„Weihnachts“ – Oberseminar / JournalClub

22.12.2016

Recent Developments in Science, Nature etc.

Ralf Zimmer

Science, November 11, 2016

- **SCIENCE 25.11.2016**
- **SCIENCE 18.11.2016**
- **SCIENCE 11.11.2016**
- **p.769 Enhancer function: Systematic mapping of functional enhancer-promoter connections with CRISPR interference, Eric Lander & Jesse Engreitz**
- **p.712 BOOK: Redesigning life: how genome editing will transform the world, CRISPR-Cas**
- **p.703 Web scientists, data scientists**
-

Science, November 4, 2016

- p. 532: Your entire genome in 3D
- p. 539: Mice made easy: genome-editing with CRISPR
- p. 557-8: one brain, many genomes: single-cell genomic techniques, somatic mutations
- p. 559: BOOK: Walter Alvarez, A most improbable journey (Burgess shale)
- p. 560: BOOK: Ezrachi&Stucke, Virtual competition: The promise and perils of the algorithm-driven economy
- p. 564-592: Pain Research
- p. 595: Keeping hearts and blood vessels young, Sci. Signal. 9, ra105 (2016)

Science, November 4, 2016

- p.596: Quantifying the evolution of individual scientific impact
- p. 553 + p. 597: Posttranslational mutagenesis: … exploring protein side-chain diversity
- p. 598: A transcription factor hierarchy defines an environmental stress response network (Ziv Bar-Joseph)
- p. 599: Exploring genetic suppression interactions on a global scale (C. Boone)
- p. 603: Optical processing: A coherent Ising machine for 2000-node optimization problems
  p. 614: Optical processing: A fully programmable 100-spin coherent Ising machine with all-to-all connections
- p. 618: Mutational signatures associated with tobacco smoking in human cancer

Science, October 28, 2016

- p. 391: BOOKS: Roger Penrose: Fashion, faith, fantasy
- p. 391: BOOKS: Tom Roossinck: Virus, an illustrated guide
- p. 391: BOOKS: Beth Shapiro: How to clone a mammoth
- p. 395: DNA.land, 3 mio consumer genetics screening
- p. 398: baby genome sequencing / screening needs more time to gestate
- p. 431: personalized medicine by another name: precision medicine, Hastings Cent Rep. 46, 21 (2016)
- p. 432: systems-level analysis of mechanisms regulating yeast metabolic flux -> VL Systems Biology
- p. 472: Senescent intimal foam cells are deleterious at all stages of atherosclerosis

Nature, December 15, 2016

- p.433 A 17-gene stemness score for rapid determination of risk in acute leukaemia
- p.423 Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism, Blencowe et al
- p. 371,379,386 Reviews: Printing soft matter in 3D, Plastic bioelectronics, Polymers for biomedical applications
- p.346 A gene-expression profile for leukemia, Ng. et al, 17 gene signature
- p.328: Predictive algorithms … for personalized cancer vaccines, cancer contest.

Nature, December 8, 2016

- **p.296 Capturing pairwise and multi-way chromosomal conformations using chromosomal walks, (anchoring topological domains: TADs), Amos Tanay**
- **p.S64-S67 Ethics: rewriting the regenerative rulebook, adult stem cells**
- **p.171 Editorial: Populism, correct misinformation**

Nature, December 1, 2016

- p.153 Technology feature. Metabolomics: small molecules, single cells
- p.151 Toolbox: a spellchecker for statistics
- p.69 The genomic basis of circadian and circalunar timing, Arndt von Haeseler et al
- p.S1-S14 nature outlook: Multiple Sclerosis
- p.37 BOOK Dance to the tune of life: Biological Relativity,Denis Noble, Cambridge Univ.Press
- p.15 Leading Alzheimer's theory survives drug failure, Solanezumab (Eli Lilly) flopped in clinical trial

Nature, November 24, 2016

- **p. 467: The power of big data must be harnessed for medical progress. But how ?**

Nature, November 17, 2016

- p. 334: Computing cancer, IBM Watson, 50 mio $ study
- p. 452. Eric Lander: Local regulation of gene expression by lncRNA promoter, transcription and splicing

ANALYSIS

nature genetics

Understanding multicellular function and disease with human tissue-specific networks

Casey S Greene[1–3,13], Arjun Krishnan[4,13], Aaron K Wong[5,13], Emanuela Ricciotti[6,7], Rene A Zelaya[1], Daniel S Himmelstein[8], Ran Zhang[9], Boris M Hartmann[10], Elena Zaslavsky[10], Stuart C Sealfon[10], Daniel I Chasman[11], Garret A FitzGerald[6,7], Kara Dolinski[4], Tilo Grosser[6,7] & Olga G Troyanskaya[4,5,12]

NATURE GENETICS VOLUME 47 | NUMBER 6 | JUNE 2015

Troyanskaya, Nature Genetics, 2015

Tissue and cell-type identity lie at the core of human physiology and disease. Understanding the genetic underpinnings of complex tissues and individual cell lineages is crucial for developing improved diagnostics and therapeutics. We present genome-wide functional interaction networks for 144 human tissues and cell types developed using a data-driven Bayesian methodology that integrates thousands of diverse experiments spanning tissue and disease states. Tissue-specific networks predict lineage-specific responses to perturbation, identify the changing functional roles of genes across tissues and illuminate relationships among diseases. We introduce NetWAS, which combines genes with nominally significant genome-wide association study (GWAS) *P* values and tissue-specific networks to identify disease-gene associations more accurately than GWAS alone. Our webserver, GIANT, provides an interface to human tissue networks through multi-gene queries, network visualization, analysis tools including NetWAS and downloadable networks. GIANT enables systematic exploration of the landscape of interacting genes that shape specialized cellular functions across more than a hundred human tissues and cell types.

NATURE GENETICS VOLUME 47 | NUMBER 6 | JUNE 2015

Troyanskaya, PLoS Comp. Biol., 2012

OPEN ACCESS Freely available online

PLOS | COMPUTATIONAL BIOLOGY

Tissue-Specific Functional Networks for Prioritizing Phenotype and Disease Genes

Yuanfang Guan[1,2], Dmitriy Gorenshteyn[3,4], Margit Burmeister[1,5], Aaron K. Wong[3], John C. Schimenti[6], Mary Ann Handel[7], Carol J. Bult[7], Matthew A. Hibbs[7,8]*, Olga G. Troyanskaya[3,9]*

Abstract

Integrated analyses of functional genomics data have enormous potential for identifying phenotype-associated genes. Tissue-specificity is an important aspect of many genetic diseases, reflecting the potentially different roles of proteins and pathways in diverse cell lineages. Accounting for tissue specificity in global integration of functional genomics data is challenging, as "functionality" and "functional relationships" are often not resolved for specific tissue types. We address this challenge by generating tissue-specific functional networks, which can effectively represent the diversity of protein function for more accurate identification of phenotype-associated genes in the laboratory mouse. Specifically, we created 107 tissue-specific functional relationship networks through integration of genomic data utilizing knowledge of tissue-specific gene expression patterns. Cross-network comparison revealed significantly changed genes enriched for functions related to specific tissue development. We then utilized these tissue-specific networks to predict genes associated with different phenotypes. Our results demonstrate that prediction performance is significantly improved through using the tissue-specific networks as compared to the global functional network. We used a testis-specific functional relationship network to predict genes associated with male fertility and spermatogenesis phenotypes, and experimentally confirmed one top prediction, *Mbyl1*. We then focused on a less-common genetic disease, ataxia, and identified candidates uniquely predicted by the cerebellum network, which are supported by both literature and experimental evidence. Our systems-level, tissue-specific scheme advances over traditional global integration and analyses and establishes a prototype to address the tissue-specific effects of genetic perturbations, diseases and drugs.

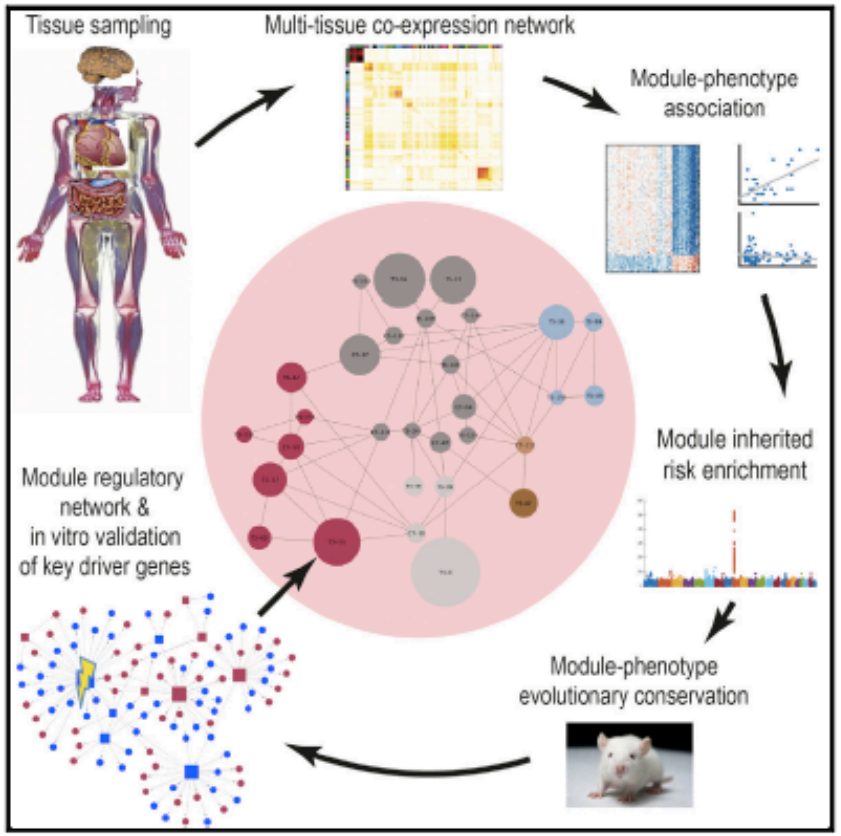
Talukdar, ... Bjorkegren, 2016

Article

Cell Systems

Cross-Tissue Regulatory Gene Networks in Coronary Artery Disease

Graphical Abstract



Authors

Husain A. Talukdar, Hassan Foroughi Asl, Rajeev K. Jain, ..., Josefin Skogsberg, Tom Michoel, Johan L.M. Björkegren

Correspondence

johan.bjorkegren@mssm.edu

In Brief

Well-established atherosclerosis risk factors and pathways are shown to operate through regulatory gene networks, active both within and across vascular and metabolic tissues, to cause coronary artery disease (CAD). Within these CAD-causal networks, the hierarchical order and connectivity patterns of both established and new genes in CAD, including so-called key disease driver genes, advance not only our global understanding of the molecular landscape in CAD but also reveal new candidate genes that may serve as suitable drug targets.

Talukdar, ... Bjorkegren, 2016

Highlights

- We reconstruct regulatory gene networks across seven vascular and metabolic tissues
- Integrative analysis using GWASs reveals 30 networks causally related to CAD
- 12 CAD-causal networks are indicated to be evolutionarily conserved from mouse
- An atherosclerotic arterial wall RNA-processing network affects foam cell formation

suitable drug targets.

Accession Numbers

GSE40231
GSE66570
GSE64768
GSE64769

Talukdar, ... Bjorkegren, 2016

SUMMARY

Inferring molecular networks can reveal how genetic perturbations interact with environmental factors to cause common complex diseases. We analyzed genetic and gene expression data from seven tissues relevant to coronary artery disease (CAD) and identified regulatory gene networks (RGNs) and their key drivers. By integrating data from genome-wide association studies, we identified 30 CAD-causal RGNs interconnected in vascular and metabolic tissues, and we validated them with corresponding data from the Hybrid Mouse Diversity Panel. As proof of concept, by targeting the key drivers *AIP*, *DRAP1*, *POLR2I*, and *PQBP1* in a cross-species-validated, arterial-wall RGN involving RNA-processing genes, we re-identified this RGN in THP-1 foam cells and independent data from CAD macrophages and carotid lesions. This characterization of the molecular landscape in CAD will help better define the regulation of CAD candidate genes identified by genome-wide association studies and is a first step toward achieving the goals of precision medicine.

Predicting effects of noncoding variants with deep learning–based sequence model

Jian Zhou[1,2] & Olga G Troyanskaya[1,3,4]

Identifying functional effects of noncoding variants is a major challenge in human genetics. To predict the noncoding-variant effects *de novo* from sequence, we developed a deep learning–based algorithmic framework, DeepSEA (http://deepsea.princeton.edu/), that directly learns a regulatory sequence code from large-scale chromatin-profiling data, enabling prediction of chromatin effects of sequence alterations with single-nucleotide sensitivity. We further used this capability to improve prioritization of functional variants including expression quantitative trait loci (eQTLs) and disease-associated variants.

Troyanskaya, Nature Methods, 2015

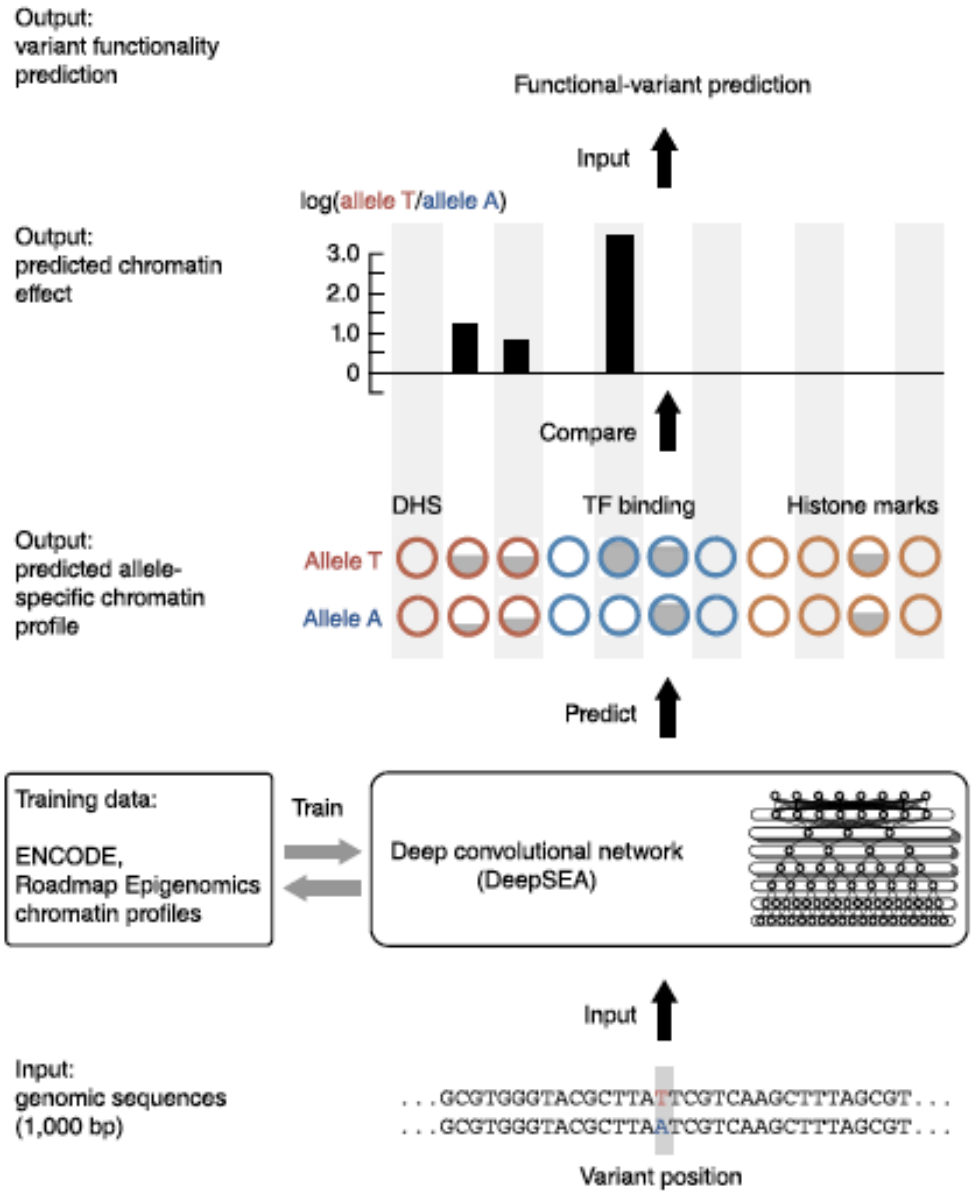
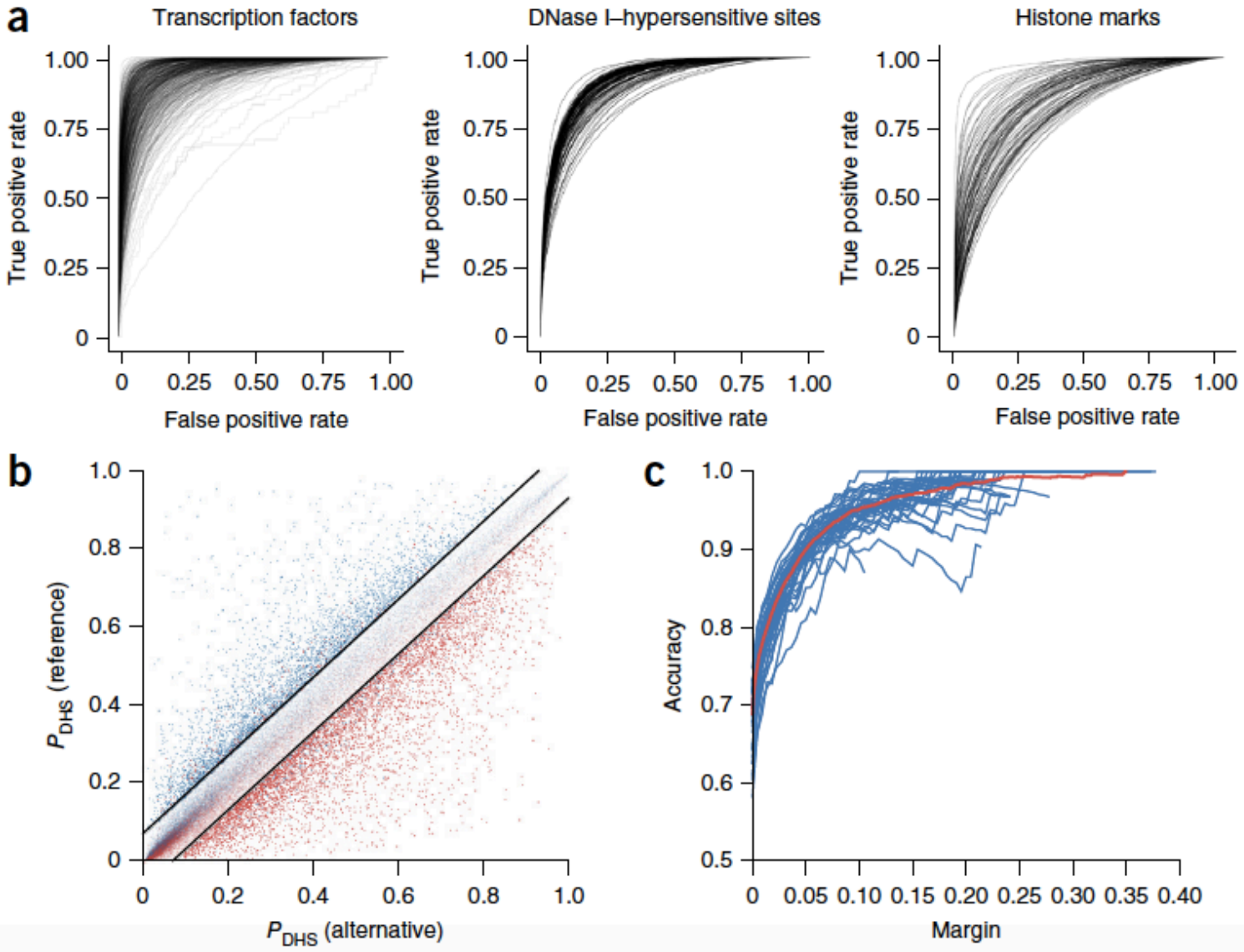


Figure 1 | Schematic overview of the DeepSEA pipeline, a strategy for predicting chromatin effects of noncoding variants.

Troyanskaya, Nature Methods, 2015



Troyanskaya, Nature Methods, 2015

- **Figure 2** | The deep-learning model accurately predicts chromatin features from sequence with single-nucleotide sensitivity. (**a**) Receiver operating characteristic (ROC) curves for each TF (left), DNase-seq (center) and histone-mark (right) profile prediction. Chromatin features with at least 50 test-positive samples were used. (**b**) DeepSEA predictions for DNase I–sensitive alleles of 57,407 allelically imbalanced variants from the digital genomic footprinting (DGF) DNase-seq data for 35 different cell types. The *y* and *x* axes show, respectively, for a variant, the predicted probabilities that the sequences carrying the reference allele and the alternative allele are DHSs within the corresponding cell type. The red and blue dots represent, respectively, the experimentally determined alternative allele–biased and reference allele–biased variants as determined by DGF data. The black lines indicate the margin, or the threshold of predicted probability differences between the two alleles for classifying high-confidence predictions (margin = 0.07 for this plot). (**c**) Accuracy. Each blue line indicates the performance for a different cell type, and the red line shows the overall performance on allelically imbalanced variants for all 35 cell types.

Troyanskaya, Nature Methods, 2015

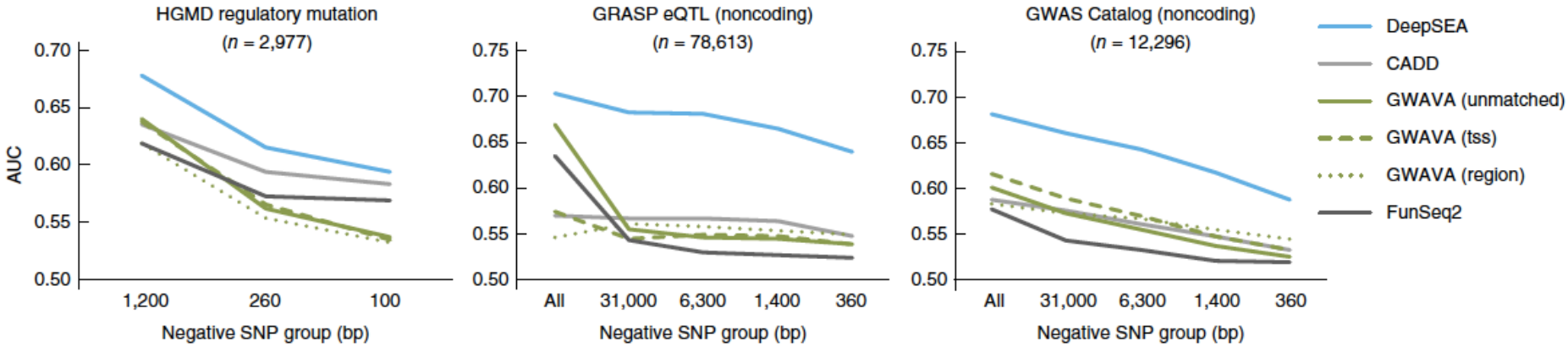


Figure 3 | Sequence-based prioritization of functional noncoding variants. Comparison of DeepSEA to other methods for prioritizing functionally annotated variants including HGMD annotated regulatory mutations, noncoding GRASP eQTLs and noncoding GWAS Catalog SNPs against noncoding 1000 Genomes Project SNPs (across multiple negative-variant groups with different scales of distances to the positive SNPs). The *x* axes show the average distances of negative-variant groups to a nearest positive variant. The "All" negative-variant groups are randomly selected negative 1000 Genomes SNPs. Because GWAVA was trained on the HGMD regulatory mutations, we filtered out GWAVA training positive-variant examples and closely located variants (within 2,000 bp) in evaluating its performance on HGMD regulatory mutations. Model performance is measured with area under the receiver operating characteristic curves (AUC).