

RESEARCH

Open Access

# Thoughts from SNP-SIG 2012: future challenges in the annotation of genetic variations

Yana Bromberg<sup>1\*</sup>, Emidio Capriotti<sup>2\*</sup>

From SNP-SIG 2012: Identification and annotation of SNPs in the context of structure, function, and disease Long Beach, CA, USA. 14 May 2012

## Overview

Advances in high-throughput sequencing, genotyping, and characterization of haplotype diversity are consistently generating vast amounts of genomic data. Single Nucleotide Polymorphisms (SNPs) are the most common type of genetic variation [1]. In the recent years the number of known SNPs has been increasing exponentially [2]; the last release of the NCBI's dbSNP database [3] contained more than 55 million human SNPs. SNPs are interesting as both markers of evolutionary history and in the context of their phenotypic manifestations (e.g. characteristic traits and diseases). However, due to their sheer number, detailed experimental annotations are impossible and computational inference is severely limited by the required resources. SNPs also present a challenge for visualization and storage.

In line with the increasing interest in the genetic variation analysis and annotation, on July 14th, 2012 [4] we organized the second SNP Special Interesting Group (SNP-SIG) meeting at the ISMB'12 in Long Beach, CA. This meeting attempted to summarize the field's research advances in the directions of "Annotation and prediction of structural/functional impacts of coding SNPs" and "SNPs and Personal Genomics: GWAS, populations and phylogenetic analysis". The discrepancy between the significant availability of the SNP data and the current lack of its interpretation requires the development of new computational annotation methods. The analysis of genetic variation is a key factor for the understanding of the genomic information. The SNP-SIG provides a forum necessary for the organization of a research network facilitating the exchange of ideas and for the establishment of new

collaborations to manage the complexity of the analysis of genetic variation. The one-day SIG attracted over 80 participants, with nine "bleeding-edge" research talks and five presentations from the leading scientists in the field.

The topics covered in the SIG presentations focused primarily on annotating the phenotypic traits associated with the specific SNPs - everything from annotating associated protein function changes [5-10] to screening pharmacogenomic variants [11] to finding driver mutations in cancer [12,13]. There was also some focus on improving method performance, e.g. a new way to efficiently analyze multi-GWAS data [14]. Finally, the issues of the quality of the current experimental variant annotations were also discussed. While the overall focus of the SIG was clearly on the side of improving variant annotation accuracy, the participants clearly recognized the need for faster methods capable of dealing with larger sets and noisy data. Although recognized as existing issues, the visualization of the SNP data and increasing data storage concerns, were not explicitly addressed. These will be a stressed focus in the future SNP-SIG meetings.

## Next meeting

We are currently preparing for the next edition of the SNP-SIG meeting to be held in the context of the ISMB 2013, Berlin, Germany. Further information about the SNP-SIG 2013 is available on our web site (<http://snpsig.biofold.org>).

## Acknowledgements

We are grateful to Sean Mooney (Buck Institute, Novato, CA) for serving as a guest editor for the manuscript submitted by the organizers. We thank Frank Schacherer and BIOBASE International for their financial support. We are very grateful for the support of the ISMB steering committee, especially Burkhard Rost (Technical University of Munich, Munich Germany), Michal Linial (Hebrew University, Jerusalem, Israel), and Hershel Safer (Tel Aviv University, Tel Aviv, Israel). We would like to extend special thanks for all help to the ISMB organizational committee and specifically Steven Leard and Jeremy Hennig. We acknowledge Sarah Headley and the editorial staff of BioMed

\* Correspondence: [yanab@rci.rutgers.edu](mailto:yanab@rci.rutgers.edu); [emidio@uab.edu](mailto:emidio@uab.edu)

<sup>1</sup>Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ, USA

<sup>2</sup>Division of Informatics, Department of Pathology, University of Alabama at Birmingham, Birmingham, AL, USA

Full list of author information is available at the end of the article

Central for their help with releasing this issue. We also thank the invited speakers: Russ Altman (Stanford University, Stanford, CA), Steven Brenner (University of California, Berkeley, CA), David Haussler (University of California, Santa Cruz, CA), Olivier Lichtarge (Baylor College of Medicine, Houston, TX) and Zemin Zhang (Genentech, San Francisco, CA). Finally, we are very grateful for the patience and help of our colleagues around the world who reviewed the presentation abstracts and the submitted manuscripts. The SIG meeting or this issue would not be possible without them:

Malay K. Basu (University of Alabama at Birmingham, Birmingham, AL), Lucia Conde (University of Alabama at Birmingham, Birmingham, AL), Xavier de la Cruz (Institut de Biologia Molecular de Barcelona, Barcelona, Spain), Andre Franke (Christian-Albrechts-University of Kiel, Kiel, Germany), Abel Gonzalez (Universitat Pompeu Fabra, Barcelona, Spain), Jae-Yoon Jung, (Harvard University, Boston, MA), Tony Kam-Thong (Max Plank Institute, Munich, Germany), Vidhya Krishnan (Genome Institute of Singapore, Singapore), Nuria Lopez (Universitat Pompeu Fabra, Barcelona, Spain), Thomas Madej (National Institute of Health, Bethesda, MD), Matthew Mort (Cardiff University, Cardiff, UK), Gaurav Pandey (Mount Sinai School of Medicine, NY), Marylyn D. Ritchie (Penn State University, University Park, PA), Janita Thusberg (Buck Institute, Novato, CA), Ali Torkmani (The Scripps Research Institute, La Jolla, CA), Iosif Vaisman (George Mason University, Manassas, VA), Andreas Zanzoni (Centre for Genomic Regulation, Barcelona, Spain), Verena Zuber, (Leipzig University, Leipzig, Germany) and the anonymous reviewers of the Capriotti *et al.* manuscripts.

This article has been published as part of *BMC Genomics* Volume 14 Supplement 3, 2013: SNP-SIG 2012: Identification and annotation of SNPs in the context of structure, function, and disease. The full contents of the supplement are available online at <http://www.biomedcentral.com/bmcgenomics/supplements/14/S3>

#### Author details

<sup>1</sup>Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ, USA. <sup>2</sup>Division of Informatics, Department of Pathology, University of Alabama at Birmingham, Birmingham, AL, USA.

Published: 28 May 2013

#### References

- 1000 Genomes Project Consortium: **A map of human genome variation from population-scale sequencing.** *Nature* 2010, **467**(7319):1061-1073.
- Capriotti E, Nehrt NL, Kann MG, Bromberg Y: **Bioinformatics for personal genome interpretation.** *Brief Bioinform* 2012, **13**(4):495-512.
- Shery ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K: **dbSNP: the NCBI database of genetic variation.** *Nucleic Acids Res* 2001, **29**(1):308-311.
- Bromberg Y, Capriotti E: **SNP-SIG Meeting 2011: identification and annotation of SNPs in the context of structure, function, and disease.** *BMC Genomics* 2012, **13**(Suppl 4):S1.
- Al-Numair N, Martin A: **The SAAP pipeline and database: tools to analyze the impact and predict the pathogenicity of mutations.** *BMC Genomics* 2013, **14**(S3):S4.
- Capriotti E, Altman R, Bromberg Y: **Collective judgment predicts disease-associated single nucleotide variants.** *BMC Genomics* 2013, **14**(Suppl 3):S2.
- Capriotti E, Calabrese R, Fariselli P, Martelli P, Altman R, Casadio R: **WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation.** *BMC Genomics* 2013, **14**(Suppl 3):S6.
- Carter H, Douville C, Stenson P, Cooper D, Karchin R: **Identifying Mendelian disease genes with the Variant Effect Scoring Tool.** *BMC Genomics* 2013, **14**(Suppl 3):S3.
- Peterson T, Park D, Kann M, *et al*: **A protein domain-centric approach for the comparative analysis of human and yeast phenotypically relevant mutations.** *BMC Genomics* 2013, **14**(Suppl 3):S5.
- Xie L, Ng C, Ali T, Valencia R, Ferreira B, Xue V, Tanweer M, Zhou D, Haddad G, Bourne P, Xie L: **Multiscale Modeling of the Causal Functional Roles of nsSNPs in a Genome-Wide Association Study: Application to Hypoxia.** *BMC Genomics* 2013, **14**(Suppl 3):S9.
- Daneshjou R, Tatonetti N, Karczewski K, Xue Y, Bourgeois S, Drozdka K, Burmester J, Tsunoda T, Nakamura Y, Kubo M, Tector M, Limdi N, Cavallari L, Perera M, Johnson J, Klein T, Altman R: **Pathway Analysis of Genome-Wide**

**Data Improves Warfarin Dose Prediction.** *BMC Genomics* 2013, **14**(Suppl 3):S11.

12. Reva B: **Revealing selection in cancer using the predicted functional impact of cancer mutations. Application to nomination of cancer drivers.** *BMC Genomics* 2013, **14**(Suppl 3):S8.
13. Gnad F, Baucom A, Mukhyala K, Manning G, Zhang Z: **Assessment of Computational Methods for Predicting the Effects of Missense Mutations in Human Cancers.** *BMC Genomics* 2013, **14**(Suppl 3):S7.
14. Goudey B, Rawlinson D, Wang Q, Shi F, Ferra H, Campbell R, Stern L, Inouye M, Ong C S, Kowalczyk A: **GWIS - Model-free, Fast and Exhaustive Search for Epistatic Interactions in Case-Control GWAS.** *BMC Genomics* 2013, **14**(Suppl 3):S10.

doi:10.1186/1471-2164-14-S3-S1

**Cite this article as:** Bromberg and Capriotti: Thoughts from SNP-SIG 2012: future challenges in the annotation of genetic variations. *BMC Genomics* 2013 **14**(Suppl 3):S1.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

