

COMMENT

Progress in genomics according to bingo: 2013 edition

Konrad J Karczewski^{1,2}

'BINGO!' - Not a phrase you might expect to hear at some of the biggest genetics meetings, such as Cold Spring Harbor's Biology of Genomes and ASHG. But it is one that has been heard in conference centers and on Twitter, in response to buzzwords thrown around in talks. The rules are simple and familiar: generate a bingo card at <http://www.interpretome.com/#bingo> and, when a speaker (not a question from the audience) says something on your card, mark it down. If you get five in a row, column, or diagonal, you win: no actual prizes, except for bragging rights and the ability to tweet your winning card (Figure 1).

2013 saw many scientific advances in genomics and genetics, as well as policy decisions. In no particular order, here were some of the popular buzzwords.

Single-cell

With the plummeting costs of sequencing, researchers are turning toward more detailed cellular measurements using these technologies. Specifically, single-cell RNA-Seq methods [1] were developed this year, which have accurately measured transcriptomes in greater than zero but less than two cells using tube- and microfluidic-based methods. These methods found that, while pooled measurements of single-cell expression recapitulate typical aggregate expression, aggregate expression is a poor predictor of single-cell expression. Additionally, single-cell Hi-C [2] methods have provided insight into cell-cell variability of chromatin structure.

Incidental

If you were to get your genome sequenced in the clinic in an attempt to diagnose a rare disease, should you be told whether you had a high risk for cardiomyopathy? The debate over whether incidental, or secondary, findings should be returned to patients has raged on. This

year, the American College of Medical Genetics has released guidelines for physicians recommending a set of genes in which incidental findings should be reported, irrespective of the age of the patient [3]. However, there is still work to be done, and efforts such as ClinVar and others will shape how genomic information is used in the clinic in the coming years.

Myriad

In a landmark decision for genetics law, the Supreme Court ruled that genomic DNA cannot be patented, but synthetically created cDNA is patentable (well summarized at <http://www.genomicslawreport.com/index.php/2013/06/18/myriad-finally-supreme-court-surprises-by-not-surprising/>). Most importantly, this decision marked the end of the debate on the claims of Myriad Genetics, which maintained patents on *BRCA1* and *BRCA2* that had been questioned extensively: if every human has these genes, how can a company own them? The court's decision effectively negated many of their claims and, now, institutions and companies are free to perform their own sequencing of *BRCA* genes. Of course, Myriad's knowledge base will keep them in business for some time, but they will face steep competition from lower-cost options. This decision was a month after Angelina Jolie published an op-ed in the New York Times about her decision to have a preventative double mastectomy based on family history and a Myriad test (<http://www.nytimes.com/2013/05/14/opinion/my-medical-choice.html>).

Large-scale transcriptomics

Move over 1000 Genomes Project! This year's large-scale RNA-Seq studies [4,5] have taken the top spot for consortium genomics buzzwords. These studies, and the GTEx project (<http://www.broadinstitute.org/gtex/>), have characterized the landscape of gene regulatory architecture, showing widespread variability in transcript structure and allelic expression, as well as variants associated with these features. Additionally, the increased

Correspondence: konradjkarczewski@gmail.com

¹Biomedical Informatics Training Program, Stanford University School of Medicine, Stanford, CA 94305, USA

²Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA

Big Data	ENCODE	Myriad	Incidental	Large-scale transcriptomics
Sequester	CRISPR	Clinically actionable	Analysis in the Cloud	NHGRI GWAS Diagram
CLIA-Certified	1000 Genomes	FREE SPACE	Exome Chip	Manhattan plot
HeLa	lncRNAs	FDA	Chromosome 21 Silencing	\$1000 genome
Re-identification	NHGRI sequencing cost graph	Nanopore	Missing heritability	Single-Cell

Figure 1 2013 bingo card. A 2013 genomics buzzword conference bingo card.

power of large datasets and meta-analyses has enabled investigation of the properties of *trans*-eQTLs and their influence on disease biology [5,6].

Big data

Our questions for 2013 and beyond have shifted from ‘how can we generate this amount of data?’ to ‘how can we analyze this amount of data?’ With terabytes of data and more, clever statistical and informatics methods are required to process and fully harness the data to arrive at the correct conclusions. Between an NIH RFA for ‘Big Data to Knowledge’, a number of companies (such as DNAnexus and Illumina) expanding the software efforts of genomics into the cloud, and Silicon Valley data-heavy company integration into health efforts (such as the Google-backed Calico), the trend toward scalable, intelligent data analysis mixed with high-performance computing is clear.

CRISPR

Clustered regularly interspaced short palindromic repeats (CRISPRs) have been developing as a molecular biology technique over the past decade, but this year has seen an explosion in their use in mammalian cells for genome editing. Their first use in human cells was published early in the year [7,8] and, since then, they have been used to correct *CFTR* gene function in cultured cells from cystic fibrosis patients, suggesting the potential for CRISPR use in gene correction therapies [9]. Just a few days ago, a number of studies were

published using CRISPRs for genome-scale knockout screens [10,11].

Nanopore

It’s unclear whether ‘nanopore’ should be noted as a word for 2012, 2013, or 2014. While the technology was announced previously, this year saw the first test runs of Oxford Nanopore’s new sequencing technology at ASHG and the launch of an early access program for MinION, their USB drive-sized sequencer.

HeLa

Of course, 2013 was not beyond some controversy. After the initial release of the HeLa cancer cell-line genome [12,13] without the knowledge of Henrietta Lacks’ family, the bioethics debates concerning the rights of - and respect for - research participants, balanced with the desire for public data availability, resurfaced. The genome was removed from public databases and, over the next few months, the Lacks family met with the NIH and medical geneticists to discuss options for data release, following which an agreement for controlled access to the data through dbGaP was struck [14].

FDA

After years of direct-to-consumer genetic testing companies selling kits and providing individuals with data on genetic risks for health, the FDA sent a warning letter to 23andMe, which instructed them to

cease marketing their genetic tests (<http://www.fda.gov/ICECI/EnforcementActions/WarningLetters/2013/ucm376296.htm>). The company, with more than half a million customers, complied and stopped offering health-related information to new customers, pending regulatory review. Interestingly, this announcement came only three days after the agency approved the first next-generation sequencer for marketing for clinical use [15].

Re-identification

Of additional interest for personal genomics companies and their customers, a study this year used publicly available personal genetic data, in the form of Y-STRs, together with genealogical records to re-identify a number of genomes [16]. These included some anonymous individuals from the HapMap project, as well as high-profile genomes such as that of Craig Venter.

lncRNAs

While long non-coding RNAs (lncRNAs) have been well-studied for years, a number of insights in 2013 have propelled them to the forefront. Thousands of novel lncRNAs were discovered using extensive *de novo* transcriptome assembly and found to be conserved and enriched for disease-associated variants [17]. Additionally, the lncRNA *Xist* was discovered to spread across the X chromosome using the three-dimensional structure of the genome [18].

Chromosome 21 silencing

Xist made another appearance in this year's highlights, when it was inserted into chromosome 21 in Down's syndrome pluripotent stem cells [19]. In doing so, the authors of the study were able to silence one copy of the chromosome, suggesting a potential use for *Xist* as a chromosome therapy for Down's syndrome.

Yes, 2013 was a great year for genomics, and the buzzwords it spawned will be around for at least a few years, but here's looking forward to a new set of advances and challenges in 2014!

Abbreviations

ASHG: American Society for Human Genetics; dbGaP: The database of Genotypes and Phenotypes; eQTL: Expression quantitative trait locus; FDA: Food and Drug Administration; GTEx: The Genotype-Tissue Expression project; lncRNAs: Long non-coding RNA; NIH: National Institutes of Health; RFA: Requests for applications; Y-STRs: Y-chromosome short tandem repeats.

Published: 24 December 2013

References

- Wu AR, Neff NF, Kalisky T, Dalerba P, Treutlein B, Rothenberg ME, Mburu FM, Mantalas GL, Sim S, Clarke MF, Quake SR: Quantitative assessment of single-cell RNA-sequencing methods. *Nat Methods* 2013. doi:10.1038/nmeth.2694.
- Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, Lue ED, Tanay A, Fraser P: Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 2013. doi:10.1038/nature12593.
- Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, McGuire AL, Nussbaum RL, O'Daniel JM, Ormond KE, Rehm HL, Watson MS, Williams MS, Biesecker LG: ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* 2013, 15:565–574.
- Lappalainen T, Sammeth M, Friedländer MR, Hoen PA T, Monlong J, Rivas MA, González-Porta M, Kurabayashi N, Griebel T, Ferreira PG, Barann M, Wieland T, Greger L, van Iterson M, Almlöf J, Ribeca P, Pulyakhina I, Esser D, Giger T, Tikhonov A, Sultan M, Bertier G, MacArthur DG, Lek M, Lizano E, Buermans HP, Padoleau I, Schwarzmayr T, Karlberg O, Ongen H, et al: Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* 2013. doi:10.1038/nature12531.
- Battle A, Mostafavi S, Zhu X, Potash JB, Weissman MM, McCormick C, Haudenschild CD, Beckman KB, Shi J, Mei R, Urban AE, Montgomery SB, Levinson DF, Koller D: Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. *Genome Res* 2013. doi:10.1101/gr.155192.113.
- Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, Christiansen MW, Fairfax BP, Schramm K, Powell JE, Zhernakova A, Zhernakova DV, Veldink JH, Van den Berg LH, Karjalainen J, Witthoff S, Uitterlinden AG, Hofman A, Rivadeneira F, Hoen PA T, Reinmaa E, Fischer K, Nelis M, Milani L, Melzer D, Ferrucci L, Singleton AB, Hernandez DG, Nalls MA, Homuth G, et al: Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013, 45:1238–1243.
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM: RNA-guided human genome engineering via Cas9. *Science* 2013, 339:823–826.
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F: Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013, 339:819–823.
- Schwank G, Koo BK, Sasselli V, Dekkers JF, Heo I, Demircan T, Sasaki N, Boymans S, Cuppen E, van der Ent CK, Nieuwennhuis EE, Beekman JM, Clevers H: Functional repair of *CfTR* by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* 2013, 13:653–658.
- Wang T, Wei JJ, Sabatini DM, Lander ES: Genetic screens in human cells using the CRISPR/Cas9 system. *Science* 2013. doi:10.1126/science.1246981.
- Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen T, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F: Genome-scale CRISPR-Cas9 knock-out screening in human cells. *Science* 2013. doi:10.1126/science.1247005.
- Adey A, Burton JN, Kitzman JO, Hiatt JB, Lewis AP, Martin BK, Qiu R, Lee C, Shendure J: The haplotype-resolved genome and epigenome of the aneuploid HeLa cancer cell line. *Nature* 2013, 500:207–211.
- Landry JJ, Pyl PT, Rausch T, Zichner T, Tekkedil MM, Stütz AM, Jauch A, Aiyar RS, Pau G, Delhomme N, Gagneur J, Korbel JO, Huber W, Steinmetz LM: The genomic and transcriptomic landscape of a HeLa cell line. *G3 (Bethesda)* 2013, 3:1213–1224.
- Hudson KL, Collins FS: Biospecimen policy: Family matters. *Nature* 2013, 500:141–142.
- Collins FS, Hamburg MA: First FDA authorization for next-generation sequencer. *N Engl J Med* 2013. doi:10.1056/NEJMmp1314561.
- Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y: Identifying personal genomes by surname inference. *Science* 2013, 339:321–324.
- Hangauer MJ, Vaughn IW, McManus MT: Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLOS Genet* 2013, 9:e1003569.
- Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, Kadri S, Xing J, Goren A, Lander ES, Plath K, Guttmann M: The *Xist* lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 2013, 341:1237973.
- Jiang J, Jing Y, Cost GJ, Chiang JC, Kolpa HJ, Cotton AM, Carone DM, Carone BR, Shivak DA, Guschin DY, Pearl JR, Rebar EJ, Byron M, Gregory PD, Brown CJ, Urnov FD, Hall LL, Lawrence JB: Translating dosage compensation to trisomy 21. *Nature* 2013, 500:296–300.

doi:10.1186/gb4148

Cite this article as: Karczewski: Progress in genomics according to bingo: 2013 edition. *Genome Biology* 2013 **14**:143.