Genome **Medicine**

RESEARCH HIGHLIGHT

Targeted next-generation sequencing for routine clinical screening of mutations

Jamie MJ Weaver* and Paul AW Edwards

Abstract

In many fields it is now desirable to sequence large panels of genes for mutation, to aid management of patients. The need for extensive sample preparation means that current approaches for assessing mutation status in the clinical setting are limited. A recent publication demonstrates a single-step, targeted, true single-molecule sequencing approach to assessing the mutational status of *BRCA1*. Fragmented DNA samples are loaded directly onto a flow cell and sequenced, thus detecting both small- and large-scale mutations with minimal sample preparation and high accuracy.

In a recent article in *Genome Research* [1], Thompson and colleagues present a development of a next-generation sequencing technology - the HeliScope sequencer - that enabled them to detect mutations in the human breast cancer 1 (*BRCA1*) gene, as a model of a clinical diagnostic protocol. Their technique is accurate and requires limited sample preparation before sequencing. It therefore holds great promise for overcoming the emerging problem in clinical genetics of target overload: the presence of a large number of gene mutations of clinical relevance in a single disease entity. In no field is this challenge more apparent than in the genomics of cancer.

The clinical application of next-generation sequencing

The past few years have seen an unprecedented deluge of data on the genes mutated in cancer and other diseases. So far, the sequences of over 50 individual cancer genomes have been published and this number is set to increase exponentially. The advent of next-generation technologies (such as the Roche 454 GS FLX+, llumina Hiseq 2000, Applied Biosystems SOLID and HeliScope single-molecule sequencer machines), which allow a human genome to be sequenced in a single week-long run, has led to a large shift in our understanding of the mutations that drive cancers [2]. For the clinician, mutations in cancer can have relevance for diagnosis, prognosis and treatment choice. However, it is now apparent that these clinically relevant mutations will be both numerous and found at low prevalence in individual cancer types. In particular, for the majority of cancers there will be no single 'magic bullet' targeted therapy. The breakpoint cluster region-Abelson murine leukemia homolog (BCR-ABL) fusion kinase, which is characteristic of chronic myeloid leukemia, is the exception not the rule when it comes to druggable cancer genes; it seems, instead, that multiple drug targets will be mutated at low frequency throughout common cancers [3]. This target overload is not restricted to the management of cancers and is also seen with monogenic disorders, for example, mental retardation and hereditary cancer predisposition [4].

It is therefore necessary to screen numerous genetic loci to decide on the best course of clinical management for an individual patient and this must be done in a rapid and cost-effective manner. The current technology for searching for mutations is to amplify a region of interest - in multiple fragments of a few hundred base pairs each in separate PCR reactions - then sequence the products by standard Sanger chain terminator sequencing. Although highly accurate for the detection of single-nucleotide variants and small insertions or deletions, this approach is expensive, labor intensive and unable to detect large-scale insertions or deletions. Furthermore, in tumor samples, mutations may be missed. This is because Sanger sequencing can reliably detect mutant alleles only when they are present in more than about 20% of the relevant DNA, and this will not be the case, for example, for a heterozygous mutation in a tumor contaminated with 60% normal DNA, a scenario that is not uncommon [5].
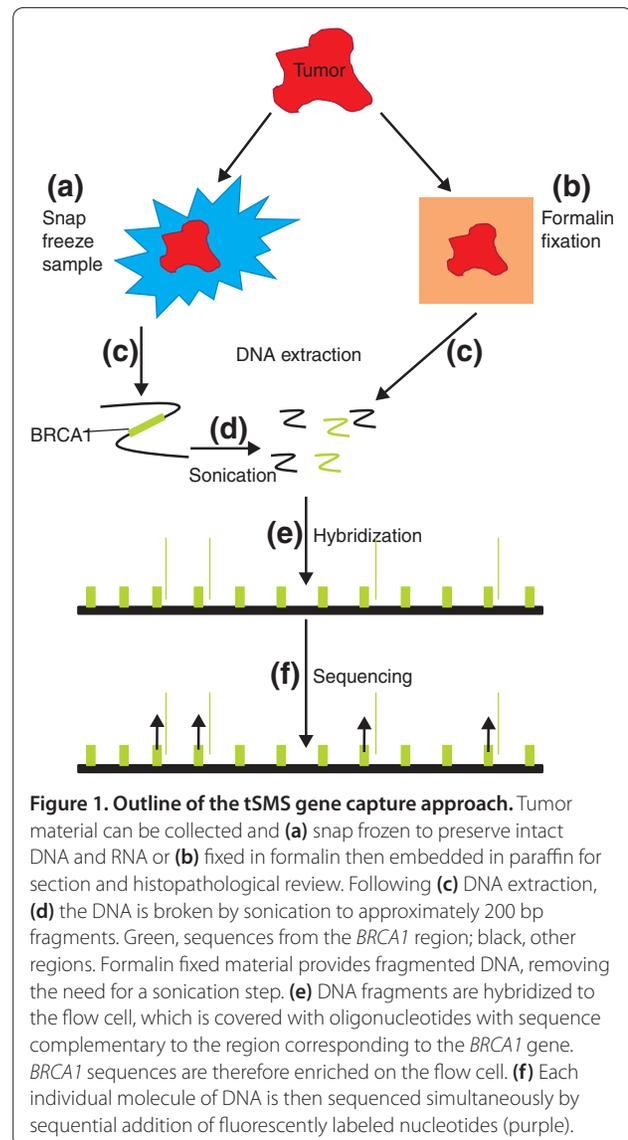
Targeted true single-molecule sequencing

Thompson and colleagues [1] have used a next-generation true single-molecule sequencing (tSMS) system, the HeliScope sequencer, to profile the mutational pattern of

*Correspondence: jmjw2@cam.ac.uk
Hutchison-MRC Research Centre and Department of Pathology, University of Cambridge, Hills Road, Cambridge, CB2 0XZ, UK

the human cancer gene *BRCA1*. Germline mutations in the genes *BRCA1* and *BRCA2* are associated with dramatically increased rates of breast and ovarian cancers and contribute to about 10% of all breast cancer cases [6]. In addition, poly-ADP ribose polymerase inhibitors, a recently developed family of pharmaceutical agents, seem to show selective toxicity for cancers with mutations in the *BRCA* genes. Cost-effective sequencing of these genes is therefore a highly desirable clinical tool for the management of breast cancer patients and those with strong family histories of the disease.

The technology used is a development of the HeliScope tSMS platform [1]. In the standard HeliScope protocol, DNA is fragmented and poly(A) adaptors are added to the ends of fragments. They are then captured on a glass slide coated with covalently bound poly(dT) oligonucleotides, and sequenced [7]. Thompson *et al.* [1] were able to dispense with the initial sample preparation steps (addition of poly(A) adaptors and 3' blocking - performed to prevent extension of the 3' end of bound poly(A)-tagged DNA molecules) and instead directly captured only the fragments of DNA belonging to the *BRCA1* locus, using oligonucleotides that match sequences in the *BRCA1* region (Figure 1). This targeted capture approach was highly successful. Approximately 20% of the sequenced reads mapped to *BRCA1*, which equates to about 100,000-fold enrichment of the target sequence. As the length of sequence that can be read in this system is limited, oligonucleotides were designed at 20-30 base pair intervals throughout the coding sequence of the gene, to ensure complete coverage. Importantly, the authors [1] were able to obtain sequencing results from as little as 100 ng of input material - thus showing that the technology could be used with samples collected as diagnostic biopsies.

This approach [1] has several advantages over the more traditional Sanger sequencing commonly used for the detection of mutations. Firstly, there is little required in the way of sample preparation - only sonication of the DNA, and in the case of archival formalin-fixed material it is possible that even this step could be dispensed with - therefore reducing cost, turn-around time and the risk of errors in sample handling. Secondly, unlike Sanger sequencing, the HeliScope system allows the detection of large deletions, by determining the number of fragments of DNA that are present from a specific location; a decreased number of fragments from a region suggests a loss of genetic material. Finally, the HeliScope system directly sequences individual molecules rather than - as with Sanger sequencing - examining the average sequence of many millions of DNA molecules. This should provide increased sensitivity for the detection of low prevalence mutations, an essential feature in the sequencing of a heterogeneous cancer sample.

Figure 1. Outline of the tSMS gene capture approach. Tumor material can be collected and **(a)** snap frozen to preserve intact DNA and RNA or **(b)** fixed in formalin then embedded in paraffin for section and histopathological review. Following **(c)** DNA extraction, **(d)** the DNA is broken by sonication to approximately 200 bp fragments. Green, sequences from the *BRCA1* region; black, other regions. Formalin fixed material provides fragmented DNA, removing the need for a sonication step. **(e)** DNA fragments are hybridized to the flow cell, which is covered with oligonucleotides with sequence complementary to the region corresponding to the *BRCA1* gene. *BRCA1* sequences are therefore enriched on the flow cell. **(f)** Each individual molecule of DNA is then sequenced simultaneously by sequential addition of fluorescently labeled nucleotides (purple).

The present technology [1] is an advance over recently presented techniques for analyzing the mutational status of target genes using next-generation sequencing instruments [8,9]. With the advent of more competitively priced next-generation sequencing machines, which can provide sequence data in a matter of hours, not days (such as Life Technologies' Ion-torrent and Illumina's Miseq), it is feasible that such approaches could be used routinely in the clinical setting. However, the instruments involved still require the selective amplification of target DNA and the relatively complex preparation of this material for sequencing. Thus, although they take advantage of the exceptionally cheap per-base cost of next-generation sequencing, competing techniques are still limited by its flaws: high complexity and slow turnaround time. Genotyping by mass spectrometry has also been

applied to the resequencing of relevant mutations [10]. This approach, although sensitive, is limited, in that it is capable of detecting only known mutations and will therefore miss novel gene lesions.

The future of gene sequencing in the clinic

For the foreseeable future, the high cost and complexity of data analysis will limit the application of whole-genome sequencing for the detection of mutations in a clinical setting. Targeted resequencing of areas of interest will therefore remain key to determining mutational status. The method published by Thompson *et al.* [1] is a stride forward in putting this into practice. Although currently only a single gene is screened, there is clearly scope for the creation of multi-gene capture arrays, allowing large numbers of loci to be analyzed rapidly and cost-effectively with low DNA input requirements. In its simplicity, this approach provides an opportunity to truly begin integrating the vast quantity of genomic data generated in this next-generation era with clinical practice.

Abbreviations
BRCA, breast cancer gene; tSMS, true single-molecule sequencing.

Competing interests
The authors declare that they have no competing interests.

References
1. Thompson JF, Reifenberger JG, Giladi E, Kerouac K, Gill J, Hansen E, Kahvejian A, Kapranov P, Knope T, Lipson D, Steinmann KE, Milos PM: **Single-step capture and sequencing of natural DNA for detection of BRCA1 mutations.** *Genome Res* 2011, doi: 10.1101/gr.122192.111.
2. Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, Varela I, Lin ML, Ordóñez GR, Bignell GR, Ye K, Alipaz J, Bauer MJ, Beare D, Butler A, Carter RJ, Chen L, Cox AJ, Edkins S, Kokko-Gonzales PI: **A comprehensive catalogue of somatic mutations from a human cancer genome.** *Nature* 2010, **463**:191-196.
3. Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, Ou SH, Dezube BJ, Jänne PA, Costa DB, Varella-Garcia M, Kim WH, Lynch TJ, Fidias P, Stubbs H, Engelman JA, Sequist LV, Tan W, Gandhi L, Mino-Kenudson M, Wei GC, Shreeve SM, Ratain MJ, Settleman J, Christensen JG, Haber DA, Wilner K, Salgia R, Shapiro GI, Clark JW, Iafrate AJ: **Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer.** *N Engl J Med* 2010, **363**:1693-1703.
4. Vissers LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wieskamp N, del Rosario M, van Bon BW, Hoischen A, de Vries BB, Brunner HG, Veltman JA: **A de novo paradigm for mental retardation.** *Nat Genet* 2010, **42**:1109-1112.
5. Tsiatis AC, Norris-Kirby A, Rich RG, Hafez MJ, Gocke CD, Eshleman JR, Murphy KM: **Comparison of Sanger sequencing, pyrosequencing, and melting curve analysis for the detection of KRAS mutations: diagnostic and clinical implications.** *J Mol Diagn* 2010, **12**:425-432.
6. King MC, Marks JH, Mandell JB: **Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2.** *Science* 2003, **302**:643-646.
7. Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, Causey M, Colonell J, Dimeo J, Efcavitch JW, Giladi E, Gill J, Healy J, Jarosz M, Lapen D, Moulton K, Quake SR, Steinmann K, Thayer E, Tyurina A, Ward R, Weiss H, Xie Z: **Single-molecule DNA sequencing of a viral genome.** *Science* 2008, **320**:106-109.
8. Morgan JE, Carr IM, Sheridan E, Chu CE, Hayward B, Camm N, Lindsay HA, Mattocks CJ, Markham AF, Bonthron DT, Taylor GR: **Genetic diagnosis of familial breast cancer using clonal sequencing.** *Hum Mutat* 2010, **31**:484-491.
9. Walsh T, Lee MK, Casadei S, Thornton AM, Stray SM, Pennil C, Nord AS, Mandell JB, Swisher EM, King MC: **Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing.** *Proc Natl Acad Sci USA* 2010, **107**:12629-33.
10. MacConaill LE, Campbell CD, Kehoe SM, Bass AJ, Hatton C, Niu L, Davis M, Yao K, Hanna M, Mondal C, Luongo L, Emery CM, Baker AC, Philips J, Goff DJ, Fiorentino M, Rubin MA, Polyak K, Chan J, Wang Y, Fletcher JA, Santagata S, Corso G, Roviello F, Shivdasani R, Kieran MW, Ligon KL, Stiles CD, Hahn WC, *et al*: **Profiling critical cancer gene mutations in clinical tumor samples.** *PLoS One* 2009, **4**:e7887.