

Hochdurchsatz-Sequenzierung von gefrorenem und in Paraffin eingebettetem Tumor- und Normalgewebe

Die Sequenzierung des humanen Genoms repräsentiert einen der größten Fortschritte der modernen Biologie. Ursprünglich dieses Vorhabens gehen zurück auf die Entdeckung der DNA-Struktur 1953 durch James Watson, Francis Crick und Rosalind Franklin. Etwa 20 Jahre später wurden die ersten Sequenzieretechnologien entwickelt: Die Sanger-Methode 1975 und die Maxam-Gilbert-Methode 1977. Einen weiteren Durchbruch auf dem Wege zu den heutigen Sequenzieretechnologien brachte die Möglichkeit, DNA mit Mitteln der PCR („polymerase chain reaction“) unbegrenzt zu vervielfältigen. Diese Methode entwickelte Kary Mullis 1983. Von diesem Zeitpunkt an vergrößerten sich die Sequenzierkapazitäten immer mehr und gipfelten in der Entwicklung von Sequenzieretechnologien der zweiten Generation.

Fortschritte auf dem Gebiet der automatischen Sequenzierung läuteten 1990 den Start für das humane Genomprojekt ein, in dessen Rahmen das gesamte menschliche Genom entschlüsselt wurde. In einer internationalen Anstrengung mit Beteiligung von über 1000 Wissenschaftlern wurde die fertige Sequenz 2001 veröffentlicht [1]. Heutzutage ist es möglich, diesen Aufwand, der über mehrere Jahre unter Einsatz enormer Ressourcen betrieben werden musste, auf wenige Tage bis Wochen zu reduzieren und Sequenzdaten für die Diagnostik bereitzustellen.

Sequenzieretechnologien der zweiten Generation

In der Tat, die Entwicklung von Sequenzieretechnologien der zweiten Generation (454 von Roche, „Genome Analyzer“ von Illumina und SOLiD von Applied Biosystems) hat einen enormen Fortschritt in der Analyse von Genomen erbracht [2, 3]. Der eigentliche Durchbruch kam durch die Entwicklung lokalisationspezifischer Sequenzierungen. Das bedeutet, dass durch Immobilisierung der zu sequenzierenden DNA-Fragmente und anschließender lokusspezifischer Datenanalyse die parallele Sequenzierung von Millionen von DNA-Fragmenten ermöglicht worden ist. Dabei emittieren Fluoreszenzmarker entweder durch eine Ligationsreaktion (SOLiD), eine Polymeraseaktivität oder eine gekoppelte Luziferaseaktivität während der Sequenzierung Licht, das anschließend in die DNA-Sequenz umgesetzt werden kann. Damit werden die Sequenzierungen in einer digitalen Art und Weise durchgeführt und können durch bioinformatische Analysealgorithmen prozessiert werden [4, 5, 6]. Diese parallele Sequenzierung ist im Besonderen für heterogenes Gewebe geeignet, da bereits geringe Mengen an Sequenzabweichungen detektiert werden können, die insbesondere in der Tumorbiologie von großer Bedeutung sein können [7, 8].

Dieser technische Fortschritt berechtigt die großen Erwartungen, die auf die Sequenzanalysen mit Technologien der

zweiten Generation gesetzt werden. Es formieren sich große internationale Konsortien wie z. B. das ICGC („International Cancer Genome Consortium“) oder TCGA („The Cancer Genome Atlas“), die sich zum Ziel gesetzt haben, durch Sequenzierung einer großen Anzahl von Tumoren Veränderungen zu detektieren, die einer Vielzahl an Tumoren zugrunde liegen.

Nachteile der FFPE-Gewebeproben

Bisherige Untersuchungen, auch solche im Rahmen dieser großen Sequenzierprojekte, basieren auf Gewebeproben, die entweder aus frischem Gewebe entnommen oder sofort nach der Entnahme in flüssigem Stickstoff eingefroren worden sind. Diese Vorgehensweise stellt einen entscheidenden Engpass dar: Einerseits haben Chirurgen bei der Tumorentnahme nicht immer die Zeit und die Gelegenheit Gewebeproben sofort einzufrieren, andererseits wurden von vielen interessanten Proben in der Vergangenheit keine Gefrierproben asserviert.

In den pathologischen Abteilungen aller Krankenhäuser werden Gewebeproben routinemäßig zunächst in Formalin fixiert und anschließend in Paraffin (FFPE) eingebettet. Diese Form der Konservierung ist auch eine gängige Methode, wenn es darum geht, große Gewebebanken anzulegen, in denen Proben über viele Jahre hinweg gesammelt werden.

M. Kerick · B. Timmermann · M.-R. Schweiger

Hochdurchsatz-Sequenzierung von gefrorenem und in Paraffin eingebettetem Tumor- und Normalgewebe

Zusammenfassung

Bisher sind Hochdurchsatz-Sequenzierungen von Tumorproben ausschließlich an DNA aus eingefrorenem Gewebe durchgeführt worden. Die Gefrierpräparation von Geweben ist jedoch relativ umständlich, und die meisten Gewebekbanken greifen deshalb auf eine Methode der Formalinfixierung und Paraffineinbettung (FFPE) zurück. Insbesondere in den Instituten für Pathologie wird Gewebe auf diese Weise konserviert und gelagert. Ein Nachteil dieser Art der Gewebearchivierung ist es allerdings, dass viele moderne molekularbiologische Techniken an solchen Proben nicht durchgeführt werden können, da in ihnen die DNA fragmentiert und teilweise quervernetzt mit Proteinen vorliegt. In den letzten Jahren wurde eine neue Technik, die Sequenzier-technologie der zweiten Generation, entwickelt. Mit dieser Technologie ist es möglich, gesamte menschliche Genome innerhalb von Tagen zu sequenzieren. Wir konnten zeigen, dass mit diesen Methoden auch FFPE-präpariertes Gewebe verwendet werden kann. Dabei können sowohl Gesamtgenom-Sequenzierungen als auch eine gezielte Anreiche-

rung von genomischen Subregionen, gefolgt von Sequenzierungen, mit ähnlich guter Qualität wie von Gefriergewebe durchgeführt werden. Darüber hinaus können die neuen Sequenzier-technologien nicht nur für die Detektion von Mutationen, sondern gleichzeitig auch für die Detektion von Kopienzahlveränderungen eingesetzt werden.

Unser Nachweis, dass auch FFPE-archiviertes Gewebe für diese modernen genomweiten Sequenzier-technologien verwendet werden kann, eröffnet neue Wege, Material bereits durchgeführter Studien zu reanalysieren und ermöglicht gleichzeitig den Zugriff auf Proben großer Gewebekbanken und anderer Institutionen. Insbesondere Sammlungen von Instituten für Pathologie bzw. Museen sind jetzt zugänglich für DNA-Sequenzanalysen.

Schlüsselwörter

Funktionelle Genomanalyse · Sequenzier-technologien · FFPE-Gewebe · Mutationsdetektion · Kopienzahldetektion

High-throughput sequencing of frozen and paraffin-embedded tumor and normal tissue

Abstract

Until now high-throughput sequencing of tumor samples relied on DNA isolated from fresh frozen tissues, the preparation of which, however, is relatively laborious. The use of preserved material, i.e. from tissue banks, could help to avoid this limitation and would enable the reanalysis of diverse clinical trials. So far we have shown that formalin-fixed paraffin-embedded (FFPE) tissue samples can be used for genomic re-sequencing processes. FFPE samples are amply available from surgical tumor resections and histopathological diagnosis, and comprise tissue from precursor lesions, primary tumors, and lymphogenic and/or hematogenic metastases. To generate models which predict the response to therapy, FFPE tissue also has the advantage that it is available from a va-

riety of clinical trials. Second generation sequencing techniques are not only applicable to snap frozen and FFPE tissues for whole genome analyses but also for targeted resequencing approaches. In addition, the detection of copy number variations and mutations in FFPE tissues can be obtained within one sequencing run.

The possibility of using genome-wide technologies irrespective of the mode of storage facilitates the retrieval of useful material and is a prerequisite for subsequent computational modelling approaches.

Keywords

Functional genome analysis · Sequencing technologies · FFPE tissue · Mutation detection · Copy number detection

Der Nachteil dieser Technik ist es jedoch, dass viele molekularbiologische Techniken an derartigen Proben nicht durchgeführt werden können, da die Verwendung von Formaldehyd zu einer Quervernetzung von DNA mit an sie gebundenen Proteinen führt und damit die DNA für viele modifizierende Enzyme nicht mehr zugänglich ist [9]. Darüber hinaus liegt DNA aus FFPE-Materialien oft fragmentiert vor, sodass Techniken wie z. B. PCR nicht durchgeführt werden können. Insgesamt ist die Qualität wie auch die Quantität derartiger DNA stark beeinträchtigt.

Material und Methoden

Im Hinblick auf die zukünftige Tumorgenetik erhebt sich daher die wichtige Frage, ob DNA aus FFPE-Geweben für Sequenzierungen mit Techniken der zweiten Generation verwendet werden kann.

Um diese Frage beantworten zu können, haben wir systematische Untersuchungen durchgeführt. Unterschiedliche Asservierungsbedingungen wie auch verschiedene DNA-Extraktions-Methoden wurden eingesetzt, um die Tauglichkeit der DNA für ihre Verwendung in den modernen Sequenzier-technologien zu prüfen. Unter anderem haben wir in unseren Versuchen Ischämiezeiten zwischen 20 bis über 360 Minuten, Fixationszeiten bis zu 72 Stunden sowie Lagerungszeiten bis zu 18 Jahren vorgegeben.

Für die Experimente haben wir Gewebeproben verwendet, deren einer Teil in Stickstoff eingefroren und der andere parallel als FFPE-Material präpariert worden ist. Der Gehalt an Tumorgewebe wurde unter dem Mikroskop bestimmt und auf mindestens 70% angesetzt. Nach der Präparation der Sequenzierbibliotheken wurden die Proben auf einem „Genome Analyzer“ der Firma Illumina sequenziert und jeweils auf Veränderungen der Kopienzahl und der Nukleotidaustausche (Polymorphismen und Mutationen) hin untersucht.

Ergebnisse

Ein bedeutender Vorteil der modernen Sequenzier-technologien besteht darin, dass sowohl Betrachtungen der Kopienzahl als auch Mutationsuntersuchungen

in ein und demselben Experiment durchgeführt werden können. Konventionelle Methoden der Kopienzahldetektion umfassten bisher Array-basierte Techniken wie CGH- (komparative genomische Hybridisierung-) Arrays oder FISH- (Fluoreszenz-in-situ-Hybridisierung-) Untersuchungen, die sich auf einzelne kleine DNA-Regionen beschränken. Für Mutationsuntersuchungen wurden meist Sanger-Sequenzierungen benutzt [10].

Umfangreiche bioinformatische Analysen unserer Ergebnisse ergaben, dass wir in beiden Arten der Gewebepreparation ähnliche Veränderungen detektieren können. So fanden sich die gleichen Kopienzahlveränderungen, und auch in der Anzahl und Art der Nukleotidaustausche traten keine Signifikanzunterschiede auf.

Interessanterweise konnten wir zeigen, dass die Sequenzierung zu einer dem 244 K Array von Agilent vergleichbaren Auflösung führt und somit bei vergleichbaren Kosten eingesetzt werden kann [11]. Die Konstanz der Abdeckung der genomischen DNA ist in beiden Präparations-techniken vergleichbar, wobei wir eine etwas höhere Sequenzierleistung bei FFPE-Gewebe benötigen, um auf die gleiche Abdeckung der DNA zu kommen.

In Bezug auf die Detektion von Veränderungen auf dem Nukleotidniveau fanden wir bei 8-facher Sequenzierung eine Übereinstimmung von mehr als 90% zwischen den beiden Gewebetypen. Allerdings muss man hier auch berücksichtigen, dass die Sequenzierungen zu einem sehr frühen Entwicklungsstand der Technologien und mit einer relativ geringen Abdeckung der DNA durchgeführt worden sind. Es finden sich etwas mehr Veränderungen in FFPE-Gewebe als in Gefriergewebe mit einem relativ höheren Anteil an Mutationen (d. h. noch nicht in der Datenbank bekannten Polymorphismen), ein Umstand, der vermutlich durch DNA-Schäden in FFPE-Gewebe verursacht worden ist. Eine Vergrößerung der Sequenzierleistung führt übrigens auch zu einer noch besseren Übereinstimmung der Veränderungen zwischen den beiden Gewebetypen, und wir sehen einen negativen Einfluss von übermäßig langen Fixationszeiten.

Fazit

Generell lässt sich sagen, dass FFPE-Gewebe für die Sequenziertechnologien der zweiten Generation verwendet werden können, allerdings sollte, wenn möglich, dem Gefriergewebe der Vorzug gegeben werden. Auf jeden Fall zeigen unsere Untersuchungen, dass für besondere Fragestellungen auf große Gewebekontingenzen und deren FFPE-Proben zurückgegriffen werden kann. Damit können z. B. bereits durchgeführte klinische Medikamentenstudien neu analysiert werden, mit dem Ziel, herauszufinden, warum eine Patientensubgruppe auf eine bestimmte Behandlung nicht angesprochen hat, während bei der anderen gute Erfolge erzielt werden konnten [12, 13, 14, 15].

Wenn auch von größtem Interesse, so ist die Verwendung dieser Sequenziertechnologien keinesfalls auf die Tumordiagnostik beschränkt. Viele genetische Erkrankungen, seien sie nun monogenetisch oder auch komplex genetisch bedingt, profitieren bereits von den Fortschritten in der DNA-Sequenzanalyse. So, wie vor wenigen Jahren die Array-CGH-Diagnostik in die Routinediagnostik eingeführt worden ist, wird in naher Zukunft diese zweite Generation an Sequenziertechnologien Einzug in der Klinik halten. Erste Ringversuche zur Ermittlung der Vergleichbarkeit der Ergebnisse an unterschiedlichen Standorten sind bereits im Gange und erste Daten zeugen von sehr guter Reproduzierbarkeit. Es bleibt somit als nächster und wichtigster Schritt, ethische Bedenken im Hinblick auf genetische Diagnostik zu adressieren und Maßnahmen zu ergreifen, die den individuellen Schutz der Patienten gewährleisten.

Korrespondenzadresse

Dr. med. Dr. rer. nat. M.-R. Schweiger
Cancer Genomics Group,
Department of Vertebrate Genomics,
Max-Planck-Institut für molekulare Genetik
Inhest. 63–73, 14195 Berlin
mschweig@molgen.mpg.de

Interessenkonflikt. Der korrespondierende Autor gibt an, dass kein Interessenkonflikt besteht.

Literatur

1. Hattori M, Fujiyama A, Taylor TD et al (2000) The DNA sequence of human chromosome 21. *Nature* 405:311–319
2. Metzker ML (o J) Sequencing technologies – the next generation. *Nat Rev Genet* 11:31–46
3. Shendure J, Ji H (2008) Next-generation DNA sequencing. *Nat Biotechnol* 26:1135–1145
4. Margulies M, Egholm M, Altman WE et al (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–380
5. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25
6. Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* 18:1851–1858
7. Bell DW (2010) Our changing view of the genomic landscape of cancer. *J Pathol* 220:231–243
8. Thomas RK, Baker AC, Debiase RM et al (2007) High-throughput oncogene mutation profiling in human cancer. *Nat Genet* 39:347–351
9. Gilbert MT, Haselkorn T, Bunce M et al (2007) The isolation of nucleic acids from fixed, paraffin-embedded tissues-which methods are useful when? *PLoS One* 2:e537
10. Chin L, Gray JW (2008) Translating insights from the cancer genome into clinical practice. *Nature* 452:553–563
11. Solinas-Toldo S, Lampel S, Stilgenbauer S et al (1997) Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 20:399–407
12. Collins I, Workman P (2006) New approaches to molecular cancer therapeutics. *Nat Chem Biol* 2:689–700
13. Mardis ER, Wilson RK (2009) Cancer genome sequencing: a review. *Hum Mol Genet* 18:R163–R168
14. Le Tourneau C, Vidal L, Siu LL (2008) Progress and challenges in the identification of biomarkers for EGFR and VEGFR targeting anticancer agents. *Drug Resist Updat* 11:99–109
15. Eberhard DA, Johnson BE, Amler LC et al (2005) Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. *J Clin Oncol* 23:5900–5909