European Journal of Human Genetics (2017) 25, 1179–1180
© 2017 European Society of Human Genetics All rights reserved 1018-4813/17
www.nature.com/ejhg

NEWS AND COMMENTARY

TAD-pathways for GWAS

The TAD-pathway for GWAS signals

Natalia Pervjakova and Inga Prokopenko*

European Journal of Human Genetics (2017) **25,** 1179–1180; doi:10.1038/ejhg.2017.148

A Latin proverb says a well-beaten path does not always make the right road. The same idea is raised by Way *et al.*[1] article, this issue, where the authors address the question of prioritisation of genes underlying signals identified by genome-wide association studies (GWAS). In many cases, there is an accepted standard when GWAS reports novel signals within non-coding regions and names these loci by the nearest gene. The idea behind such a simplification is that the functional interpretation of DNA variability is challenging and often research groups do not have enough supportive datasets or methods available to assign genomic effects to the underlying pathways or causal genes. Way *et al.*[1] address this issue by developing an innovative computational method that performs a Gene Ontology (GO) analysis and implicates the genetic content of topologically associating domains (TADs)[2] through a pathway overrepresentation test.

TADs are large self-interacting chromatin domains that were proposed for the first time following a series of Hi-C data experiments.[2] Both the strong conservation between species and the remarkable stability of such regions provide the connection to the fundamental organising principle of a genome. Briefly, TADs are collections of many chromatin loops separated by clear borders enriched with the insulator binding protein CTCF, housekeeping genes, transfer RNAs, constant timing regions (CTRs), and short interspersed nuclear element (SINE) repeats (Figure 1).[2–5] Moreover, TADs are stable during replication-timing period and are suggested to coordinate gene expression.[4] Indeed, a large extent of the TADs borders are conserved between different cell types in humans and even more between mouse and human species making them a valuable source for dissecting molecular mechanisms underlying variability of a trait or disease.[2,5,6]

The method implemented in the TAD Pathways, proposed by Way *et al*, is a simple and elegant solution for the researches interested in revealing the causal genes underlying association signals that are located in both coding and non-coding regions.[1] The criticism of this approach could be based on the usage of TAD regions previously determined by Dixon *et al.*;[2] however, the sophism of this situation is that we are yet limited in understanding of the exact TADs organisation. While working with any mapping of the TADs, it is important to remember that there is a variety of tools assessing the TADs organisational division, whereas almost each tool is based on a range of assumptions that are specific and important in each particular experiment, that is, type of Hi-C signal, size distribution, chromosomal organisation of different species etc.[7–9] Thus, in the context of today's reality, the implementation of already established TAD boundaries is a rational way to improve the quality of functional characterisation for association signals in GWAS reports.

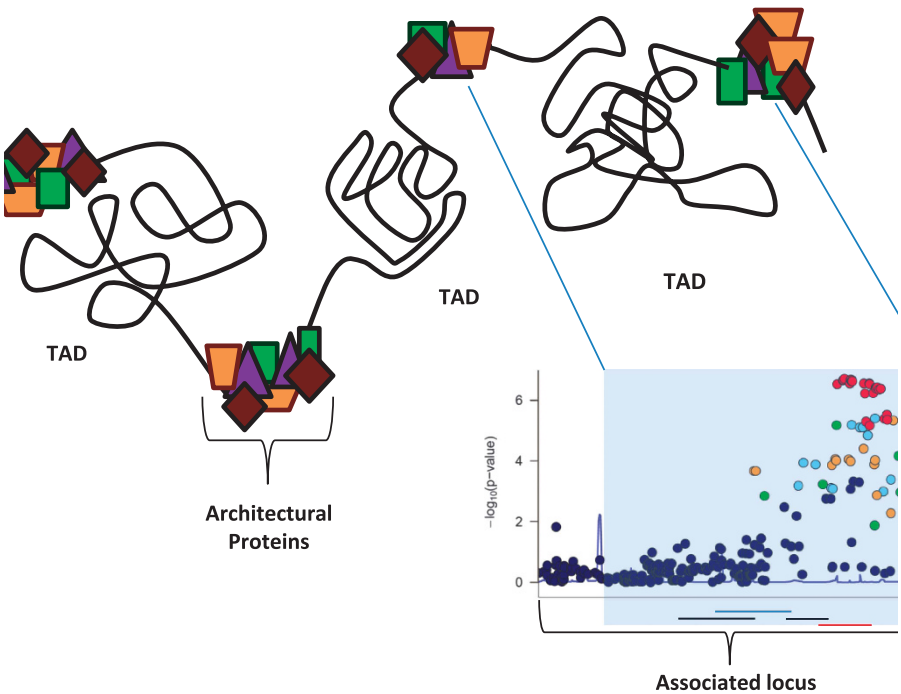

Figure 1 The schematic representation of topologically associating domains, TADs, in relation to a signal of association in GWAS. TADs are collections of chromatin loops (represented by the blue lines) separated by borders enriched with a range of regulatory elements, also called as architectural proteins. The association pattern observed in GWAS within a region is shown in a plot, where blue background represents the region covered by TAD, *y* axis shows the strength of association, *x* axis—relative positions of polymorphic variants, recombination pattern (blue patterned line) and genes (horizontal lines, with red line representing the gene closest to the lead variant, name of such a gene usually is used for a locus, blue line—gene suggested within a TAD region, and black lines—other genes in the region).

N Pervjakova is at the Department of Medicine, Section of Genomics of Common Disease, Imperial College London, London, UK; Estonian Genome Center, University of Tartu, Tartu, Estonia; Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia and I Prokopenko, Department of Medicine, Section of Genomics of Common Disease, Imperial College London, London, UK
E-mail: i.prokopenko@imperial.ac.uk

The TAD Pathways method is the first step to relate the 3D genomic structure genome-wide to the annotation of novel associated variants. The authors demonstrate application of their approach on the example of bone mineral density (BMD) phenotype where they define the previously suggested *ARHGAP1* gene at the BMD-associated rs7932354 variant as a false-positive, and implicate as causal *ACP2* gene, a novel regulator of osteoblast metabolism, located nearby.[1] Such situations could potentially be exemplar to the researchers studying phenotypes that do not have adequate animal models, for instance, age at natural menopause or traits associated with human-specific imprinted genes.[10,11] In addition, while analyses of *cis*- and *trans*-eQTLs are among the most popular follow-ups to determine the altered expression of a gene affected by a detected signal, their power is often limited by the tissue in which expression was measured as it does not necessarily correspond to the expression pattern predicted for a specific variant.[12] An improved mapping of TADs could also potentially be of interest to address this issue due to their ubiquitous distribution across cell types. Whilst the quality of Hi-C resolution is constantly improving, it is worth expecting development of other competitive approaches and therefore more discoveries relating TADs genic content to novel and already established GWAS signals.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Way GP, Youngstrom DW, Hankenson KD *et al*: Implicating genes at GWAS signals by leveraging topologically associated domains. *Eur J Hum Genet*, this issue.
2 Dixon JR, Selvaraj S, Yue F *et al*: Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 2012; **485**: 376–380.
3 Nguyen HQ, Bosco G: Gene Positioning Effects on Expression in Eukaryotes. *Annu Rev Genet* 2015; **49**: 627–646.
4 Pope BD, Ryba T, Dileep V *et al*: Topologically associating domains are stable units of replication-timing regulation. *Nature* 2014; **515**: 402–405.
5 Bonev B, Cavalli G: Organization and function of the 3D genome. *Nat Rev Genet* 2016; **17**: 661–678.
6 Lonfat N, Duboule D: Structure, function and evolution of topologically associating domains (TADs) at HOX loci. *FEBS Lett* 2015; **589**: 2869–2876.
7 Dali R, Blanchette M: A critical assessment of topologically associating domain prediction tools. *Nucleic Acids Res* 2016; **45**: 2994–3005.
8 Ciabrelli F, Cavalli G: Chromatin-driven behavior of topologically associating domains. *J Mol Biol* 2015; **427**: 608–625.
9 Wang XT, Dong PF, Zhang HY, Peng C: Structural heterogeneity and functional diversity of topologically associating domains in mammalian genomes. *Nucleic Acids Res* 2015; **43**: 7237–7246.
10 Day FR, Ruth KS, Thompson DJ *et al*: Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. *Nat Genet* 2015; **47**: 1294–1303.
11 Peters J: The role of genomic imprinting in biology and disease: an expanding view. *Nat Rev Genet* 2014; **15**: 517–530.
12 Bonder MJ, Kasela S, Kals M *et al*: Genetic and epigenetic regulation of gene expression in fetal and adult human livers. *BMC Genomics* 2014; **15**: 860.