MOLECULAR ONCOLOGY XXX (2014) 1–15

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/molonc

Review

Standardized decision support in next generation sequencing reports of somatic cancer variants

Rodrigo Dienstmann[a,*], Fei Dong[a], Darrell Borger[b], Dora Dias-Santagata[a], Leif W. Ellisen[b], Long P. Le[a], A. John Iafrate[a]

[a]Massachusetts General Hospital and Harvard Medical School, Molecular Pathology Lab, USA
[b]Massachusetts General Hospital Cancer Center and Harvard Medical School, 55 Fruit St GRJ, Boston, MA 02114, USA

ARTICLE INFO

Article history:
Received 9 February 2014
Received in revised form
18 March 2014
Accepted 26 March 2014
Available online ■

Keywords:
Cancer
Genomics
Next-generation sequencing
Report
Variant annotation

ABSTRACT

Of hundreds to thousands of somatic mutations that exist in each cancer genome, a large number are unique and non-recurrent variants. Prioritizing genetic variants identified via next generation sequencing technologies remains a major challenge. Many such variants occur in tumor genes that have well-established biological and clinical relevance and are putative targets of molecular therapy, however, most variants are still of unknown significance. With large amounts of data being generated as high throughput sequencing assays enter the clinical realm, there is a growing need to better communicate relevant findings in a timely manner while remaining cognizant of the potential consequences of misuse or overinterpretation of genomic information. Herein we describe a systematic framework for variant annotation and prioritization, and we propose a structured molecular pathology report using standardized terminology in order to best inform oncology clinical practice. We hope that our experience developing a comprehensive knowledge database of emerging predictive markers matched to targeted therapies will help other institutions implement similar programs.

© 2014 Federation of European Biochemical Societies.
Published by Elsevier B.V. All rights reserved.

1. Introduction

With the advent of Next Generation Sequencing (NGS) or massively parallel sequencing technologies, we have the promise of a complete genetic description of patient tumors to optimally direct therapy. Clinical laboratories increasingly view large cancer genes panels as a cost-effective — and tissue-saving — alternative to running a series of multiple single-gene companion tests. Tremendous amounts of genomic data are being generated, with hundreds to thousands of variants observed in the coding regions of an individual's cancer genome, including somatic single

Abbreviations: CAP, College of American Pathologists; COSMIC, Catalog of Somatic Mutations in Cancer; EMR, electronic medical records; MAF, mutant allele fraction; NCCN, National Comprehensive Cancer Network; NGS, Next Generation Sequencing; SNP, single nucleotide polymorphisms; TCGA, The Cancer Genome Atlas.

* *Corresponding author.* Present address: Sage Bionetworks, Fred Hutchinson Cancer Research Center, 1100 Fairview N, Seattle, WA 98109, USA. Tel.: +1 206 667 4265.
E-mail address: rodrigo.dienstmann@sagebase.org (R. Dienstmann).

1574-7891/$ – see front matter © 2014 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.molonc.2014.03.021

Please cite this article in press as: Dienstmann, R., et al., Standardized decision support in next generation sequencing reports of somatic cancer variants, Molecular Oncology (2014), http://dx.doi.org/10.1016/j.molonc.2014.03.021

ARTICLE IN PRESS

nucleotide variants, insertions and deletions (indels), copy number alterations, rearrangements and germline susceptibility loci (Garraway and Lander, 2013; Vogelstein et al., 2013).

Molecular pathologists and cancer genomicists face a particular challenge in the reporting of the cancer genome. Manually annotating each single variant in terms of clinical significance in every possible tumor type is a daunting challenge. The large amount of data generated by high throughput assays and strain on the turnaround time drive the need for prioritization strategies for the identification and reporting of clinically significant genetic variants. Routine testing of full gene sequences as opposed to hotspots (Dias-Santagata et al., 2010) frequently identifies mutations of low frequency and unknown functional consequences, most of which are likely to be neutral or passenger alterations. On the other hand, some of these rare variants occur in cancer genes that have well-established clinical utility, driving tumorigenesis and tumor progression. The available scientific knowledge on these mutations should be presented in the report, so that physicians and patients can make evidence-based decisions in a responsible fashion. In addition, the availability of genetic results may provide a strong rationale for treatment with matched targeted agents in clinical trials, with the potential of directly benefitting the patient and accelerating the drug development process (Dienstmann et al., 2013). Consolidating so much information into a very discrete report that clearly identifies the clinical significance while preserving observations that can be further looked into by the clinician is not an easy undertaking. As physicians trained in fields other than genetics are playing a more central role in the ordering and reviewing of genetic test results, the importance of translating genomic data into informative reports is further increased.

In cancer genomics, performing NGS in the clinical laboratory is a multistep process that typically involves sample acquisition and quality control, DNA extraction, library preparation, sequencing and genomic data generation. As illustrated in Figure 1, the process continues with three dynamic pipelines for data analysis: (i) bioinformatics tools for variant identification; (ii) variant annotation and prioritization; and (iii) interpretation of clinical significance and reporting to clinicians (Van Allen et al., 2013; Watt et al., 2013). In this manuscript we discuss the challenges involved in the variant annotation and prioritization process and describe how genomic data can be translated into structured evidence-based reports. We hope that our experience developing the framework for clinical interpretation of somatic cancer variants and a comprehensive knowledge database of emerging predictive markers matched to targeted therapies will help other institutions implement similar programs.

2. Variant annotation and prioritization

The report generation process starts with standardized definitions by the molecular pathology laboratory of "reportable" and "not reportable" variants. Following variant identification using bioinformatics pipelines, a computational engine is needed in order to parse the variants and suppress those that are irrelevant, highlight the ones which need manual curation and identify pertinent "wild-types" in each tumor sample. As discussed below, several annotation and prioritization parameters are taken into consideration so as to provide a stronger estimation of the functional significance of unknown and novel mutations. Useful tools include sequencing metrics variables, external germline single nucleotide polymorphisms (SNPs) and cancer databases for comparison of variants across populations, as well as prediction models for defining damaging/deleterious or potentially driver mutations.

2.1. *Computational/bioinformatics tools*

When analyzing large cancer gene panels based on exome or whole genome sequencing, pairwise comparison with germline DNA plays a pivotal role. Subtracting the genetic variation of a non-cancerous "normal" genome from its cancerous counterpart allows the identification of the somatic mutations. Eliminating known harmless variants that are present in public (dbSNP) or in-house polymorphism databases and published studies such as the 1000 Genomes Project (Abecasis et al., 2012) and the Exome Sequencing Project (ESP6500)(Fu et al., 2013) is a very helpful strategy for reducing the candidate list of deleterious mutations.

Additionally, different bioinformatic adjustments can be used in order to improve variant detection and deal with library preparation or sequencing artifacts along with sample characteristics, including tumor purity and heterogeneity. Adequate coverage in target regions needs to be assured not only for variant detection but also to accurately define pertinent "wild-types" in specific tumor types. The next step involves prioritizing missense, nonsense or splice-site mutations over synonymous and intronic variants. In order to consider the variant as real and reportable, it is also advised to establish a minimum threshold of mutant allele fraction (MAF), the number of alternate reads at the genomic position divided by the total number of reads – coverage – at the same site. This threshold should take into consideration tumor cellularity and also clinical context, as rare resistant subclones in the treatment-refractory setting might be of relevance. Therefore, known gene variants previously clinically annotated are generally prioritized irrespective of MAF. If available, comparison with gene expression data (RNA sequencing) of the same sample can help determine the possible functional effects of a mutation, as variants not transcribed are less likely driver genomic events and therefore are not prioritized in the annotation process.

A fundamental aspect of the bioinformatics pipeline includes dealing with the mutational heterogeneity across the genome of a particular tumor, across different regions of the same tumor, across patients in a given tumor type, and across multiple tumor types. Sophisticated algorithms that incorporate DNA replication timing and transcriptional activity in the mutation call pipeline, such as MutSigCV, are able to identify true driver genomic events with higher accuracy (Lawrence et al., 2014, 2013). These models should be taken into consideration when implementing computational genomic methods in clinical molecular pathology labs.

Please cite this article in press as: Dienstmann, R., et al., Standardized decision support in next generation sequencing reports of somatic cancer variants, Molecular Oncology (2014), http://dx.doi.org/10.1016/j.molonc.2014.03.021

Download English Version:

https://daneshyari.com/en/article/10914630

Download Persian Version:

https://daneshyari.com/article/10914630

Daneshyari.com