Libertas Academica
FREEDOM TO RESEARCH

Open Access: Full open access to this and thousands of other papers at http://www.la-press.com.

Gene Expression to Genetical Genomics

Recent Advances in Studying of Copy Number Variation and Gene Expression

George E. Liu[1], Lingyang Xu[1,2] and Kevin S. Huang[1,3]

[1]Bovine Functional Genomics Laboratory, USDA, ARS, Beltsville, MD, USA. [2]Department of Animal and Avian Sciences, University of Maryland, College Park, MD, USA. [3]Centennial High School, 4300 Centennial Lane, Ellicott City, MD, USA.

ABSTRACT: Copy number variations (CNVs) are gains and losses of genomic sequence between two individuals of a species. CNV can have a significant impact on the expression of genes mapped within, nearby and far away from the CNV event itself. Earlier studies have revealed that CNV effects can be caused by affecting gene dosage, through position effect, altering downstream pathways and regulatory networks, or modifying the chromosome structure and position within the nucleus. We will first review recently published results based on high-resolution CNV and transcriptome data. We will then discuss new insights brought by emerging new technologies of next generation sequencing and single-cell sequencing.

KEYWORDS: copy number variation, gene expression, mechanism, next generation sequencing, single-cell sequencing

CITATION: Liu et al. Recent Advances in Studying of Copy Number Variation and Gene Expression. *Gene Expression to Genetical Genomics* 2014:7 1–5 doi:10.4137/GGG.S14286.

RECEIVED: March 17, 2014. **RESUBMITTED:** April 22, 2014. **ACCEPTED FOR PUBLICATION:** April 23, 2014.

ACADEMIC EDITOR: Shiva Singh, Editor in Chief

TYPE: Review

FUNDING: This work was supported in part by USDA Project 1245-31000-101-00 from the USDA's ARS and the AFRI grant No. 2011-67015-30183 from USDA NIFA.

DISCLAIMER: Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. The USDA is an equal opportunity provider and employer.

COMPETING INTERESTS: Authors disclose no potential conflicts of interest.

CORRESPONDENCE: george.liu@ars.usda.gov

This paper was subject to independent, expert peer review by a minimum of two blind peer reviewers. All editorial decisions were made by the independent academic editor. All authors have provided signed confirmation of their compliance with ethical and legal obligations including (but not limited to) use of any copyrighted material, compliance with ICMJE authorship and competing interests disclosure guidelines and, where applicable, compliance with legal and ethical guidelines on human and animal research participants. Provenance: the authors were invited to submit this paper.

Introduction

One major theme of modern genetics is to understand the relationship between genotype and phenotype. Variation of gene expression (narrowly defined here as the transcription of DNA into mRNA) may serve as a proxy for phenotypic variation.[1–3] Gene expression has been studied in many species in order to understand its genetic basis of variation and its regulatory networks. These studies provide multiple models to explain the causes of phenotypic variation.

Along with single nucleotide polymorphisms (SNPs), recent studies dealing with larger, more complex forms of genetic variation, structural variations, have attracted much attention.[4] Among the largest of the genetic variant types, copy number variations (CNVs or unbalanced structural variations) are deletions and insertions of genomic sequence between two individuals of a species.[5] Additionally, balanced structural variants include chromosomal fusions, translocations, and inversions. In this review, we generally refer all of them as CNV unless we state otherwise. Phenotypic effects of CNVs are caused by changes in gene expression, directly affecting gene dosage, indirectly through position effect, or downstream pathways and regulatory networks. One extreme example of CNV is chromosome aneuploidy (ie imbalances in chromosome number), which has been shown to globally influence gene expression on the variant chromosome and is usually lethal to the host.[6,7] Early studies of CNV and gene expression and their potential mechanisms have been reviewed before.[8,9] In this review, we will focus on newly published CNV studies based on high-resolution CNV and transcriptome data. We will then discuss influences of emerging new technologies of next generation sequencing (NGS) and single-cell sequencing on the future studies.

CNVs in Human and Model Organisms

CNVs have been intensively studied in humans[10–14] and several model organisms, including mouse, rat, and fruit fly.[15–20] Besides a number of plants, CNV reports also appeared in other animals, such as chimpanzee, macaque, dog, and cattle.[21–25] The newest version (July 2013) of Database of Genomic Variants (DGV) collected 55 human CNV studies (http://dgv.tcag.ca/dgv/app/home), containing 230,4349 human CNVs and 109,863 CNV regions.[26] The Database of Genomic Structural Variation (dbVar) hosts 106 published studies, including 85 human, 12 mouse, and several other animal datasets (http://www.ncbi.nlm.nih.gov/dbvar/).[27,28]

Different CNV detection platforms (hybridization or sequence based) and calling algorithms resulted in different CNV datasets in varying resolution.[29] Several major CNV formation mechanisms have been proposed, including NAHR (Non-Allelic Homologous Recombination), NHEJ (Non-Homologous End Joining), MEI (Mobile Element Insertion), and FoSTeS/MMBIR (Fork Stalling and Template Switching and Microhomology-Mediated Break-Induced Replication).[30] Specifically, large CNVs are mainly formed by NAHR mediated by repetitive sequences, including segmental duplications (SDs).[31,32] Although tandem duplications seem the most likely archetypical organization as shown in mouse, rat, dog, cattle, and fruit fly,[15,17,23,33] human and great ape species show a preponderance of interspersed duplications. As a result, large CNVs are associated with SDs in mammals and enriched in centromeres and telomeres.

Dozens of CNVs have been shown to be important in both normal phenotypic variability and disease susceptibility in human.[34–40] Although analyses of a subset of CNVs provided evidence of linkage disequilibrium with flanking SNPs,[41] a significant portion of CNVs fell in genomic regions not well covered by SNP arrays, such as SD regions, and thus were not genotyped.[42–44] Combining CNV and SNP data in human genome-wide association studies has associated CNVs with diseases such as intellectual disability, autism, schizophrenia, neuroblastoma, Crohn's disease, and severe early-onset obesity.[45–54]

Early Studies of CNVs and Gene Expression

By year 2011, several reports have reported associations of CNVs with gene expression variation in human, rodents, and fruit fly based on early low-resolution CNV and gene expression data.[15,18,20,55,56] For example, a comprehensive survey by Stranger et al reported that SNPs and CNVs captured about 82% and 18% of the total detected genetic variation in gene expression, respectively.[55] Chaignat et al studied brain and liver tissues of mouse embryos at different stages and observed the spatial and temporal influence of CNVs on tissue transcriptomes throughout development.[57] Notably, they detected that CNVs are significantly enriched within differentially expressed genes in both adulthood and at embryonic stage among mouse strains. Additionally, they found that only in brain, some expressed genes within CNVs appear to be under compensatory loops only at specific time points, suggesting that the effect of CNVs on these genes in brain is modulated during development. These studies and potential mechanisms have been reviewed before.[8,9] We will briefly summarize molecular mechanisms below. All of them may play a role, either individually or in combination, highlighting complex relationships between CNV and gene expression.

Dosage sensitive. Only a weak positive correlation was found between CNV and gene expression, dependent on tissue type. This correlation is driven by a small fraction of genes (5–18% in rodents).[18,56,58]

Dosage insensitive and dosage compensation. In over 70% of the fruit fly genes, CNV had no effect on gene expression, suggesting either dosage compensation mechanisms or the incomplete inclusion of regulatory elements in the CNV event.[20] In addition, mechanisms similar to genetic imprinting may also inhibit the expression at a CNV locus.

Dosage reversed. For 2–15% of the genes, gene expression levels were significantly inversely correlated with CNV.[18,55,56,59,60] This observation can be explained by either a negative feedback loop, which reduces the expression of the CNV gene or a steric hindrance of the extra copies of a gene, which impairs their access to a specific transcription factory.[61] For example, deletion of a transcriptional repressor could serve to elevate gene expression.[55]

Position effect. Position effect or *cis*-ruption is used to describe the physical dissociation of the gene and its regulatory *cis* elements.[62,63] Position effect has been reported for the expression of genes of up to 1.5 Mb from the CNV event itself.[64]

Chromosome structure and positioning. There are several hypotheses explaining disruption of CNV of long-range *trans* regulations: CNVs can result in modification of transcriptional control through alteration of chromosome structure[65–67] and modification of the positioning of chromosome within the nucleus and/or within a chromosome territory of a genomic region.[68,69]

Other effects. Deletion of CNVs can alter a phenotype by unmasking recessive mutations. CNVs may also influence gene expression through perturbation of transcript structure.[70,71]

Altogether, this first batch of studies only evaluated the effects of large CNVs at a low resolution, ascertained with the early versions of microarray platforms with low probe densities. Indeed, array comparative genomic hybridization (aCGH) can only assay the sequence present in the reference assembly. It cannot distinguish between tandem and non-tandem duplications. Another caveat is the reliability of CNV detection, particularly in defining CNV breakpoints with confidence when aCGH data are used.[12,14] Unreliable boundary calling can lead to genes being erroneously included or excluded from CNVs, which can artificially skew gene expression profiles and affect correlation results.

Recent Studies of CNVs and Gene Expression

Technological advances in microarray and NGS have recently led to "second-generation" CNV maps with increased resolution and broadened CNV size range.[5,14] The 1000 Genomes Project reported CNVs of 50 bp to a 1 Mb in size (median size 730 bp) in more than 150 individuals, with over 50% of CNV were mapped to nucleotide resolution.[5] The count of detected CNVs has increased from below a hundred to several thousand per individual.[5,14] Improved CNV genotyping technology further facilitated associating CNVs with phenotypic data.[5,14,41,72,73]

Recent human studies. By relating second-generation CNV genotypes to transcriptome sequencing (RNA-Seq) data, Schlattl et al evaluated the impact of CNV on gene expression focusing on proximal effects less than 200 kb.[74] Totally, they identified CNVs associated with the expression of 110 genes. They observed an enrichment of large CNVs, including large intergenic CNVs, relative to the entire set of expression-associated CNVs. While a minor part of gene expression may be affected by CNV, their results suggest that dosage compensation in human cells may be less pronounced than in *Drosophila melanogaster*.[20] Furthermore, 48% of the genes they identified in the context of CNV-associated expression QTLs were also associated with SNPs before.[75,76] This portion is higher than the fraction of genes (18%) associated with CNVs by Stranger et al, which also displayed a significant SNP association.[55] They attribute this higher overlap to the increased resolution and accuracy of the second-generation CNV data. Altogether, their results suggest that association studies can gain in resolution and power by including fine-scale CNV information. It is noted that this study was bias against duplications/insertions as most of associated CNVs were deletions. Also, a low count of associated genes (110) may be an underestimation of the impact of CNVs on gene expression, because CNVs involving long-range *cis*- and *trans*-regulatory elements were not investigated in this study.

Hulse and Cai recently reported a genome-wide screening for expression variability quantitative trait loci (evQTL) in humans and found that most *cis*-acting evQTL are located in CNV regions.[77] On the one hand, their results indicate that evQTL account for the mixture of positive and negative correlations between CNV and gene expression. On the other hand, their results could suggest that control of gene copy number is associated with reducing the variability in gene expression.[78] In human tumor cells, copy number changes of oncogenes and tumor suppressor genes may disrupt gene expression and stoichiometric relationships in cell metabolism and physiology, leading to cancer development and progression. It is estimated that ~60% of the genes show differential expression concordant to their copy number status, suggesting that the global correlation between CNV and gene expression is consistent across cancer studies.[79]

Recent studies in nonhuman primates. CNV synteny sequences have been found in primates with 22% and 25% of chimpanzee and macaque CNVs overlapping human CNVs, suggesting the existence of CNV hotspots in primates.[21,22] Iskow et al reported that regulatory element copy number differences can shape gene expression profiles in primates.[80] They discovered 964 copy number differences of conserved sequences across human, chimpanzee, and macaque and detected their impacts on species-specific gene expression profiles. Individual samples with copy number different genes had significantly different expression than samples with neutral copy number. Transcription factor genes and other regulatory elements (ultraconserved elements and long intergenic noncoding RNAs) differed in copy number and were associated with significant expression differences. Using NGS-based CNV maps, Gokcumen et al showed that the lineage-specific formation rates of NAHR and Alu repeats resulted in markedly different CNV landscapes in chimpanzee, orangutan, and macaque.[81] They further described several gene duplications, which led to evolutionary innovation through the gain of gene expression in new tissues.

Future Directions

Single-cell sequencing is generating new biological insights, due to the improved protocols for isolation of single cells and amplification of their genomes (DNA) or transcriptomes (RNA).[82–84] Traditionally, pooling large quantity of cells from a tissue obscures the heterogeneity of complex systems. Single-cell sequencing helped to discover clonal mutations, cryptic cell types, or transcriptional features that would be diluted or averaged out in bulk tissue.[85–87]

To provide the proof of concept, genome-wide SNPs and CNVs were successfully detected by using a single human cell.[88] Using single-cell sequencing, McConnell et al further showed that mosaic CNV is abundant in human neurons.[89] Neurons taken from postmortem human frontal cortex tissue and neurons derived from induced pluripotent stem cell differentiation *in vitro* showed surprising diversity in individual cell genomes: 13–41% of neurons had at least one megabase-scale *de novo* CNV with twice more deletions than duplications. On the other hand, single neuron sequencing performed by Evrony et al suggested that LINE-1 retrotransposition is not a major generator of neuronal diversity in brain tissue.[90] It will be interesting to see how these CNV results are correlated with the strictly regulated expression of CNV genes in the brain as shown by the mouse embryo study.[57] Therefore, more detailed RNA-Seq studies at single neuron level are warranted. Actually, single-cell RNA-Seq already revealed dynamic, random monoallelic gene expression in mouse cells, which increases the heterogeneity among cells and likely contributes to the phenotypic variance among individuals of identical genotype.[91] Single-cell DNA methylome landscapes of mouse embryonic stem cells and early embryos were also analyzed using reduced representation bisulfite sequencing.[92]

As single-cell amplification matures and sequencing cost drops, this approach will become a standard tool for understanding CNV and gene expression relationship at high resolution.

Many of the mechanisms underlying the CNV-associated gene expression changes remain to be elucidated. Detailed investigations at single-cell level are warranted to shed light on how CNVs impact the expression of genes. When NGS cost reduces, more paired datasets will provide information on CNV and gene expression. This will remove probe bias and limitations of the microarray platforms and provide additional information on alternative splicing and miRNA expression. Whole genome sequence with NGS will provide additional information like SNPs, and balanced structural variations such as translocation and inversion. Furthermore, additional paired information such as epigenetic DNA methylation and histone modification will provide more detailed insights on gene expression.

Author Contributions

Conceived the concept: GEL. Analyzed the data: GEL. Wrote the first draft of the manuscript: GEL. Contributed to the writing of the manuscript: GEL, LX, KSH. Agree with manuscript results and conclusions: GEL, LX, KSH. Jointly developed the structure and arguments for the paper: GEL, LX, KSH. Made critical revisions and approved final version: GEL, LX, KSH. All authors reviewed and approved of the final manuscript.

REFERENCES

1. Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, Burdick JT. Mapping determinants of human gene expression by regional and genome-wide association. *Nature*. 2005;437(7063):1365–1369.
2. Schadt EE, Molony C, Chudin E, et al. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol*. 2008;6(5):e107.
3. Myers AJ, Gibbs JR, Webster JA, et al. A survey of genetic human cortical gene expression. *Nat Genet*. 2007;39(12):1494–1499.
4. Scherer SW, Lee C, Birney E, et al. Challenges and standards in integrating surveys of structural variation. *Nat Genet*. 2007;39(7 suppl):S7–S15.
5. Mills RE, Walter K, Stewart C, et al. Mapping copy number variation by population-scale genome sequencing. *Nature*. 2011;470(7332):59–65.
6. Sheltzer JM, Torres EM, Dunham MJ, Amon A. Transcriptional consequences of aneuploidy. *Proc Natl Acad Sci U S A*. 2012;109(31):12644–12649.
7. Torres EM, Williams BR, Amon A. Aneuploidy: cells losing their balance. *Genetics*. 2008;179(2):737–746.
8. Henrichsen CN, Chaignat E, Reymond A. Copy number variants, diseases and gene expression. *Hum Mol Genet*. 2009;18(R1):R1–R8.
9. Harewood L, Chaignat E, Reymond A. Structural variation and its effect on expression. *Methods Mol Biol*. 2012;838:173–186.
10. Sebat J, Lakshmi B, Troge J, et al. Large-scale copy number polymorphism in the human genome. *Science*. 2004;305(5683):525–528.
11. Tuzun E, Sharp AJ, Bailey JA, et al. Fine-scale structural variation of the human genome. *Nat Genet*. 2005;37(7):727–732.
12. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. *Nature*. 2006;444(7118):444–454.
13. Kidd JM, Cooper GM, Donahue WF, et al. Mapping and sequencing of structural variation from eight human genomes. *Nature*. 2008;453(7191):56–64.
14. Conrad DF, Pinto D, Redon R, et al. Origins and functional impact of copy number variation in the human genome. *Nature*. 2009;464(7289):704–712.
15. Dopman EB, Hartl DL. A portrait of copy-number polymorphism in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A*. 2007;104(50):19920–19925.
16. Graubert TA, Cahan P, Edwin D, et al. A high-resolution map of segmental DNA copy number variation in the mouse genome. *PLoS Genet*. 2007;3(1):e3.
17. She X, Cheng Z, Zollner S, Church DM, Eichler EE. Mouse segmental duplication and copy number variation. *Nat Genet*. 2008;40(7):909–914.
18. Guryev V, Saar K, Adamovic T, et al. Distribution and functional impact of DNA copy number variation in the rat. *Nat Genet*. 2008;40(5):538–545.
19. Yalcin B, Wong K, Agam A, et al. Sequence-based characterization of structural variation in the mouse genome. *Nature*. 2011;477(7364):326–329.
20. Zhou J, Lemos B, Dopman EB, Hartl DL. Copy-number variation: the balance between gene dosage and expression in *Drosophila melanogaster*. *Genome Biol Evol*. 2011;3:1014–1024.
21. Perry GH, Yang F, Marques-Bonet T, et al. Copy number variation and evolution in humans and chimpanzees. *Genome Res*. 2008;18(11):1698–1710.
22. Lee AS, Gutierrez-Arcelus M, Perry GH, et al. Analysis of copy number variation in the rhesus macaque genome identifies candidate loci for evolutionary and human disease studies. *Hum Mol Genet*. 2008;17(8):1127–1136.
23. Nicholas TJ, Cheng Z, Ventura M, Mealey K, Eichler EE, Akey JM. The genomic architecture of segmental duplications and associated copy number variants in dogs. *Genome Res*. 2009;19(3):491–499.
24. Chen WK, Swartz JD, Rush LJ, Alvarez CE. Mapping DNA structural variation in dogs. *Genome Res*. 2009;19(3):500–509.
25. Liu GE, Hou Y, Zhu B, et al. Analysis of copy number variations among diverse cattle breeds. *Genome Res*. 2010;20(5):693–703.
26. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The database of genomic variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res*. 2014;42(Database Issue):D986–D992.
27. Church DM, Lappalainen I, Sneddon TP, et al. Public data archives for genomic structural variation. *Nat Genet*. 2010;42(10):813–814.
28. Lappalainen I, Lopez J, Skipper L, et al. DbVar and DGVa: public archives for genomic structural variation. *Nucleic Acids Res*. 2013;41(Database Issue): D936–D941.
29. Pinto D, Darvishi K, Shi XH, et al. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. *Nat Biotechnol*. 2011;29(6):512–U76.
30. Zhang F, Carvalho CM, Lupski JR. Complex human chromosomal and genomic rearrangements. *Trends Genet*. 2009;25(7):298–307.
31. Sharp AJ, Locke DP, McGrath SD, et al. Segmental duplications and copy-number variation in the human genome. *Am J Hum Genet*. 2005;77(1):78–88.
32. Kim PM, Lam HY, Urban AE, et al. Analysis of copy number variants and segmental duplications in the human genome: evidence for a change in the process of formation in recent evolutionary history. *Genome Res*. 2008;18(12):1865–1874.
33. Liu GE, Ventura M, Cellamare A, et al. Analysis of recent segmental duplications in the bovine genome. *BMC Genomics*. 2009;10:571.
34. Yang Y, Chung EK, Wu YL, et al. Gene copy-number variation and associated polymorphisms of complement component C4 in human systemic lupus erythematosus (SLE): low copy number is a risk factor for and high copy number is a protective factor against SLE susceptibility in European Americans. *Am J Hum Genet*. 2007;80(6):1037–1054.
35. Craddock N, Hurles ME, Cardin N, et al. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature*. 2010;464(7289):713–720.
36. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med*. 2010;61:437–455.
37. Le Marechal C, Masson E, Chen JM, et al. Hereditary pancreatitis caused by triplication of the trypsinogen locus. *Nat Genet*. 2006;38(12):1372–1374.
38. Fellermann K, Stange DE, Schaeffeler E, et al. A chromosome 8 gene-cluster polymorphism with low human beta-defensin 2 gene copy number predisposes to Crohn disease of the colon. *Am J Hum Genet*. 2006;79(3):439–448.
39. Fanciulli M, Norsworthy PJ, Petretto E, et al. FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity. *Nat Genet*. 2007;39(6):721–723.
40. Aitman TJ, Dong R, Vyse TJ, et al. Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. *Nature*. 2006;439(7078):851–855.
41. McCarroll SA, Kuruvilla FG, Korn JM, et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet*. 2008;40(10):1166–1174.
42. Estivill X, Armengol L. Copy number variants and common disorders: filling the gaps and exploring complexity in genome-wide association studies. *PLoS Genet*. 2007;3(10):1787–1799.
43. Locke DP, Sharp AJ, McCarroll SA, et al. Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome. *Am J Hum Genet*. 2006;79(2):275–290.
44. Campbell CD, Sampas N, Tsalenko A, et al. Population-genetic properties of differentiated human copy-number polymorphisms. *Am J Hum Genet*. 2011;88(3): 317–332.
45. de Vries BBA, Pfundt R, Leisink M, et al. Diagnostic genome profiling in mental retardation. *Am J Hum Genet*. 2005;77(4):606–616.
46. Pinto D, Pagnamenta AT, Klei L, et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*. 2010;466(7304):368–372.
47. Sharp AJ, Hansen S, Selzer RR, et al. Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat Genet*. 2006;38(9):1038–1042.

48. Sebat J, Lakshmi B, Malhotra D, et al. Strong association of de novo copy number mutations with autism. *Science*. 2007;316(5823):445–449.
49. Cook EH Jr, Scherer SW. Copy-number variations associated with neuropsychiatric conditions. *Nature*. 2008;455(7215):919–923.
50. Diskin SJ, Hou C, Glessner JT, et al. Copy number variation at 1q21.1 associated with neuroblastoma. *Nature*. 2009;459(7249):987–991.
51. Shi J, Levinson DF, Duan J, et al. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature*. 2009;460(7256):753–757.
52. Glessner JT, Wang K, Cai G, et al. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature*. 2009;459(7246):569–573.
53. Stefansson H, Ophoff RA, Steinberg S, et al. Common variants conferring risk of schizophrenia. *Nature*. 2009;460(7256):744–747.
54. Bochukova EG, Huang N, Keogh J, et al. Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature*. 2009;463(7281):666–670.
55. Stranger BE, Forrest MS, Dunning M, et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science*. 2007;315(5813):848–853.
56. Henrichsen CN, Vinckenbosch N, Zollner S, et al. Segmental copy number variation shapes tissue transcriptomes. *Nat Genet*. 2009;41(4):424–429.
57. Chaignat E, Yahya-Graison EA, Henrichsen CN, et al. Copy number variation modifies expression time courses. *Genome Res*. 2011;21(1):106–113.
58. Orozco LD, Cokus SJ, Ghazalpour A, et al. Copy number variation influences gene expression and metabolic traits in mice. *Hum Mol Genet*. 2009;18(21):4118–4129.
59. Beckmann JS, Estivill X, Antonarakis SE. Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nat Rev Genet*. 2007;8(8):639–646.
60. Schuster-Bockler B, Conrad D, Bateman A. Dosage sensitivity shapes the evolution of copy-number varied regions. *PLoS One*. 2010;5(3):e9474.
61. Sexton T, Umlauf D, Kurukuti S, Fraser P. The role of transcription factories in large-scale structure and dynamics of interphase chromatin. *Semin Cell Dev Biol*. 2007;18(5):691–697.
62. Kleinjan DJ, van Heyningen V. Position effect in human genetic disease. *Hum Mol Genet*. 1998;7(10):1611–1618.
63. Bhatia S, Bengani H, Fish M, et al. Disruption of autoregulatory feedback by a mutation in a remote, ultraconserved PAX6 enhancer causes aniridia. *Am J Hum Genet*. 2013;93(6):1126–1134.
64. Benko S, Fantes JA, Amiel J, et al. Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. *Nat Genet*. 2009;41(3):359–364.
65. Gabellini D, Green MR, Tupler R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell*. 2002;110(3):339–348.
66. Gabellini D, D'Antona G, Moggio M, et al. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. *Nature*. 2006;439(7079):973–977.
67. Lee JA, Lupski JR. Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. *Neuron*. 2006;52(1):103–121.
68. Fraser P, Bickmore W. Nuclear organization of the genome and the potential for gene regulation. *Nature*. 2007;447(7143):413–417.
69. Heard E, Bickmore W. The ins and outs of gene regulation and chromosome territory organisation. *Curr Opin Cell Biol*. 2007;19(3):311–316.
70. Cahan P, Li Y, Izumi M, Graubert TA. The impact of copy number variation on local gene expression in mouse hematopoietic stem and progenitor cells. *Nat Genet*. 2009;41(4):430–437.
71. Reymond A, Henrichsen CN, Harewood L, Merla G. Side effects of genome structural changes. *Curr Opin Genet Dev*. 2007;17(5):381–386.
72. Park H, Kim JI, Ju YS, et al. Discovery of common Asian copy number variants using integrated high-resolution array CGH and massively parallel DNA sequencing. *Nat Genet*. 2010;42(5):400–405.
73. Sudmant PH, Kitzman JO, Antonacci F, et al. Diversity of human copy number variation and multicopy genes. *Science*. 2010;330(6004):641–646.
74. Schlattl A, Anders S, Waszak SM, Huber W, Korbel JO. Relating CNVs to transcriptome data at fine resolution: assessment of the effect of variant size, type, and overlap with functional regions. *Genome Res*. 2011;21(12):2004–2013.
75. Pickrell JK, Marioni JC, Pai AA, et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature*. 2010;464(7289):768–772.
76. Montgomery SB, Sammeth M, Gutierrez-Arcelus M, et al. Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature*. 2010;464(7289):773–777.
77. Hulse AM, Cai JJ. Genetic variants contribute to gene expression variability in humans. *Genetics*. 2013;193(1):95–108.
78. Raser JM, O'Shea EK. Noise in gene expression: origins, consequences, and control. *Science*. 2005;309(5743):2010–2013.
79. Huang N, Shah PK, Li C. Lessons from a decade of integrating cancer copy number alterations with gene expression profiles. *Brief Bioinform*. 2012;13(3):305–316.
80. Iskow RC, Gokcumen O, Abyzov A, et al. Regulatory element copy number differences shape primate expression profiles. *Proc Natl Acad Sci U S A*. 2012;109(31): 12656–12661.
81. Gokcumen O, Tischler V, Tica J, et al. Primate genome architecture influences structural variation mechanisms and functional consequences. *Proc Natl Acad Sci USA*. 2013;110(39):15764–15769.
82. Xue Z, Huang K, Cai C, et al. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature*. 2013;500(7464):593–597.
83. Lu S, Zong C, Fan W, et al. Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. *Science*. 2012;338(6114):1627–1630.
84. Gole J, Gore A, Richards A, et al. Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. *Nat Biotechnol*. 2013;31(12):1126–1132.
85. Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. *Nature*. 2011;472(7341):90–94.
86. Ni X, Zhuo M, Su Z, et al. Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. *Proc Natl Acad Sci U S A*. 2013;110(52):21083–21088.
87. Xu X, Ge H, Gu C, et al. Modeling spatial correlation of DNA deformation: DNA allostery in protein binding. *J Phys Chem B*. 2013;117(42):13378–13387.
88. Zong C, Lu S, Chapman AR, Xie XS. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science*. 2012;338(6114):1622–1626.
89. McConnell MJ, Lindberg MR, Brennand KJ, et al. Mosaic copy number variation in human neurons. *Science*. 2013;342(6158):632–637.
90. Evrony GD, Cai X, Lee E, et al. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell*. 2012;151(3):483–496.
91. Deng Q, Ramskold D, Reinius B, Sandberg R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science*. 2014;343(6167):193–196.
92. Guo H, Zhu P, Wu X, Li X, Wen L, Tang F. Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res*. 2013;23(12):2126–2135.