

Paper of the month, April 2014

MutationTaster2 - mutation prediction for the deep sequencing age.

Schwarz JM, Cooper DN, Schuelke M, Seelow D.

Nature Methods. 2014 Apr;11(4):361-2

Die rasche Entwicklung der Hochdurchsatz-Sequenzierung hat in den letzten Jahren viele Bereiche der Medizin revolutioniert – nicht zuletzt auch das Gebiet der seltenen Erkrankungen. Alleine 2012 konnten etwa 200 neue Erkrankungen durch Ganz-Exom-Sequenzierungen identifiziert werden. Der nächste wichtige Schritt ist die breite Anwendung der Ganz-Genom-Sequenzierung. Dies ist sinnvoll, schließlich macht das Exom nur etwa 1% unseres Genoms aus. Obwohl sich die meisten bislang bekannten kausalen Mutationen im Exom befinden, können auch Mutationen in nicht-kodierenden Genom-Abschnitten Krankheiten verursachen. Diese sind bei wenigen Erkrankungen wie der Thalassämie ansatzweise untersucht, insgesamt jedoch kaum verstanden.

Ein wesentliches Problem der Hochdurchsatzsequenzierungen wird mit dem Wechsel von der Ganz-Exom- zur Ganz-Genom-Sequenzierung jedoch noch zunehmen: die schiere Menge an Daten, aus denen einzelne krankheitsverursachende Variablen herausgefiltert werden sollen. Bildlich gesprochen: bei der Suche nach der Nadel im Heuhaufen wird letzterer deutlich grösser.

Bereits bei Ganz-Exom-Sequenzierungen werden in der Regel ca. 20.000 Varianten identifiziert. Bei der Filterung und Priorisierung dieser Varianten kommen bioinformatische Tools zum Einsatz, welche auf Basis von publizierten Genomen Vorhersagen über die funktionelle Auswirkungen der Mutation treffen. Wichtige Beispiele sind *SIFT*, *PolyPhen* und *MutationTaster*. Schwarz *et al.* stellen nun in einer Arbeit aus dem MitoNET die nächste Generation des *MutationTaster* vor. Mit dessen Hilfe können neben klassischen Aminosäure-Substitutionen nun auch synonyme Basenveränderungen, Intron-Mutationen sowie kurze Insertionen und/oder Deletionen evaluiert werden.

Die Arbeit trägt daher wesentlich dazu bei, die neuen Sequenzier-Technologien praktisch nutzbar zu machen. Auf diese Weise sollten in den kommenden Jahren auch weiterhin viele seltene Erkrankungen identifiziert werden.

The recent rapid development of high-throughput sequencing has revolutionized many areas of medicine - not least the field of rare diseases. Alone in 2012, about 200 new rare diseases were identified by whole-exome sequencing. The next important step will be the widespread use of whole-genome sequencing. This is reasonable as the exome amounts to only about 1% of the genome. Although most of the known causative mutations are in the exons, there are also mutations in non-coding areas that can cause diseases. This has been shown for a few diseases, notably for thalassemia, but overall it is still poorly understood.

One problem of high-throughput sequencing will actually increase with the change from whole-exome to whole-genome sequencing: the sheer amount of data from which individual disease-causing variables need to be filtered out. Figuratively, when searching for a needle in a haystack, the haystack gets bigger and bigger.

Approximately 20,000 variants are usually identified using whole-exome sequencing. These variants are filtered and prioritized using bioinformatics tools which predict functional consequences of the particular mutation based on published genomes. Important examples are *SIFT*, *PolyPhen* and *MutationTaster*. The discussed paper by Schwarz *et al.* from the mitoNET presents the next generation of the *MutationTaster* software. With its help, also synonymous base changes, intron mutations and short insertions and / or deletions can be evaluated in addition to traditional amino acid substitutions.

This work helps to put the new sequencing technologies to practical use. As a result, the pace of discoveries of new rare diseases should continue at the same breathtaking pace.

Comment by

Dr. Tilmann Schober, Dr. von Hauner Children's Hospital, Ludwig Maximilians University Munich