Using iPSCs and genomics to catch CNVs in the act

Alexander Eckehart Urban & Carolin Purmann

Large copy number variants (CNVs) are strongly associated with morphogenetic processes and common neurodevelopmental disorders. A new study uses the example of Williams-Beuren syndrome (WBS) and Williams-Beuren region duplication syndrome to illustrate how induced pluripotent stem cells (iPSCs) and next-generation genomics can lead to a better understanding of complex genetics.

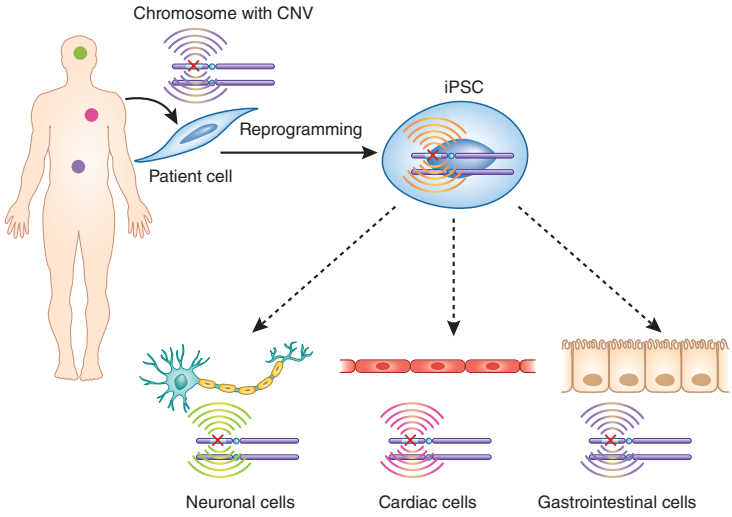
CNVs are deletions or duplications of stretches of genome sequence relative to the reference genome, normally ranging in size from hundreds of base pairs to, rarely, hundreds of thousands of base pairs[1–4]. Several even larger CNVs are strongly associated with diseases, particularly developmental and neuropsychiatric disorders[5,6]. A detailed understanding of the effects of large CNVs on the organism could lead to important general insights into these common but still very poorly understood disorders. However, molecular and functional analysis of large CNVs is not straightforward. A new study by Giuseppe Testa and colleagues[7] aimed to avoid the limitations of existing approaches by taking advantage of two revolutionary new methods: iPSC-based modeling and next-generation genomics. The authors used iPSCs derived from patients with disease-associated large CNVs, differentiated them into relevant cell lineages and carried out high-resolution, comprehensive molecular analysis of the effects of the large CNVs in several precursor cell types (**Fig. 1**).

There are typically dozens of genes within the boundaries of a large CNV, and identifying which are relevant to the disease under study is a monumental challenge. Existing approaches have limitations, and a complementing concept would be very useful. Animal models are powerful tools for neurophysiology, but it is not clear how far they can be taken in the modeling of higher brain functions and complex behavior in humans[8,9]. Primary tissues such as adult autopsy brain or blood in many ways provide the correct context for studying the effects of large CNVs, but only by means of a snapshot of their effects. Adult autopsy tissues will probably be less useful in identifying when and how during brain development a large CNV contributes to the disease phenotype. Meanwhile, blood samples will only show those effects of a large CNV that are either relevant to blood or not cell-type specific.

CNV effects caught in the act

Large CNVs on chromosome 7q11.23 occur as deletions (WBS) or duplications (Williams-Beuren region duplication syndrome). They are around 1.6 million base pairs in size and contain approximately 28 genes[10,11]. The spectrum of associated symptoms is quite variable and ranges from facial dysmorphology to problems with heart function and hearing and to cognitive and behavioral consequences[12,13].

The authors assembled a rather large cohort of iPSC lines by the current standards of the iPSC field (considering workload and expense), including lines for five WBS cases, two Williams-Beuren region duplication syndrome cases and six controls; for most cases, duplicate or triplicate clonal lines were obtained. The iPSC lines were differentiated into neural progenitor cells (NPCs), neural crest stem cells (NCSCs) and mesenchymal stem cells (MSCs). In the various cell types, changes in transcript levels were determined by transcriptome-wide RNA sequencing. Genome-wide binding patterns and interaction partners for a transcription factor encoded within the large CNV, GTF2I, were investigated using chromatin immunoprecipitation and sequencing (ChIP-seq), RNA interference (RNAi) manipulations and proteomics.



Marina Corral Spence/Nature Publishing Group

Figure 1 Analyzing the molecular effects of CNVs using iPSC models and genomics. Primary cells from patients with a disease-associated large CNV are reprogrammed into iPSCs and differentiated toward various relevant cell types. The cell type–specific molecular effects of the CNV (represented by different colored 'radiowaves') can then be studied using genomic technologies.

Alexander Eckehart Urban and Carolin Purmann are in the Departments of Psychiatry and Behavioral Sciences and the Department of Genetics, Stanford University, Stanford, California, USA.
e-mail: aeurban@stanford.edu

Gene dosage effects seeded, retained and amplified

Almost all of the genes within the boundaries of the large CNVs had changes in their expression levels consistent with the given change in gene dosage. For example, *GTF2I* expression levels were approximately three times as high in iPSCs derived from patients with the duplication (Williams-Beuren region duplication syndrome) than in cells derived from patients with the deletion (WBS). However, this was not the case for all genes in the region, which should make for very interesting follow-up analyses.

At the genome-wide level, there were several hundred differentially expressed genes (DEGs), suggesting a network effect emanating from the CNV at 7q11.23. Some of these expression changes were cell-type specific. Interestingly, many of the affected molecular pathways already showed signs of dysregulation in the iPSC state. This transcriptional dysregulation had a tendency to become amplified in the more differentiated cell states, such that the retained DEGs would often possess a lineage-specific function. For example, DEGs and Gene Ontology (GO) categories related to axon formation emerged more clearly in NPCs, as did those for synapses in NCSCs and for smooth muscle tissue in MSCs. Changes in GTF2I genome-wide binding patterns also indicated lineage-specific effects of the large CNV. Interestingly, overlap with the DEG patterns was limited, pointing to a more indirect effect of this transcription factor on transcription networks. The authors were able to identify a few specific target genes as potentially relevant for the disease phenotypes, namely *PDLIM1* (which acts in neurites and is associated with cardiovascular defects), *MYH14* (involved in hearing) and, in particular, *BEND4* (another transcription factor, involved in neural processes). The targeting of *BEND4* by GTF2I provides a glimpse at how the effects of the large CNV might ripple across longer distances on the molecular network.

Toward a new research paradigm

This study shows that the somewhat notorious variance between iPSC lines[14,15] is not too severe to mask clear effects of large CNVs, some of which are tissue specific and related to disease. There are some deviations between expression levels and discrete dosage changes at the DNA level that remain unexplained, but such discrepancies point the way to interesting studies to be done on the epigenetic level.

The study identified a few specific genes that can be further explored in the context of WBS and Williams-Beuren region duplication syndrome, although, before too much effort is expanded on these individual genes, it would be prudent to explore the effects of the 7q11.23 CNVs in disease-relevant terminal cell differentiation states, not just in precursor cells. In such terminal differentiation states, it will then also be possible and important to carry out functional analyses such as neurophysiological measurements in relevant neuronal subtypes.

Larger cohorts will be used once iPSC methods have sufficiently evolved, and there will have to be an accounting for the typically rather large degree of phenotypic variance between patients with nearly identical large CNVs. But the general approach is sound and already quite reliable. It will be very interesting to see what parallels and differences are shown by the studies that are well under way in laboratories around the world using the same concept for other large CNVs.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Iafrate, A.J. *et al. Nat. Genet.* **36**, 949–951 (2004).
2. Sebat, J. *et al. Science* **305**, 525–528 (2004).
3. Scherer, S.W. *et al. Nat. Genet.* **39**, S7–S15 (2007).
4. 1000 Genomes Project Consortium. *Nature* **491**, 56–65 (2012).
5. Malhotra, D. & Sebat, J. *Cell* **148**, 1223–1241 (2012).
6. Southard, A.E., Edelmann, L.J. & Gelb, B.D. *Pediatrics* **129**, 755–763 (2012).
7. Adamo, A. *et al. Nat. Genet.* **47**, 132–141 (2015).
8. Tordjman, S. *et al. Behav. Genet.* **37**, 61–78 (2007).
9. Nestler, E.J. & Hyman, S.E. *Nat. Neurosci.* **13**, 1161–1169 (2010).
10. Pérez Jurado, L.A. *et al. Am. J. Hum. Genet.* **59**, 781–792 (1996).
11. Urbán, Z. *et al. Am. J. Hum. Genet.* **59**, 958–962 (1996).
12. Osborne, L.R. *Mol. Genet. Metab.* **67**, 1–10 (1999).
13. Merla, G. *et al. Hum. Genet.* **128**, 3–26 (2010).
14. Thatava, T. *et al. Mol. Ther.* **21**, 228–239 (2013).
15. Cahan, P. & Daley, G.Q. *Nat. Rev. Mol. Cell Biol.* **14**, 357–368 (2013).

New insight into a complex plant–fungal pathogen interaction

Peter J Balint-Kurti & James B Holland

The coevolution of plants and microbes has shaped plant mechanisms that detect and repel pathogens. A newly identified plant gene confers partial resistance to a fungal pathogen not by preventing initial infection but by limiting its spread through the plant.

Plants, like other higher organisms, display a tremendous diversity of associations with microbes. At one end of the spectrum is mutualism, as observed between nitrogen-fixing rhizobacteria and their legume hosts. At the other extreme is the parasitism of obligate biotrophs, such as the rust fungi that colonize and derive nutrients from living host plant tissue, and of necrotrophs, such as the botrytis fungi that kill and rot plant tissues as a means to obtain their nutrients. Endophytic fungi grow systemically throughout a plant without causing obvious disease symptoms and in some cases appear to benefit their hosts by producing compounds that inhibit insect herbivory[1] or contribute to stress tolerance[2]. On page 151 of this issue, Mingliang Xu and colleagues identify a plant gene, *ZmWAK*, that regulates the interaction between maize and the fungus *Sporisorium reilianum*, causal agent of head smut disease[3]. *S. reilianum* infects the roots of maize seedlings and grows systemically, often causing only subtle observable effects on host morphology or physiology[4,5] until flowering, when the fungus forms spectacular large black sori filled with teliospores that replace the ears and tassels of the host (**Fig. 1**).

Peter J. Balint-Kurti is in the Plant Science Research Unit, US Department of Agriculture–Agricultural Research Service (USDA-ARS) and the Department of Plant Pathology, North Carolina State University, Raleigh, North Carolina, USA. James B. Holland is in the Plant Science Research Unit, US Department of Agriculture–Agricultural Research Service (USDA-ARS) and the Department of Crop Science, North Carolina State University, Raleigh, North Carolina, USA.
e-mail: peter_balintkurti@ncsu.edu

Plant defense responses

Plant pattern recognition receptors recognize general microbial features (pathogen-associated molecular patterns), triggering a basal defense response sufficient to prevent most pathogen infections. Pathogen-derived 'effector' proteins