

The relationship between a neuron's molecular phenotype and parameters such as location, morphology, connectivity, and excitability is still unexplored. The authors developed a method, called Patch-Seq, to obtain full transcriptome data from single neocortical pyramidal cells and interneurons after whole-cell patch-clamp recordings in brain slices obtained from a mouse model. Briefly, after a patch-clamp stimulus, the procedure includes the aspiration of the entire somatic compartment into a recording pipette, reverse transcription of RNA including addition of unique molecular identifiers, cDNA amplification, library preparation, and sequencing. The application of the method revealed a close link between electrophysiological characteristics, responses to acute chemical challenges, and RNA expression of neurotransmitter receptors and channels. Moreover, the authors were able to distinguish established neuronal subpopulations as well as undescribed neuronal subtypes.

Illumina Technology: HiSeq 2000

Lovatt D., Ruble B. K., Lee J., Dueck H., Kim T. K., et al. (2014) Transcriptome *in vivo* analysis (TIVA) of spatially defined single cells in live tissue. Nat Methods 11: 190-196.

Transcriptome *in vivo* analysis (TIVA) is a protocol that captures mRNA from live cells. In this method, a TIVA tag is loaded into cells. Photoactivation cleaves photocleavable linkers, allowing the tag to hybridize to mRNA. The biotin-bound mRNA is captured using streptavidin-coated magnetic beads, transcribed into cDNA, and sequenced. Sequencing the cDNA provides single-cell transcriptome analysis from complex tissues.

Illumina Technology: HiSeq 1000

Usoskin D., Furlan A., Islam S., Abdo H., Lonnerberg P., et al. (2015) Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. Nat Neurosci 18: 145-153.

The primary sensory system is composed of multiple cell types, and its full complexity is unclear. In this study, the authors used transcriptome analysis of 622 single neurons from 6 mice, and classified them in an unbiased manner. Their results revealed a total of 11 types, confirming previously anticipated neuronal subtypes and also providing markers for new, functionally distinct subtypes.

Illumina Technology: Illumina GA$_{IIx}$, HiSeq 2000

Macaulay I. C., Svensson V., Labalette C., Ferreira L., Hamey F., et al. (2016) Single-Cell RNA-Sequencing Reveals a Continuous Spectrum of Differentiation in Hematopoietic Cells. Cell Rep 14: 966-977.

The authors analyzed the continuous nature of hematopoietic cell differentiation by applying single-cell RNA-Seq to a population of hematopoietic cells in zebrafish, as they underwent thrombocyte lineage commitment. By a computational developmental chronology reconstruction, they were able to place each cell along a continuum from stem cell to mature cell, refining the traditional lineage tree. Overall, the total number of genes expressed, as well as the total mRNA content of cells, decreases as cells undergo lineage commitment.

Illumina Technology: Nextera XT DNA Library Prep Kit, HiSeq 2000

Gaublomme J. T., Yosef N., Lee Y., Gertner R. S., Yang L. V., et al. (2015) Single-Cell Genomics Unveils Critical Regulators of Th17 Cell Pathogenicity. Cell 163: 1400-1412.

The authors studied the molecular mechanisms that govern heterogeneity and pathogenicity of Th17 cells isolated from the central nervous system (CNS) and lymph nodes (LN) at the peak of autoimmune encephalomyelitis (EAE), or differentiated *in vitro* under either pathogenic or nonpathogenic conditions. They performed single-cell RNA-Seq on murine Th17 cells. Through computational analysis, they relate a spectrum of cellular states *in vivo* to *in vitro* differentiated Th17 cells and unveil genes governing pathogenicity and disease susceptibility.

Illumina Technology: Nextera XT DNA Sample Prep Kit

Hanchate N. K., Kondoh K., Lu Z., Kuang D., Ye X., et al. (2015) Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. Science 350: 1251-1255.

Smell allows chemicals to be perceived as diverse scents. This study used single-neuron RNA-Seq on murine neurons to explore the developmental mechanisms that shape the ability to smell as nasal olfactory neurons mature. The authors found that most mature neurons expressed only 1 out of approximately 1000 odorant receptor genes (*Olfrs*) available, and at a high level. In contrast, several immature neurons expressed low levels of multiple *Olfrs*. The coexpression of Olfrs localized to overlapping zones of the nasal epithelium suggested regional biases but not to single genomic loci. A single immature neuron was able to express *Olfrs* from up to 7 different chromosomes. The mature state, in which expression of *Olfrs* genes is restricted to 1 per neuron, emerges over a developmental progression that appears to be independent of neuronal activity involving sensory transduction molecules.

Klein A. M., Mazutis L., Akartuna I., Tallapragada N., Veres A., et al. (2015) Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161: 1187-1201.

This study describes a newly developed high-throughput droplet-microfluidic approach for barcoding the RNA from thousands of single cells for subsequent NGS analysis. The authors report that the method has very low noise and can be adapted to different sequence-based assays. They applied it to mouse ESCs, revealing in detail the population structure and the heterogeneous onset of differentiation after withdrawal of leukemia inhibitory factor. The reproducibility of the high-throughput single-cell data enabled the authors to deconstruct cell populations and infer gene expression relationships.

Illumina Technology: MiSeq

Paul F., Arkin Y., Giladi A., Jaitin D. A., Kenigsberg E., et al. (2015) Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. Cell 163: 1663-1677.

Bone marrow stem cells differentiate and give rise to diverse blood cell types. This study comprehensively mapped myeloid progenitor subpopulations by transcriptional sorting of single cells from the bone marrow of a murine model. The authors described multiple progenitor subgroups, showing transcriptional priming toward 7 differentiation fates but no progenitors with a mixed state. Transcriptional differentiation correlated with combinations of known and previously undefined transcription factors, suggesting a tight regulation of the process.

Illumina Technology: HiSeq 1500, NextSeq 500

Tasic B., Menon V., Nguyen T. N., Kim T. K., Jarsky T., et al. (2016) Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat Neurosci 19: 335-346.

The nervous system is composed of various cell types. To investigate the extent of such cell diversity, the authors performed single-cell RNA-Seq in mouse neurons, constructing a cellular taxonomy of the primary visual cortex. They identified 49 transcriptomic cell types, including 23 GABAergic, 19 glutamatergic, and 7 nonneuronal types. They also analyzed cell-specific mRNA processing. Finally, they found that some of their transcriptomic cell types displayed specific and differential electrophysiological and axon projection properties, providing evidence that single cell-transcriptomic signatures can be associated with specific cellular properties.

Illumina Technology: Nextera XT DNA Library Prep Kit, MiSeq, HiSeq 2000, HiSeq 2500

For Research Use Only. Not for use in diagnostic procedures.

30 Genomic solutions for cell biology and complex disease research

THE ROLE OF THE ENVIRONMENT

A major challenge in the field of complex diseases has always been quantifying how much genetic factors contribute to disease, compared to the environment.[150] The environment and genetics have been working side by side since the origin of life, and the organisms that are alive today are the result of interactions between environmental factors and genetics.[151]

In fact, environmental factors can be genetic factors themselves (as in the case of metagenomics),[152, 153] or have a direct effect on genomics or epigenomics (as in the case of chemicals and stress).[154] Other factors have shaped human genetics in the course of evolution (e.g., diet and metabolic genes, pathogens, and immunity-related genes).[155, 156] For these reasons, NGS is a powerful tool in the study of environmental factors, as demonstrated by research in the fields of agrigenomics, metagenomics, and epigenomics.

150. Polderman T. J., Benyamin B., de Leeuw C. A., Sullivan P. F., van Bochoven A., et al. (2015) Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nat Genet 47: 702-709

151. Anderson J. T., Wagner M. R., Rushworth C. A., Prasad K. V. and Mitchell-Olds T. (2014) The evolution of quantitative traits in complex environments. Heredity (Edinb) 112: 4-12

152. Wang Z., Gerstein M. and Snyder M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57-63

153. Moore-Connors J. M., Dunn K. A., Bielawski J. P. and Van Limbergen J. (2016) Novel Strategies for Applied Metagenomics. Inflamm Bowel Dis 22: 709-718

154. Stankiewicz A. M., Swiergiel A. H. and Lisowski P. (2013) Epigenetics of stress adaptations in the brain. Brain Res Bull 98: 76-92

155. Leonard W. R., Snodgrass J. J. and Robertson M. L. (2010) Fat Detection: Taste, Texture, and Post Ingestive Effects. Fat Detection: Taste, Texture, and Post Ingestive Effects.

156. Flajnik M. F. and Kasahara M. (2010) Origin and evolution of the adaptive immune system: genetic events and selective pressures. Nat Rev Genet 11: 47-59

Figure 13. The role of genetics in complex diseases is not limited only to human genetics and predisposition to disease. Genetics and the environment are a continuum, and they interact and influence each other at many different levels.

Tran P. V., Kennedy B. C., Pisansky M. T., Won K. J., Gewirtz J. C., et al. (2016) Prenatal Choline Supplementation Diminishes Early-Life Iron Deficiency-Induced Reprogramming of Molecular Networks Associated with Behavioral Abnormalities in the Adult Rat Hippocampus. J Nutr 146: 484-493.
Early-life iron deficiency is associated with several neurological disorders. These deficits are recapitulated by the rat model of diet-induced fetal-neonatal iron deficiencies. In this study, the authors examined whether early-life iron deficiency permanently reprograms the hippocampal transcriptome and assessed the effects of maternal dietary choline supplementation. They provided pregnant and nursing dams an iron-deficient (ID) diet in order to obtain ID pups. Controls were obtained by dams that were provided an iron-sufficient (IS) diet. Choline was provided to half of the dams in each group between gestational days 11 and 18. Hippocampal transcriptomes were assayed by RNA-Seq at postnatal day 65 and data were processed with knowledge-based Ingenuity Pathway Analysis. They observed that ID rats had altered hippocampal expression of 619 genes, many of which mapped to molecular networks implicated in psychological disorders, including autism, AD, and schizophrenia.

Illumina Technology: TruSeq RNA Sample Preparation v2 Kit, HiSeq 2000

Urak K. T., Shore S., Rockey W. M., Chen S. J., McCaffrey A. P., et al. (2016) In vitro RNA SELEX for the generation of chemically-optimized therapeutic RNA drugs. Methods.
Aptamers are single-stranded oligonucleotides that can bind with high affinity and specificity to target molecules. They are often referred as "nucleic acid antibodies" and are commonly obtained by a chemical process known as systematic evolution of ligands by exponential enrichment (SELEX). This method is a chemical equivalent of Darwinian evolution and was first described in 1990. Since then, it has yielded aptamers for a wide range of applications, including *in vitro* diagnostics, biomarker discovery, and therapeutics. In this study, the authors outline the key steps of the SELEX process, enabling the rapid identification of RNA aptamers for *in vivo* applications. They also discuss methods for performing NGS of the RNAs from each round of selection.

For Research Use Only. Not for use in diagnostic procedures.

Wang W., Jovel J., Halloran B., Wine E., Patterson J., et al. (2015) Metagenomic analysis of microbiome in colon tissue from subjects with inflammatory bowel diseases reveals interplay of viruses and bacteria. Inflamm Bowel Dis 21: 1419-1427.
The current model for inflammatory bowel diseases (IBD) suggests that genetically susceptible patients develop intolerance to gut microflora and chronic inflammation develops as a result of environmental insults. To study the contribution of viral infection in the pathogenesis of IBD, the authors performed a metagenomics analysis to document the basic virome in RNA extracted from colonic biopsies from 10 IBD patients and 5 controls. They found differences in gut microflora and in the abundance of mammalian viruses and human endogenous retroviruses. Specifically, patients with Herpesviridae sequences in their colon demonstrated increased expression of human endogenous viral sequences and differences in the diversity of their microbiome.

Illumina Technology: HiSeq 2000

He B., Nohara K., Ajami N. J., Michalek R. D., Tian X., et al. (2015) Transmissible microbial and metabolomic remodeling by soluble dietary fiber improves metabolic homeostasis. Sci Rep 5: 10604.
Dietary management can help improve metabolic diseases by promoting health benefits not yet totally characterized. Digestion-resistant plant-derived fibers, like maltodextrin (RM), improve glucose and lipid homeostasis and help reduce weight gain. The authors fed RM to obese mice models and tested its effects in glucose tolerance. They monitored gut microbiota changes by amplicon sequencing of the 16S V4 rRNA gene. RM improved glycemic control by decreasing fasting glucose levels and improved glucose control. RM induced beneficial gut microbiota remodeling by increasing abundance of beneficial bacteria (*Lactobacillus* and *Bifidobacterium*) and decreasing fat-associated bacteria (*Alistipes*). Fecal transplantation corroborated the positive effects of RM-remodeled gut microbiota, which was accompanied by metabolic changes, such as improved metabolism of cholesterol and glucose. The study sheds light on mechanisms underlying the beneficial effects of RM.

Illumina Technology: MiSeq, CASAVA® v1.8.3

Jones-Hall Y. L., Kozik A. and Nakatsu C. (2015) Ablation of tumor necrosis factor is associated with decreased inflammation and alterations of the microbiota in a mouse model of inflammatory bowel disease. PLoS One 10: e0119441.
The chronic inflammation characteristic of IBD is closely related to prolonged secretion of tumor necrosis factor (TNF). The effects of TNF in colitis and gut microbiota are not fully characterized, which may explain why anti-TNF therapy is not always successful. The authors assessed microbial composition by 16S V3-V4 rRNA amplicon sequencing in WT and tnf-/- mice (acute colitis model) to elucidate TNF effects in colitis and gut microbiota. Inflammation caused significant differences in microbiota composition according to mouse genotype, whereas absence of TNF resulted in milder colitis and less microbial   -diversity than in WT mice. The authors conclude that TNF-mediated inflammation modifies gut microbiota; therefore, combined therapies that inhibit TNF and alter microbial communities could be beneficial.

Morgan X. C., Kabakchiev B., Waldron L., Tyler A. D., Tickle T. L., et al. (2015) Associations between host gene expression, the mucosal microbiome, and clinical outcome in the pelvic pouch of patients with inflammatory bowel disease. Genome Biol 16: 67.
Ileal pouch-anal anastomosis (IPAA) surgery for ulcerative colitis (UC) is often complicated by pouchitis associated with anatomical and microbiota changes that resemble a colon-like environment. How and why these changes take place, and their relation to UC and IBD, is not fully understood. The authors obtained paired-host microbiomes (by 16S V4 amplicons sequencing) and transcriptomes from a large cohort of IPAA patients to study the microbiome-host gene expression axis. Microbiomes were variable across individuals and influenced by clinical variables such as antibiotic therapy, whereas host epithelial transcription was influenced by tissue location. Associations between microbiome and host transcription patterns were related to the level of host tissue inflammation: the strongest microbe-host association pattern was enriched in complement system and cytokine IL-12 pathways inversely correlated with the abundance of *Bifidobacteria* and others. However, it was not possible to generate a pouchitis outcome model based on microbial composition and/or transcriptional activity, suggesting that the role of sectional changes in epithelial transcripts may not be critical for the host-microbiome interface during IPAA.

Illumina Technology: MiSeq

Schaubeck M., Clavel T., Calasan J., Lagkouvardos I., Haange S. B., et al. (2015) Dysbiotic gut microbiota causes transmissible Crohn's disease-like ileitis independent of failure in antimicrobial defence.
Intestinal dysbiosis is associated with intestinal inflammatory disease (CD), although functional explanation of this frequent observation is still missing. In this study, the authors obtained temporal metaproteomic (by LC-MS) and metagenomic (by 16S rRNA gene sequencing) profiles of gut microbiota from TNF^{deltaARE} mice, a model that resembles CD pathology. Disease severity and location were microbiota-dependent, as it is absent in germ-free TNF^{deltaARE} mice and attenuated after antibiotic treatment. Several compositional and functional alterations were observed in microbiota communities in inflamed mice, features that were

reproducible through microbiome transplantation that resulted in CD-like ileitis accompanied by loss of Paneth cell function. The study provides evidence of causal role for gut dysbiosis in the development of chronic ileal inflammatory disease.

Illumina Technology: MiSeq

Chassaing B., Koren O., Goodrich J. K., Poole A. C., Srinivasan S., et al. (2015) Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome. Nature 519: 92-96.
Gut microbes are kept at a safe distance from the intestinal epithelium by layers of mucus. Emulsifiers disrupt this mucus layer *in vitro*, and their inclusion in processed foods may be a contributing factor for the increased incidence of metabolic diseases. The authors exposed gut tissue and microbiomes of mice to low concentrations of 2 common emulsifiers, carboxymethylcellulose (CMC) or polysorbate-80 (P80), and monitored them by 16S V4 rRNA sequencing. Both emulsifiers caused mild inflammation and obesity in WT mice, and colitis in a mouse sensitive model. Obesity was correlated with increased microbiota invasion of the epithelial layer, altered species composition, and increased proinflammatory potential. Fecal transplants into germ-free mice corroborated the role of this altered gut microbiota in the development of metabolic syndrome. The novel effects of common emulsifiers in the gut environment and their possible association with metabolic syndromes and inflammatory diseases need to be reconsidered in light of their increased incidence.

Illumina Technology: MiSeq 250

Silva P. E., Costa P. S., Avila M. P., Suhadolnik M. L., Reis M. P., et al. (2015) Leprous lesion presents enrichment of opportunistic pathogenic bacteria. Springerplus 4: 187.
M. leprae causes leprosy, a chronic disease that affects skin. Changes in skin microbiota composition caused by the disease have not been characterized in the past, but now high-throughput sequencing can address this question. This study examined the skin microbiota of healthy skin (previously published) and leprous lesions by 16S V3-V4 rRNA gene NGS and Sanger sequencing. Profound shifts in skin microbiota taxa composition and abundance were associated with leprous lesions that favor overgrowth of potentially pathogenic bacteria not usually associated with normal skin (*Burkhordelia*, *Pseudomonas*, and *Bacillus*) at the expense of normal flora (*Propionibacterium*, *Staphylococcus*, and *Corynebacterium*). This shift suggests that potentially pathogenic bacteria may have gained a competitive advantage over normal resident microbes in the environment provided by leprous lesions.

Illumina Technology: MiSeq

Watanabe Y., Arase S., Nagaoka N., Kawai M. and Matsumoto S. (2016) Chronic Psychological Stress Disrupted the Composition of the Murine Colonic Microbiota and Accelerated a Murine Model of Inflammatory Bowel Disease. PLoS One 11: e0150559.
Psychological stress has an effect on the gastrointestinal microbiota, and it might be associated with increased disease activity in IBD. This study investigated the relationship between psychological stress, gastrointestinal microbiota, and severity of colitis. The authors analyzed the impact of 12-week repeated water-avoidance stress on the microbiota of different strains of mice (TCRαKO on BALB/c background, TCRαKO on C57BL/6 background, and background controls) by NGS of bacterial 16S rRNA genes. They observed that, while the knockout of the TCRα gene caused a loss of gastrointestinal microbial diversity and stability in both strains, the chronic exposure to repeated water stress altered the composition of the colonic microbiota of C57BL/6 mice but not BALB/c mice. In C57BL/6 mice, species from the genus *Clostridium* were relatively abundant and weakly positively associated with colitis severity, an effect that was not seen in individuals with a relatively diverse microbiota. Exposure to stress also altered the concentration of free IgA in colonic contents, possibly affecting the loss of bacterial diversity in the colonic microbiota and the severity of the colitis exacerbation.

Illumina Technology: MiSeq

Afshinnekoo E., Meydan C., Chowdhury S., Jaroudi D., Boyer C., et al. (2015) Geospatial Resolution of Human and Bacterial Diversity with City-Scale Metagenomics. Cell Syst 1: 72-87.
Identifying the urban microbiome of dense populated areas is critical to identify potentially harmful disturbances. The authors geotagged swabs from 466 subway stations along 5 NYC boroughs were geotagged and performed metagenomic sequencing to establish the urban microbiome baseline of the NYC transit system. They found that 48% of the DNA did not match any known organism and identified 1688 taxa—bacterial (47%), viral (mostly phages, 0.03%), archaeal (0.003%), and eukaryotic (0.8%)—mostly associated with skin. Human DNA mirrored geospatial demographics of census data, whereas DNA from potential pathogens (*B. anthracis*, *Y. pestis*) was found in several stations. Antibiotic-resistance genes were present and active in cultured samples, and an abandoned subway station flooded by Hurricane Sandy still resembled a cold marine environment. The authors conclude this highly integrative microbiome study can be applied to cities to enable adequate responses to system perturbations.

Illumina Technology: TruSeq Nano DNA Library Prep Kit, HiSeq 2500, MiSeq

BIBLIOGRAPHY

Afshinnekoo E., Meydan C., Chowdhury S., Jaroudi D., Boyer C., et al. (2015) Geospatial Resolution of Human and Bacterial Diversity with City-Scale Metagenomics. Cell Syst 1: 72-87

Albert F. W. and Kruglyak L. (2015) The role of regulatory variation in complex traits and disease. Nat Rev Genet 16: 197-212

Alberts B., Johnson A., Lewis J., Raff M., Roberts K., et al. (2015) Molecular Biology of the Cell.

Anderson J. T., Wagner M. R., Rushworth C. A., Prasad K. V. and Mitchell-Olds T. (2014) The evolution of quantitative traits in complex environments. Heredity (Edinb) 112: 4-12

Angermueller C., Clark S. J., Lee H. J., Macaulay I. C., Teng M. J., et al. (2016) Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. Nat Methods

Avior Y., Sagi I. and Benvenisty N. (2016) Pluripotent stem cells in disease modelling and drug discovery. Nat Rev Mol Cell Biol 17: 170-182

Bakulski K. M., Dolinoy D. C., Sartor M. A., Paulson H. L., Konen J. R., et al. (2012) Genome-wide DNA methylation differences between late-onset Alzheimer's disease and cognitively normal controls in human frontal cortex. J Alzheimers Dis 29: 571-588

Bates G. P. (2005) History of genetic disease: the molecular genetics of Huntington disease - a history. Nat Rev Genet 6: 766-773

Becker A. J., Mc C. E. and Till J. E. (1963) Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. Nature 197: 452-454

Benfey P. N. and Mitchell-Olds T. (2008) From genotype to phenotype: systems biology meets natural variation. Science 320: 495-497

Bentham J., Morris D. L., Cunninghame Graham D. S., Pinder C. L., Tombleson P., et al. (2015) Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. Nat Genet 47: 1457-1464

Bjorklund A. K., Forkel M., Picelli S., Konya V., Theorell J., et al. (2016) The heterogeneity of human CD127 innate lymphoid cells revealed by single-cell RNA sequencing. Nat Immunol

Bonomi R., Mukhopadhyay U., Shavrin A., Yeh H. H., Majhi A., et al. (2015) Novel Histone Deacetylase Class IIa Selective Substrate Radiotracers for PET Imaging of Epigenetic Regulation in the Brain. PLoS One 10: e0133512

Boyer L. A., Lee T. I., Cole M. F., Johnstone S. E., Levine S. S., et al. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122: 947-956

Brennecke P., Reyes A., Pinto S., Rattay K., Nguyen M., et al. (2015) Single-cell transcriptome analysis reveals coordinated ectopic gene-expression patterns in medullary thymic epithelial cells. Nat Immunol 16: 933-941

Camp J. G., Badsha F., Florio M., Kanton S., Gerber T., et al. (2015) Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. Proc Natl Acad Sci U S A 112: 15672-15677

Castro-Mejia J., Jakesevic M., Krych L., Nielsen D. S., Hansen L. H., et al. (2016) Treatment with a Monoclonal Anti-IL-12p40 Antibody Induces Substantial Gut Microbiota Changes in an Experimental Colitis Model. Gastroenterol Res Pract 2016: 4953120

Chassaing B., Koren O., Goodrich J. K., Poole A. C., Srinivasan S., et al. (2015) Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome. Nature 519: 92-96

Chattopadhyay P. K., Gierahn T. M., Roederer M. and Love J. C. (2014) Single-cell technologies for monitoring immune systems. Nat Immunol 15: 128-135

Cirulli E. T. and Goldstein D. B. (2010) Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet 11: 415-425

Civelek M. and Lusis A. J. (2014) Systems genetics approaches to understand complex traits. Nat Rev Genet 15: 34-48

consortium C. (2015) Sparse whole-genome sequencing identifies two loci for major depressive disorder. Nature 523: 588-591

Consortium U. K., Walter K., Min J. L., Huang J., Crooks L., et al. (2015) The UK10K project identifies rare variants in health and disease. Nature 526: 82-90

Cooper D. N., Krawczak M., Polychronakos C., Tyler-Smith C. and Kehrer-Sawatzki H. (2013) Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease. Hum Genet 132: 1077-1130

Coppieters N., Dieriks B. V., Lill C., Faull R. L., Curtis M. A., et al. (2014) Global changes in DNA methylation and hydroxymethylation in Alzheimer's disease human brain. Neurobiol Aging 35: 1334-1344

Crick F. (1970) Central dogma of molecular biology. Nature 227: 561-563

Darmanis S., Sloan S. A., Zhang Y., Enge M., Caneda C., et al. (2015) A survey of human brain transcriptome diversity at the single cell level. Proc Natl Acad Sci U S A 112: 7285-7290

Davies G., Armstrong N., Bis J. C., Bressler J., Chouraki V., et al. (2015) Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). Mol Psychiatry 20: 183-192

Davis H. P. and Squire L. R. (1984) Protein synthesis and memory: a review. Psychol Bull 96: 518-559

De Coppi P., Bartsch G., Jr., Siddiqui M. M., Xu T., Santos C. C., et al. (2007) Isolation of amniotic stem cell lines with potential for therapy. Nat Biotechnol 25: 100-106

de la Calle Mustienes E., Gomez-Skarmeta J. L. and Bogdanovic O. (2015) Genome-wide epigenetic cross-talk between DNA methylation and H3K27me3 in zebrafish embryos. Genom Data 6: 7-9

DeKosky B. J., Kojima T., Rodin A., Charab W., Ippolito G. C., et al. (2015) In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. Nat Med 21: 86-91

Deng W. (2010) Induced pluripotent stem cells: paths to new medicines. A catalyst for disease modelling, drug discovery and regenerative therapy. EMBO Rep 11: 161-165

Dichmann D. S., Walentek P. and Harland R. M. (2015) The alternative splicing regulator Tra2b is required for somitogenesis and regulates splicing of an inhibitory Wnt11b isoform. Cell Rep 10: 527-536

Do C. B., Hinds D. A., Francke U. and Eriksson N. (2012) Comparison of family history and SNPs for predicting risk of complex disease. PLoS Genet 8: e1002973

Dolinoy D. C., Weidman J. R., Waterland R. A. and Jirtle R. L. (2006) Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. Environ Health Perspect 114: 567-572

Dolinoy D. C. (2008) The agouti mouse model: an epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome. Nutr Rev 66 Suppl 1: S7-11

For Research Use Only. Not for use in diagnostic procedures.

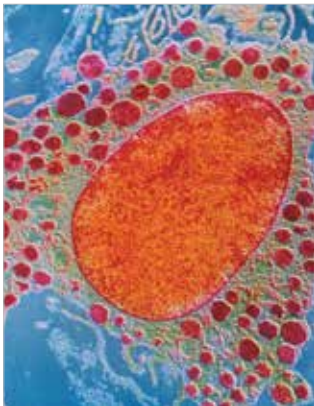**34** Genomic solutions for cell biology and complex disease research

Don E. K., de Jong-Curtain T. A., Doggett K., Hall T. E., Heng B., et al. (2016) Genetic basis of hindlimb loss in a naturally occurring vertebrate model. Biol Open

Dong X., Chen K., Cuevas-Diaz Duran R., You Y., Sloan S. A., et al. (2015) Comprehensive Identification of Long Non-coding RNAs in Purified Cell Types from the Brain Reveals Functional LncRNA in OPC Fate Determination. PLoS Genet 11: e1005669

Dooley J., Tian L., Schonefeldt S., Delghingaro-Augusto V., Garcia-Perez J. E., et al. (2016) Genetic predisposition for beta cell fragility underlies type 1 and type 2 diabetes. Nat Genet

Dowell R. D., Ryan O., Jansen A., Cheung D., Agarwala S., et al. (2010) Genotype to phenotype: a complex problem. Science 328: 469

Dueck H., Eberwine J. and Kim J. (2016) Variation is function: Are single cell differences functionally important?: Testing the hypothesis that single cell variation is required for aggregate function. Bioessays 38: 172-180

Duff M. O., Olson S., Wei X., Garrett S. C., Osman A., et al. (2015) Genome-wide identification of zero nucleotide recursive splicing in Drosophila. Nature 521: 376-379

Dura B., Dougan S. K., Barisa M., Hoehl M. M., Lo C. T., et al. (2015) Profiling lymphocyte interactions at the single-cell level by microfluidic cell pairing. Nat Commun 6: 5940

Ebert A. D., Liang P. and Wu J. C. (2012) Induced pluripotent stem cells as a disease modeling and drug screening platform. J Cardiovasc Pharmacol 60: 408-416

Eberwine J. and Bartfai T. (2011) Single cell transcriptomics of hypothalamic warm sensitive neurons that control core body temperature and fever response Signaling asymmetry and an extension of chemical neuroanatomy. Pharmacol Ther 129: 241-259

Eberwine J., Sul J. Y., Bartfai T. and Kim J. (2014) The promise of single-cell sequencing. Nat Methods 11: 25-27

Egger G., Liang G., Aparicio A. and Jones P. A. (2004) Epigenetics in human disease and prospects for epigenetic therapy. Nature 429: 457-463

Elboudwarej E., Cole M., Briggs F. B., Fouts A., Fain P. R., et al. (2016) Hypomethylation within gene promoter regions and type 1 diabetes in discordant monozygotic twins. J Autoimmun

Eugster A., Lindner A., Catani M., Heninger A. K., Dahl A., et al. (2015) High diversity in the TCR repertoire of GAD65 autoantigen-specific human CD4+ T cells. J Immunol 194: 2531-2538

Evrony G. D., Cai X., Lee E., Hills L. B., Elhosary P. C., et al. (2012) Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. Cell 151: 483-496

Evrony G. D., Lee E., Mehta B. K., Benjamini Y., Johnson R. M., et al. (2015) Cell lineage analysis in human brain using endogenous retroelements. Neuron 85: 49-59

Fan R., Wang Y., Mills J. L., Wilson A. F., Bailey-Wilson J. E., et al. (2013) Functional linear models for association analysis of quantitative traits. Genet Epidemiol 37: 726-742

Fava V. M., Manry J., Cobat A., Orlova M., Van Thuc N., et al. (2016) A Missense LRRK2 Variant Is a Risk Factor for Excessive Inflammatory Responses in Leprosy. PLoS Negl Trop Dis 10: e0004412

Feng H., Qin Z. and Zhang X. (2013) Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. Cancer Lett 340: 179-191

Fields S. and Johnston M. (2005) Cell biology. Whither model organism research? Science 307: 1885-1886

Flajnik M. F. and Kasahara M. (2010) Origin and evolution of the adaptive immune system: genetic events and selective pressures. Nat Rev Genet 11: 47-59

Franke B., Stein J. L., Ripke S., Anttila V., Hibar D. P., et al. (2016) Genetic influences on schizophrenia and subcortical brain volumes: large-scale proof of concept. Nat Neurosci

Fuzik J., Zeisel A., Mate Z., Calvigioni D., Yanagawa Y., et al. (2016) Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. Nat Biotechnol 34: 175-183

Gaublomme J. T., Yosef N., Lee Y., Gertner R. S., Yang L. V., et al. (2015) Single-Cell Genomics Unveils Critical Regulators of Th17 Cell Pathogenicity. Cell 163: 1400-1412

Girirajan S., Campbell C. D. and Eichler E. E. (2011) Human copy number variation and complex genetic disease. Annu Rev Genet 45: 203-226

Gjoneska E., Pfenning A. R., Mathys H., Quon G., Kundaje A., et al. (2015) Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. Nature 518: 365-369

Grskovic M., Javaherian A., Strulovici B. and Daley G. Q. (2011) Induced pluripotent stem cells--opportunities for disease modelling and drug discovery. Nat Rev Drug Discov 10: 915-929

Gudbjartsson D. F., Helgason H., Gudjonsson S. A., Zink F., Oddson A., et al. (2015) Large-scale whole-genome sequencing of the Icelandic population. Nat Genet 47: 435-444

Guo H., Zhu P., Wu X., Li X., Wen L., et al. (2013) Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. Genome Res 23: 2126-2135

Gusella J. F., Wexler N. S., Conneally P. M., Naylor S. L., Anderson M. A., et al. (1983) A polymorphic DNA marker genetically linked to Huntington's disease. Nature 306: 234-238

Hanchate N. K., Kondoh K., Lu Z., Kuang D., Ye X., et al. (2015) Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. Science 350: 1251-1255

Hannon E., Chand A. N., Evans M. D., Wong C. C., Grubb M. S., et al. (2015) A role for Ca1 and calcineurin signaling in depolarization-induced changes in neuronal DNA methylation. Neuroepigenetics 3: 1-6

He B., Nohara K., Ajami N. J., Michalek R. D., Tian X., et al. (2015) Transmissible microbial and metabolomic remodeling by soluble dietary fiber improves metabolic homeostasis. Sci Rep 5: 10604

Hibar D. P., Stein J. L., Renteria M. E., Arias-Vasquez A., Desrivieres S., et al. (2015) Common genetic variants influence human subcortical brain structures. Nature 520: 224-229

Hoh R. A., Joshi S. A., Liu Y., Wang C., Roskin K. M., et al. (2016) Single B-cell deconvolution of peanut-specific antibody responses in allergic patients. J Allergy Clin Immunol 137: 157-167

Holliday R. (2006) Epigenetics: a historical overview. Epigenetics 1: 76-80

Holoch D. and Moazed D. (2015) RNA-mediated epigenetic regulation of gene expression. Nat Rev Genet 16: 71-84

Holt C. E. and Schuman E. M. (2013) The central dogma decentralized: new perspectives on RNA function and local translation in neurons. Neuron 80: 648-657

Hrdlickova B., Westra H. J., Franke L. and Wijmenga C. (2011) Celiac disease: moving from genetic associations to causal variants. Clin Genet 80: 203-313

Hu Y., Shmygelska A., Tran D., Eriksson N., Tung J. Y., et al. (2016) GWAS of 89,283 individuals identifies genetic variants associated with self-reporting of being a morning person. Nat Commun 7: 10448

Huang K., Shen Y., Xue Z., Bibikova M., April C., et al. (2014) A panel of CpG methylation sites distinguishes human embryonic stem cells and induced pluripotent stem cells. Stem Cell Reports 2: 36-43

Imperio C. G., McFalls A. J., Colechio E. M., Masser D. R., Vrana K. E., et al. (2015) Assessment of individual differences in the rat nucleus accumbens transcriptome following taste-heroin extended access. Brain Res Bull

Ingolia N. T., Ghaemmaghami S., Newman J. R. and Weissman J. S. (2009) Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. Science 324: 218-223

Inoue H. and Yamanaka S. (2011) The use of induced pluripotent stem cells in drug development. Clin Pharmacol Ther 89: 655-661

International HapMap C. (2003) The International HapMap Project. Nature 426: 789-796

International Stem Cell I., Adewumi O., Aflatoonian B., Ahrlund-Richter L., Amit M., et al. (2007) Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. Nat Biotechnol 25: 803-816

Iourov I. Y., Vorsanova S. G. and Yurov Y. B. (2012) Single cell genomics of the brain: focus on neuronal diversity and neuropsychiatric diseases. Curr Genomics 13: 477-488

Jiang Y., Jahagirdar B. N., Reinhardt R. L., Schwartz R. E., Keene C. D., et al. (2002) Pluripotency of mesenchymal stem cells derived from adult marrow. Nature 418: 41-49

Johnson A. D. and O'Donnell C. J. (2009) An open access database of genome-wide association results. BMC Med Genet 10: 6

Johnson M. R., Shkura K., Langley S. R., Delahaye-Duriez A., Srivastava P., et al. (2016) Systems genetics identifies a convergent gene network for cognition and neurodevelopmental disease. Nat Neurosci 19: 223-232

Jones-Hall Y. L., Kozik A. and Nakatsu C. (2015) Ablation of tumor necrosis factor is associated with decreased inflammation and alterations of the microbiota in a mouse model of inflammatory bowel disease. PLoS One 10: e0119441

Kahrstrom C. T. (2013) Epigenetics: Legionella makes its mark on histones. Nat Rev Genet 14: 370

Kao C. Y., He Z., Zannas A. S., Hahn O., Kuhne C., et al. (2016) Fluoxetine treatment prevents the inflammatory response in a mouse model of posttraumatic stress disorder. J Psychiatr Res 76: 74-83

Karve T. M. and Cheema A. K. (2011) Small changes huge impact: the role of protein posttranslational modifications in cellular homeostasis and disease. J Amino Acids 2011: 207691

Kashi K., Henderson L., Bonetti A. and Carninci P. (2016) Discovery and functional analysis of lncRNAs: Methodologies to investigate an uncharacterized transcriptome. Biochim Biophys Acta 1859: 3-15

Kaut O., Ramirez A., Pieper H., Schmitt I., Jessen F., et al. (2014) DNA methylation of the TNF-alpha promoter region in peripheral blood monocytes and the cortex of human Alzheimer's disease patients. Dement Geriatr Cogn Disord 38: 10-15

Kendler K. S. (2015) A joint history of the nature of genetic variation and the nature of schizophrenia. Mol Psychiatry 20: 77-83

Kim C. (2014) Disease modeling and cell based therapy with iPSC: future therapeutic option with fast and safe application. Blood Res 49: 7-14

Kim S., Becker J., Bechheim M., Kaiser V., Noursadeghi M., et al. (2014) Characterizing the genetic basis of innate immune response in TLR4-activated human monocytes. Nat Commun 5: 5236

Kimura M. (1991) The neutral theory of molecular evolution: a review of recent evidence. Jpn J Genet 66: 367-386

Klein A. M., Mazutis L., Akartuna I., Tallapragada N., Veres A., et al. (2015) Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161: 1187-1201

Kloosterman W. P. and Plasterk R. H. (2006) The diverse functions of microRNAs in animal development and disease. Dev Cell 11: 441-450

Kraiczy J., Nayak K., Ross A., Raine T., Mak T. N., et al. (2015) Assessing DNA methylation in the developing human intestinal epithelium: potential link to inflammatory bowel disease. Mucosal Immunol

Lander E. S. and Schork N. J. (1994) Genetic dissection of complex traits. Science 265: 2037-2048

Lee Y. S., Lewis J. A., Ippolito D. L., Hussainzada N., Lein P. J., et al. (2016) Repeated exposure to neurotoxic levels of chlorpyrifos alters hippocampal expression of neurotrophins and neuropeptides. Toxicology 340: 53-62

Leonard W. R., Snodgrass J. J. and Robertson M. L. (2010) Fat Detection: Taste, Texture, and Post Ingestive Effects. Fat Detection: Taste, Texture, and Post Ingestive Effects.

Linnemann C., Heemskerk B., Kvistborg P., Kluin R. J., Bolotin D. A., et al. (2013) High-throughput identification of antigen-specific TCRs by TCR gene capture. Nat Med 19: 1534-1541

Liu C., Lou C. H., Shah V., Ritter R., Talley J., et al. (2016) Identification of microRNAs and microRNA targets in Xenopus gastrulae: The role of miR-26 in the regulation of Smad1. Dev Biol 409: 26-38

Liu Z., Li X., Zhang J. T., Cai Y. J., Cheng T. L., et al. (2016) Autism-like behaviours and germline transmission in transgenic monkeys overexpressing MeCP2. Nature 530: 98-102

Lockett G. A., Patil V. K., Soto-Ramirez N., Ziyab A. H., Holloway J. W., et al. (2013) Epigenomics and allergic disease. Epigenomics 5: 685-699

Lossius A., Johansen J. N., Vartdal F. and Holmoy T. (2016) High-throughput sequencing of immune repertoires in multiple sclerosis. Annals of CLinical and Translational Neurology

Lovatt D., Ruble B. K., Lee J., Dueck H., Kim T. K., et al. (2014) Transcriptome *in vivo* analysis (TIVA) of spatially defined single cells in live tissue. Nat Methods 11: 190-196

Lun M. P., Johnson M. B., Broadbelt K. G., Watanabe M., Kang Y. J., et al. (2015) Spatially heterogeneous choroid plexus transcriptomes encode positional identity and contribute to regional CSF production. J Neurosci 35: 4903-4916

Macaulay I. C. and Voet T. (2014) Single cell genomics: advances and future perspectives. PLoS Genet 10: e1004126

Macaulay I. C., Svensson V., Labalette C., Ferreira L., Hamey F., et al. (2016) Single-Cell RNA-Sequencing Reveals a Continuous Spectrum of Differentiation in Hematopoietic Cells. Cell Rep 14: 966-977

Macosko E. Z., Basu A., Satija R., Nemesh J., Shekhar K., et al. (2015) Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 161: 1202-1214

Malecova B. and Morris K. V. (2010) Transcriptional gene silencing through epigenetic changes mediated by non-coding RNAs. Curr Opin Mol Ther 12: 214-222

Manolio T. A., Collins F. S., Cox N. J., Goldstein D. B., Hindorff L. A., et al. (2009) Finding the missing heritability of complex diseases. Nature 461: 747-753

Marre M. L., James E. A. and Piganelli J. D. (2015) beta cell ER stress and the implications for immunogenicity in type 1 diabetes. Front Cell Dev Biol 3: 67

Marwick J. A., Kirkham P. A., Stevenson C. S., Danahay H., Giddings J., et al. (2004) Cigarette smoke alters chromatin remodeling and induces proinflammatory genes in rat lungs. Am J Respir Cell Mol Biol 31: 633-642

Mattson M. P. (2010) Acetylation unleashes protein demons of dementia. Neuron 67: 900-902

Maza I., Caspi I., Zviran A., Chomsky E., Rais Y., et al. (2015) Transient acquisition of pluripotency during somatic cell transdifferentiation with iPSC reprogramming factors. Nat Biotechnol 33: 769-774

McCarroll S. A. and Altshuler D. M. (2007) Copy-number variation and association studies of human disease. Nat Genet 39: S37-42

McCarthy M. I., Abecasis G. R., Cardon L. R., Goldstein D. B., Little J., et al. (2008) Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 9: 356-369

McConnell M. J., Lindberg M. R., Brennand K. J., Piper J. C., Voet T., et al. (2013) Mosaic copy number variation in human neurons. Science 342: 632-637

McGettigan P. A. (2013) Transcriptomics in the RNA-seq era. Curr Opin Chem Biol 17: 4-11

Mistry V., Bockett N. A., Levine A. P., Mirza M. M., Hunt K. A., et al. (2015) Exome sequencing of 75 individuals from multiply affected coeliac families and large scale resequencing follow up. PLoS One 10: e0116845

Mittal A., Pachter L., Nelson J. L., Kjaergaard H., Smed M. K., et al. (2015) Pregnancy-Induced Changes in Systemic Gene Expression among Healthy Women and Women with Rheumatoid Arthritis. PLoS One 10: e0145204

Moore-Connors J. M., Dunn K. A., Bielawski J. P. and Van Limbergen J. (2016) Novel Strategies for Applied Metagenomics. Inflamm Bowel Dis 22: 709-718

Morgan X. C., Kabakchiev B., Waldron L., Tyler A. D., Tickle T. L., et al. (2015) Associations between host gene expression, the mucosal microbiome, and clinical outcome in the pelvic pouch of patients with inflammatory bowel disease. Genome Biol 16: 67

Muraro P. A., Robins H., Malhotra S., Howell M., Phippard D., et al. (2014) T cell repertoire following autologous stem cell transplantation for multiple sclerosis. J Clin Invest 124: 1168-1172

Neale B. M. (2010) Introduction to linkage disequilibrium, the HapMap, and imputation. Cold Spring Harb Protoc 2010: pdb top74

Need A. C., Ge D., Weale M. E., Maia J., Feng S., et al. (2009) A genome-wide investigation of SNPs and CNVs in schizophrenia. PLoS Genet 5: e1000373

Ooi L., Sidhu K., Poljak A., Sutherland G., O'Connor M. D., et al. (2013) Induced pluripotent stem cells as tools for disease modelling and drug discovery in Alzheimer's disease. J Neural Transm (Vienna) 120: 103-111

Paul F., Arkin Y., Giladi A., Jaitin D. A., Kenigsberg E., et al. (2015) Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. Cell 163: 1663-1677

Peleg S., Sananbenesi F., Zovoilis A., Burkhardt S., Bahari-Javan S., et al. (2010) Altered histone acetylation is associated with age-dependent memory impairment in mice. Science 328: 753-756

Pfeifer G. P., Kadam S. and Jin S. G. (2013) 5-hydroxymethylcytosine and its potential roles in development and cancer. Epigenetics Chromatin 6: 10

Pinello L., Xu J., Orkin S. H. and Yuan G. C. (2014) Analysis of chromatin-state plasticity identifies cell-type-specific regulators of H3K27me3 patterns. Proc Natl Acad Sci U S A 111: E344-353

Plomin R., Haworth C. M. and Davis O. S. (2009) Common disorders are quantitative traits. Nat Rev Genet 10: 872-878

Poduri A., Evrony G. D., Cai X. and Walsh C. A. (2013) Somatic mutation, genomic variation, and neurological disease. Science 341: 1237758

Polderman T. J., Benyamin B., de Leeuw C. A., Sullivan P. F., van Bochoven A., et al. (2015) Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nat Genet 47: 702-709

Qureshi I. A. and Mehler M. F. (2014) Epigenetic mechanisms underlying the pathogenesis of neurogenetic diseases. Neurotherapeutics 11: 708-720

Revill P. (2007) Genomics of common diseases. Drug News Perspect 20: 475-479

Rhrissorrakrai K., Belcastro V., Bilal E., Norel R., Poussin C., et al. (2015) Understanding the limits of animal models as predictors of human biology: lessons learned from the sbv IMPROVER Species Translation Challenge. Bioinformatics 31: 471-483

Rivera C. M. and Ren B. (2013) Mapping human epigenomes. Cell 155: 39-55

Roelfsema P. R. and Treue S. (2014) Basic neuroscience research with nonhuman primates: a small but indispensable component of biomedical research. Neuron 82: 1200-1204

Ryu S., Atzmon G., Barzilai N., Raghavachari N. and Suh Y. (2016) Genetic landscape of APOE in human longevity revealed by high-throughput sequencing. Mech Ageing Dev

Safka Brozkova D., Deconinck T., Griffin L. B., Ferbert A., Haberlova J., et al. (2015) Loss of function mutations in HARS cause a spectrum of inherited peripheral neuropathies. Brain 138: 2161-2172

Sanchez-Mut J. V., Aso E., Panayotis N., Lott I., Dierssen M., et al. (2013) DNA methylation map of mouse and human brain identifies target genes in Alzheimer's disease. Brain 136: 3018-3027

Sareddy G. R., Zhang Q., Wang R., Scott E., Zou Y., et al. (2015) Proline-, glutamic acid-, and leucine-rich protein 1 mediates estrogen rapid signaling and neuroprotection in the brain. Proc Natl Acad Sci U S A 112: E6673-6682

Schafer S., Adami E., Heinig M., Rodrigues K. E., Kreuchwig F., et al. (2015) Translational regulation shapes the molecular landscape of complex disease phenotypes. Nat Commun 6: 7200

Schaubeck M., Clavel T., Calasan J., Lagkouvardos I., Haange S. B., et al. (2015) Dysbiotic gut microbiota causes transmissible Crohn's disease-like ileitis independent of failure in antimicrobial defence. Gut

Schizophrenia Psychiatric Genome-Wide Association Study C. (2011) Genome-wide association study identifies five new schizophrenia loci. Nat Genet 43: 969-976

Schizophrenia Working Group of the Psychiatric Genomics C. (2014) Biological insights from 108 schizophrenia-associated genetic loci. Nature 511: 421-427

Scobie K. N., Damez-Werno D., Sun H., Shao N., Gancarz A., et al. (2014) Essential role of poly(ADP-ribosyl)ation in cocaine action. Proc Natl Acad Sci U S A 111: 2005-2010

Sekar A., Bialas A. R., de Rivera H., Davis A., Hammond T. R., et al. (2016) Schizophrenia risk from complex variation of complement component 4. Nature

Shi L., Zhang X., Golhar R., Otieno F. G., He M., et al. (2013) Whole-genome sequencing in an autism multiplex family. Mol Autism 4: 8

Silva P. E., Costa P. S., Avila M. P., Suhadolnik M. L., Reis M. P., et al. (2015) Leprous lesion presents enrichment of opportunistic pathogenic bacteria. Springerplus 4: 187

Simchovitz A., Soreq L. and Soreq H. (2016) Transcriptome profiling in Parkinson's leukocytes: from early diagnostics to neuroimmune therapeutic prospects. Curr Opin Pharmacol 26: 102-109

Soldner F. and Jaenisch R. (2012) Medicine. iPSC disease modeling. Science 338: 1155-1156

Stamnaes J. and Sollid L. M. (2015) Celiac disease: Autoimmunity in response to food antigen. Semin Immunol 27: 343-352

Stankiewicz A. M., Swiergiel A. H. and Lisowski P. (2013) Epigenetics of stress adaptations in the brain. Brain Res Bull 98: 76-92

Steinberg S., Stefansson H., Jonsson T., Johannsdottir H., Ingason A., et al. (2015) Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease. Nat Genet 47: 445-447

Sterneckert J. L., Reinhardt P. and Scholer H. R. (2014) Investigating human disease using stem cell models. Nat Rev Genet 15: 625-639

Streets A. M., Zhang X., Cao C., Pang Y., Wu X., et al. (2014) Microfluidic single-cell whole-transcriptome sequencing. Proc Natl Acad Sci U S A 111: 7048-7053

Stubbington M. J., Lonnberg T., Proserpio V., Clare S., Speak A. O., et al. (2016) T cell fate and clonality inference from single-cell transcriptomes. Nat Methods

Takahashi K. and Yamanaka S. (2016) A decade of transcription factor-mediated reprogramming to pluripotency. Nat Rev Mol Cell Biol 17: 183-193

Tanaka T. (2005) [International HapMap project]. Nihon Rinsho 63 Suppl 12: 29-34

For Research Use Only. Not for use in diagnostic procedures.

Tasan M., Musso G., Hao T., Vidal M., MacRae C. A., et al. (2015) Selecting causal genes from genome-wide association studies via functionally coherent subnetworks. Nat Methods 12: 154-159

Tasic B., Menon V., Nguyen T. N., Kim T. K., Jarsky T., et al. (2016) Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat Neurosci 19: 335-346

Thomas D. C. (2009) Genome-wide association studies for discrete traits. Genet Epidemiol 33 Suppl 1: S8-12

Thomson J. A., Itskovitz-Eldor J., Shapiro S. S., Waknitz M. A., Swiergiel J. J., et al. (1998) Embryonic stem cell lines derived from human blastocysts. Science 282: 1145-1147

Thomson J. P., Lempiainen H., Hackett J. A., Nestor C. E., Muller A., et al. (2012) Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. Genome Biol 13: R93

Tipton C. M., Fucile C. F., Darce J., Chida A., Ichikawa T., et al. (2015) Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. Nat Immunol 16: 755-765

Tran P. V., Kennedy B. C., Pisansky M. T., Won K. J., Gewirtz J. C., et al. (2016) Prenatal Choline Supplementation Diminishes Early-Life Iron Deficiency-Induced Reprogramming of Molecular Networks Associated with Behavioral Abnormalities in the Adult Rat Hippocampus. J Nutr 146: 484-493

Trapnell C. (2015) Defining cell types and states with single-cell genomics. Genome Res 25: 1491-1498

Treutlein B., Brownfield D. G., Wu A. R., Neff N. F., Mantalas G. L., et al. (2014) Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature 509: 371-375

Tsujioka H., Kunieda T., Katou Y., Shirahige K. and Kubo T. (2015) Unique gene expression profile of the proliferating Xenopus tadpole tail blastema cells deciphered by RNA-sequencing analysis. PLoS One 10: e0111655

Urak K. T., Shore S., Rockey W. M., Chen S. J., McCaffrey A. P., et al. (2016) In vitro RNA SELEX for the generation of chemically-optimized therapeutic RNA drugs. Methods

Usoskin D., Furlan A., Islam S., Abdo H., Lonnerberg P., et al. (2015) Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. Nat Neurosci 18: 145-153

Valenci I., Yonai L., Bar-Yaacov D., Mishmar D. and Ben-Zvi A. (2015) Parkin modulates heteroplasmy of truncated mtDNA in Caenorhabditis elegans. Mitochondrion 20: 64-70

van der Sijde M. R., Ng A. and Fu J. (2014) Systems genetics: From GWAS to disease pathways. Biochim Biophys Acta 1842: 1903-1909

Vermunt M. W., Tan S. C., Castelijns B., Geeven G., Reinink P., et al. (2016) Epigenomic annotation of gene regulatory alterations during evolution of the primate brain. Nat Neurosci 19: 494-503

Voet T., Kumar P., Van Loo P., Cooke S. L., Marshall J., et al. (2013) Single-cell paired-end genome sequencing reveals structural variation per cell cycle. Nucleic Acids Res 41: 6119-6138

Waddington C. H. (2012) The epigenotype. 1942. Int J Epidemiol 41: 10-13

Wang W., Jovel J., Halloran B., Wine E., Patterson J., et al. (2015) Metagenomic analysis of microbiome in colon tissue from subjects with inflammatory bowel diseases reveals interplay of viruses and bacteria. Inflamm Bowel Dis 21: 1419-1427

Wang Z., Gerstein M. and Snyder M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57-63

Warnault V. and Ron D. (2013) Chromatin remodeling: a new landscape to treat harmful alcohol-use disorders. Future Med Chem 5: 2011-2013

Watanabe Y., Arase S., Nagaoka N., Kawai M. and Matsumoto S. (2016) Chronic Psychological Stress Disrupted the Composition of the Murine Colonic Microbiota and Accelerated a Murine Model of Inflammatory Bowel Disease. PLoS One 11: e0150559

Wellcome Trust Case Control C. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661-678

Wu C. C., Kruse F., Vasudevarao M. D., Junker J. P., Zebrowski D. C., et al. (2016) Spatially Resolved Genome-wide Transcriptional Profiling Identifies BMP Signaling as Essential Regulator of Zebrafish Cardiomyocyte Regeneration. Dev Cell 36: 36-49

Wu S. M. and Hochedlinger K. (2011) Harnessing the potential of induced pluripotent stem cells for regenerative medicine. Nat Cell Biol 13: 497-505

Xiao Y., Liu F., Zhang Z., Tang J., Zou C. G., et al. (2016) Gut-Colonizing Bacteria Promote *C. elegans* Innate Immunity by Producing Nitric Oxide. Cell Rep 14: 1301-1307

Xue Z., Huang K., Cai C., Cai L., Jiang C. Y., et al. (2013) Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature 500: 593-597

Yang C. P., Fu C. C., Sugino K., Liu Z., Ren Q., et al. (2016) Transcriptomes of lineage-specific Drosophila neuroblasts profiled by genetic targeting and robotic sorting. Development 143: 411-421

Yuen R. K., Thiruvahindrapuram B., Merico D., Walker S., Tammimies K., et al. (2015) Whole-genome sequencing of quartet families with autism spectrum disorder. Nat Med 21: 185-191

Zaratiegui M., Irvine D. V. and Martienssen R. A. (2007) Noncoding RNAs and gene silencing. Cell 128: 763-776

Zhang J., Yuan L., Zhang X., Hamblin M. H., Zhu T., et al. (2016) Altered long non-coding RNA transcriptomic profiles in brain microvascular endothelium after cerebral ischemia. Exp Neurol 277: 162-170

Zhang K. X., Tan L., Pellegrini M., Zipursky S. L. and McEwen J. M. (2016) Rapid Changes in the Translatome during the Conversion of Growth Cones to Synaptic Terminals. Cell Rep 14: 1258-1271

Scientific Publication Reviews can be accessed at
www.illumina.com/pubreviews

Immunology Research Review

Repertoire sequencing has enabled researchers to identify unique receptor variants found in individuals with susceptibility to hematological malignancies, autoimmune diseases, and allergen response.

Illumina next-generation sequencing provides the quality, throughput, and read lengths required by the research community to map the human immune response at high resolution. The emergence of new approaches such as phase-defined sequencing and single-cell sequencing can be expected to accelerate this knowledge base.
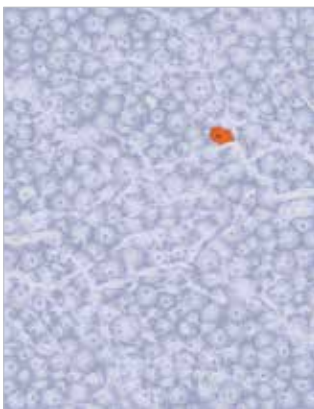
Single Cell Research Review

Most of the impetus for single-cell tissue sequencing has come from cancer research, where cell lineage and the detection of residual disease are of paramount concern. The same approaches are being used to improve our understanding of massively complex biological systems, such as neural development and immunology.

This document highlights recent publications that demonstrate the use of Illumina technology for single-cell sequencing and very low input applications and techniques.

This Scientific Publication Review is brought to you by Illumina, Inc.

illumına®