

Einführung in die Grundlagen der Hochdurchsatzsequenzierung

Das menschliche diploide Genom besteht aus etwa 6 Mrd. Nukleotiden. Schon ein einziger Fehler hierin kann zu einer schwerwiegenden Krankheit führen. Klassische Methoden zur Detektion solcher Mutationen waren bisher die radioaktive Sequenzierung nach Maxam und Gilbert und die kapilläre Sanger-Sequenzierung. Während die radioaktive Sequenzierung schon lange nicht mehr verwendet wird, hat sich die kapilläre Sanger-Sequenzierung zum Goldstandard zur Bestimmung der Sequenzabfolge eines DNA-Moleküls entwickelt. Im Jahre 2006 wurde dann allerdings die erste Hochdurchsatzsequenzierung beschrieben. Diese neueste technologische Entwicklung macht es nun möglich, alle Nukleotide des menschlichen Genoms (oder zumindest seines proteinkodierenden Teils, dem sog. Exom) lesen zu können. Inzwischen ist Next Generation Sequencing (NGS) in aller Munde. In diesem Übersichtsartikel werden die Grundlagen der Hochdurchsatzsequenzierung, die zugrunde liegenden Techniken verschiedener Maschinen mit einigen Vor- und Nachteilen sowie einige Anwendungsmöglichkeiten beschrieben.

Sanger-Sequenzierung

Die Sequenzierung nach Frederick Sanger (1918–2013) wird auch als Didesoxymethode nach Sanger oder Kettenabbruchsynthese bezeichnet [10]. Hierbei werden zunächst Polymerase-Kettenreaktions(PCR)-Produkte der zu untersuchenden Region (z. B. eines Exons) hergestellt, um eine Anreicherung der Sequenz dieser Region zu erzielen. In einer folgenden Sequenzierreaktion werden dann zunächst die DNA-Doppelstränge per Hitze-denaturierung voneinander getrennt, sodass die resultierenden Einzelstränge für

die Amplifikation zur Verfügung stehen. Für die nachfolgende Reaktion werden allerdings nun außer Polymerase, Sequenzierprimer und normalen Trinukleotidphosphaten (dNTPs) auch Didesoxynukleotidtriphosphate (ddNTPs) zugefügt, wobei jedes der 4 ddNTPs an einen eigenen Fluoreszenzfarbstoff gekoppelt ist. Die ddNTPs besitzen keine 3'-Hydroxylgruppe, sodass nach Einbau eines solchen ddNTPs eine Verlängerung durch die DNA-Polymerase nicht mehr möglich ist. Auf diese Weise erfolgt eine lineare Amplifikation der DNA, die in amplifizierten DNA-Abschnitten unterschiedlicher Länge resultiert. Statistisch gesehen sollte hierbei nach jedem einzelnen eingebauten Nukleotid ein Abbruch stattfinden, sodass alle möglichen unterschiedlichen Fragmentlängen entstehen. Diese unterschiedlich langen Sequenzierprodukte werden anschließend in einem Sequenziergerät kapillarelektrophoretisch nach ihrer Größe aufgetrennt. Mithilfe eines eingebauten Lasers werden die Fluoreszenzfarbstoffe angeregt und die Fluoreszenz mit einer Kamera detektiert. Die resultierenden Farbsignale werden anschließend als Elektropherogramm wiedergegeben, das entsprechend ausgewertet werden kann, indem die Basenabfolge des DNA-Abschnitts bestimmt wird.

Man kann sich vorstellen, dass die elektrophoretische Auftrennung eines amplifizierten Sequenzierprodukts die Kapazität der Sanger-Sequenzierung bestimmt. Pro Kapillare wird i. d. R. ein einzelnes PCR-Produkt aufgetrennt, wobei die resultierende Sequenz je nach Kapillare etwa 300–1000 bp lang ist. Obwohl es heute Sanger-Sequenziergeräte mit bis zu 96 Kapillaren gibt, bleibt der Durchsatz daher relativ gering. Weitere Nachteile sind, dass i. d. R. für jedes einzelne zu untersuchen-

de DNA-Fragment eine separate PCR benötigt wird, man keine Aussage zur Allelverteilung machen kann und eine quantitative Aussage nicht möglich ist. Der Vorteil der Sanger-Sequenzierung ist die hohe Qualität, die pro Amplikon erreicht werden kann.

Next Generation Sequencing

Eigentlich ist der Begriff NGS sehr verwirrend, denn es handelt sich hierbei nicht um eine einzelne neue Technik, sondern eher um viele verschiedene neue Methoden. Es ist also wichtig zu wissen, dass NGS nur ein zusammenfassender Begriff für viele Techniken ist und nicht selber eine Technik darstellt. Insgesamt gibt es mehr als 20 verschiedene Hochdurchsatzsequenziergeräte, die entweder z. Z. entwickelt werden oder bereits als kommerzielle Plattformen zur Verfügung stehen. Die bekanntesten und gebräuchlichsten werden in diesem Artikel vorgestellt. Obwohl alle diese Maschinen ihre eigene Technik besitzen, folgen doch die meisten dieser Sequenziergeräte einem ähnlichen Prinzip aus 3 aufeinander folgenden Schritten: Library-Präparation, Amplifikation und Sequenzierung.

Library-Präparation

Zunächst muss man die zu untersuchende DNA auf die Sequenzierung vorbereiten, ein Vorgang der üblicherweise als Library-Präparation bezeichnet wird. Hierbei wird die DNA z. B. enzymatisch oder mechanisch in Fragmente bestimmter Länge geschnitten. Welche Fragmentlänge gewählt wird, hängt dabei davon ab, welche Sequenziermaschine benutzt werden soll. Nach der Fragmentierung werden meist die DNA-Enden repariert, um

anschließend Adaptoren (doppelsträngige NGS-Primer) an beide Seiten des DNA-Fragments anhängen zu können (Adaptterligation). Die fertigen DNA-Abschnitte der richtigen Länge, die an beiden Seiten die zur Sequenzierung nötigen Primersequenzen enthalten, werden zusammen als DNA-Bibliothek („DNA library“) bezeichnet.

Zumeist wird durch die Art der Library-Präparation auch die Lesart des jeweiligen NGS-Geräts bestimmt. Hierbei unterscheidet man Single-end-Reads, Paired-end-Reads sowie Mate-pair-Reads. Bei diesen unterschiedlichen Sequenzierarten werden entweder einzelne Abschnitte der DNA-Moleküle (je nach Technologie 50–10.000 bp), die jeweiligen Enden desselben Moleküls (etwa 2×50–250 bp) oder die beiden Enden eines sehr langen zirkularisierten Moleküls (etwa 50–100 bp eines bis zu 20 kb langen Moleküls) gelesen.

Amplifikation

Wie bei der Sanger-Sequenzierung gilt auch bei den meisten NGS-Plattformen, dass eine große Anzahl identischer Moleküle pro DNA-Fragment benötigt wird, um genügend messbares Signal zu erzeugen. Diese Amplifikation erfolgt analog zur PCR bei der Sanger-Sequenzierung und ist bei fast allen kommerziellen NGS-Plattformen zu finden. Die 2 wichtigsten Methoden sind die Bridge-PCR und die Emulsions-PCR. Mithilfe dieser Methoden wird die DNA-Bibliothek amplifiziert, um diese anschließend sequenzieren zu können.

Emulsions-PCR

Die Emulsions-PCR basiert auf einer Wasser-in-Öl-Emulsion, in der Wassertropfen im Öl wie kleine Mikroreaktoren fungieren können. Die PCR findet dann i. d. R. an der Oberfläche kleiner Kügelchen (Beads) statt. In jedem Mikroreaktor befindet sich theoretisch jeweils genau eines dieser Beads mit einem Oligonukleotidprimer sowie einem DNA-Molekül, dem Gegenprimer und allen anderen für die PCR benötigten Reagenzien in Lösung. Um dies zu erreichen, werden die DNA-Library-Fragmente in solch einem Verhältnis zu den Beads gegeben,

dass theoretisch an jedes Kügelchen genau ein einzelnes DNA-Molekül bindet. Dies folgt einer Poisson-Verteilung, sodass de facto bis zu 85% leere Beads in Kauf genommen werden. Auf diese Weise kann in jedem einzelnen Mikroreaktor eine einzelne PCR-Reaktion stattfinden. Nach der Amplifikation müssen die Emulsionen aufgebrochen und Ölreste sowie andere restliche Reagenzien entfernt werden. Leider ist es bei einer Emulsions-PCR nicht immer möglich, wirklich genau ein DNA-Fragment und ein Bead in jeweils einen Mikroreaktor zusammenzubringen. Manchmal geschieht es, dass mehrere DNA-Fragmente an ein Bead binden, sodass man man von polyklonalen Beads spricht. Diese Beads resultieren später bei der Sequenzierreaktion in nichtbrauchbaren Signalen und sollten daher so weit wie möglich vermieden werden. Ebenso ist es möglich, dass gar kein DNA-Molekül an ein Bead bindet. Ist diesem Fall erhält man ein leeres Bead, das in einer Sequenzierreaktion kein Signal produziert. Glücklicherweise kann man die leeren Beads aus der Masse der mit DNA belegten Beads weitgehend entfernen. Dieser Vorgang wird auch als Anreicherung bezeichnet, weil man hier die erwünschten Beads anreichert. Meist erfolgt das Entfernen der leeren Beads durch Zugabe von größeren Kügelchen (z. B. aus Polystyren). Nach der Beadanreicherung müssen diese anschließend auf dem Sequenzierträger aufgebracht werden. Bei den 454™- und Ion-Torrent™-Technologien werden die Beads in Vertiefungen („wells“) einer Picotiterplatte (PTP) bzw. eines Halbleiterchips deponiert. Bei der SOLiD™-Technologie werden sie an einen vorbehandelten Glasobjektträger gebunden. Der Vorteil dieser Technologie ist, dass hier Millionen von Amplifikationsreaktionen gleichzeitig durchgeführt werden können. Nachteilig ist, dass es auch zu dem Szenario eines DNA-Moleküls mit 2 Beads in einer PCR kommen kann, wodurch identische klonale Beads entstehen können. Diese sind sog. Duplikate, die unerwünscht sind. Ein anderer Nachteil ist, dass meist zusätzliche Geräte benötigt werden.

Bridge-Amplifikation

Neben der Emulsions-PCR hat sich v. a. eine Technologie durchgesetzt, bei der aus

Einzelmolekülen durch isothermale Polymerasereaktionen DNA-Kolonien („DNA colonies“ oder „colonies“) entstehen [12]. Die am häufigsten angewandte kommerzielle Variante hiervon ist die von Solexa® entwickelte Clusteramplifikation. Hierbei werden die DNA-Fragmente an Glasplatten einer Durchflusszelle („flow cell“) gebunden, auf der die Sequenzierung stattfindet. Diese Glasplatte enthält auf ihrer Oberfläche Oligonukleotide, an die die Adaptoren der DNA-Bibliothek binden können. Ausgehend von den gebundenen DNA-Fragmenten findet die Amplifikation als klonale Amplifikation (auch Bridge-Amplifikation genannt) auf dieser Glasplatte statt, bis diese vollständig geladen ist. In einer anschließenden Denaturierung werden die reversen Stränge entfernt, sodass Einzelstränge zur Sequenzierung zur Verfügung stehen. An diese kann nun der Sequenzierprimer gebunden werden, und die eigentliche Sequenzreaktion kann beginnen. Vorteil dieser Technologie ist wieder, dass sehr viele DNA-Fragmente gleichzeitig amplifiziert werden können. Weiterhin kann diese Technik mittlerweile in der Sequenziermaschine selbst angewandt werden, sodass keine zusätzliche Ausstattung benötigt wird. Nachteilig ist, dass die Cluster nicht immer klar voneinander abgegrenzt sind (hier verspricht der neue HiSeq™X Ten eine Verbesserung) und unterschiedlich groß sein können.

Sequenzierung

Die eigentliche Sequenzieretechnologie kann momentan in 3 Kategorien unterteilt werden:

- Pyrosequenzierung,
- „sequencing by synthesis“ (SBS) und
- „sequencing by ligation“ (SBL).

Wie diese Reaktionen im Detail aussehen, hängt von der jeweiligen Maschine ab und wird nachfolgend bei der Beschreibung der einzelnen Sequenziergeräte dargestellt. Sehr detailliert werden diese Technologien z. B. durch Metzker beschrieben [8].

Die 4 meist verbreiteten NGS-Plattformen

Roche 454™

Das 454™-Sequenziergerät war der erste kommerziell erwerbliche NGS-Apparat, und wurde im Jahre 2005 erstmals von seinem Erbauer Jonathan Rothberg publiziert [6]. Seit dem Jahr 2007 ist es unter dem Namen Roche 454™ erhältlich. Mit dieser Maschine wurde auch das Prinzip der bereits beschriebenen Library-Präparation und Emulsions-PCR für NGS zuerst kommerziell angeboten. Nach der Emulsions-PCR werden die angereicherten Beads auf eine Sequenzierplatte (Picotiterplatte) geladen und in die Maschine gegeben, wo die eigentliche Sequenzierreaktion stattfindet. Hierbei handelt es sich um eine Pyrosequenzierung. Gestartet wird der Sequenziervorgang mithilfe eines Sequenzierprimers, der an die (beadgekoppelten) einzelsträngigen DNA-Fragmente gebunden wird. Nun werden nacheinander die 4 verschiedenen dNTPs über die Picotiterplatte gespült. Bei jedem Einbau eines Nukleotids wird ein Pyrophosphat (PPi) abgespalten. Mithilfe von Sulfurase und ASP entsteht so ATP, das wiederum mit Luciferase und Luciferin reagiert. Hierbei entsteht Oxyluciferin und Licht, das von einer Kamera detektiert werden kann. Bei jedem Einbau eines Nukleotids entsteht also proportional zur Menge des freien PPi und damit zur Anzahl der eingebauten Nukleotide ein Lichtsignal. Diese Lichtsignale werden von einer Kamera detektiert und anschließend in die Abfolge der eingebauten Nukleotide umgeschrieben, sodass die sequenzierte DNA gelesen werden kann. Eine Besonderheit des 454™-Sequenziergeräts sind die langen DNA-Fragmente (Reads), die gelesen werden können. Eine Read-Länge von ungefähr 400 bp für den GS Junior und eine Read-Länge von bis zu 1000 bp für das GS-FLX+-System machten den 454™ besonders für Labore interessant, die auf lange Reads, z. B. für die Human-leukocyte-antigene (HLA)-Sequenzierung oder für De-novo-Assemblierung, angewiesen waren. Allerdings hat die Fa. Roche, obwohl sie im Besitz der ersten NGS-Maschine war, im Laufe der letzten Jahre im Vergleich zu an-

medgen 2014 · 26:231–238 DOI 10.1007/s11825-014-0447-7
© Springer-Verlag Berlin Heidelberg 2014

K. Neveling · A. Hoischen

Einführung in die Grundlagen der Hochdurchsatzsequenzierung

Zusammenfassung

Hintergrund. Next Generation Sequencing ist die neue Sequenziermethode für DNA. Aber was verbirgt sich eigentlich dahinter und was ist der Unterschied zur Sanger-Sequenzierung? In dieser Übersicht wird die neue Technologie ein wenig näher erläutert, und es wird erklärt, dass es sich hierbei nicht um eine einzige, sondern um viele neue Techniken handelt.

Technologie und Anwendung. Die momentan bekanntesten Sequenziergeräte und -techniken werden im Detail erklärt und die Gemeinsamkeiten der Maschinen, aber gerade auch die Unterschiede sowie Vor- und Nachteile dargestellt. Auf diese Weise soll der Leser erkennen, dass es nicht die perfekte Maschine für alle Applikationen gibt, sondern dass man für die jeweilige Fragestellung die Maschine aussuchen sollte, deren Spezifikationen sich hierfür am ehesten eignen. Auch die Möglichkeit des Outsourcings wird

besprochen, die sicherlich für einige Laboratorien interessant sein könnte. Desweiteren wird kurz erklärt, dass, analog zur Polymerase-Kettenreaktion bei der Sanger-Sequenzierung, auch beim Next Generation Sequencing zuvor oft die zu untersuchenden Regionen anreichert werden. Hierfür existieren verschiedene Methoden, deren Wahl i. d. R. von der Anzahl der zu untersuchenden Patienten und Gene abhängt.

Ausblick. Es wird ein Ausblick auf neueste Entwicklungen gegeben, die deutlich anzeigen, dass das Ende der genetischen Revolution noch nicht in Sicht ist.

Schlüsselwörter

Hochdurchsatzsequenzierung · Massive parallele Sequenzierung · DNA-Sequenzanalyse · Next Generation Sequencing (NGS) · Gezielte Resequenzierung

Introduction to the basics of next generation sequencing

Abstract

Background. Next generation sequencing is the new sequencing method for DNA. But what does this actually mean, and how does it differ from Sanger sequencing? In this review, insights into next generation sequencing are provided, while this does not represent a single technology, but rather comprises many different new techniques.

Technology and application. The currently most commonly used sequencing machines and techniques are explained in detail. Thereby, similarities in techniques, but also the differences, advantages and disadvantages are described. One has to realize that the reader will learn that not one machine is perfect for all applications, but that the best machine has to be chosen for a given application. In addition, the possibility of outsourc-

ing is discussed and could be interesting for some laboratories. Furthermore, analogous to the polymerase chain reaction for Sanger sequencing, one also has to enrich for the region of interest for most NGS applications. For this purpose, various methods can be selected, depending on the number of genes and samples to be investigated.

Future perspectives. Insights into future technologies are provided, underlining that the genetic revolution is ongoing.

Keywords

High-throughput nucleotide sequencing · Massively-parallel sequencing · Sequence analysis, DNA · Next generation sequencing · Targeted re-sequencing

deren NGS-Anbietern die Technologie wenig weiterentwickelt. Auch der Preis für Instrument und Reagenzien ist vergleichsweise hoch. Im Oktober 2013 hat die Fa. Roche verkündet, den Service der 454™-Technologie bis zum Jahr 2016 einzustellen.

Ion Torrent™

Die Ion-Torrent™-Technologie wurde genauso wie die Roche-454™-Technologie von Jonathan Rothberg entwickelt und weist daher große Ähnlichkeiten mit der Methode der ersten NGS-Maschine auf [9]. Allerdings handelt es sich bei den Ion-Torrent™-Sequenziergeräten um die ersten sog. Post-light-Sequenziergeräte, die nicht mehr mit fluoreszierenden Farb-

stoffen arbeiten und daher auch keine Kameras und Laser benötigen. Die Methode des Ion-Torrent™-Sequenziergeräts besteht darin, dass bei jedem Einbau eines Nukleotids in ein DNA-Fragment ein Proton freigesetzt wird. Dieses Proton kann in einem Puffer aufgefangen werden, wobei es den pH-Wert dieses Puffers verändert. Diese pH-Veränderung wird wiederum detektiert. Wie beim Roche-454™-Sequenziergerät werden auch beim Ion-Torrent™-Gerät DNA-Bibliotheken hergestellt, die an Beads (hier „ion sphere particles“ genannt) gekoppelt werden und mithilfe einer Emulsions-PCR amplifiziert werden. Die angereicherten Beads werden nachfolgend allerdings nicht auf eine Picotiterplatte geladen, sondern auf einen Halbleiterchip. Daher spricht man auch von „ion semiconductor sequencing“. Genau wie die Picotiterplatte enthält der Halbleiterchip viele kleine Vertiefungen, in die die einzelnen Micorbeads hineinfallen und in denen die Sequenzierreaktion stattfindet. Jede Vertiefung hat dabei ihren eigenen Sensor, der jede pH-Veränderung direkt in ein elektrisches Signal umwandelt. Die Detektion, ob ein Nukleotid eingebaut wurde oder nicht, kann demnach ganz genau per Vertiefung durchgeführt werden. Auch hier werden die 4 verschiedenen dNTPs wieder in Zyklen über den Chip gespült, sodass per Vertiefung ganz genau detektiert werden kann, welches Nukleotid eingebaut wurde.

Es gibt im Moment 2 verschiedene Modelle des Ion-Torrent™-Sequenziergeräts: Ion PGM™ („personal genome machine“) und Ion Proton™. Beim Ion PGM™ handelt es sich um eine sog. Benchtop-Maschine, ein Name der daher rührt dass diese Maschine klein genug ist, um auf eine Laborbank („bench“) zu passen. Solche Benchtop-Maschinen produzieren i. d. R. relativ kleine Datenmengen, sind aber dementsprechend kostengünstig und schnell, was insbesondere für diagnostische Laboratorien interessant ist. Für den Ion PGM™ sind im Moment 3 verschiedenen Chipformate erhältlich, die unterschiedlich viele Vertiefungen besitzen (1,2–11,1 Mio. Vertiefungen) und demnach unterschiedliche Datenmengen (etwa 50 Mb bis etwa 1 Gb) produzieren. Neben Sequenzierkits für die im Moment noch am weitesten

verbreiteten Read-Länge von 200 bp sind seit einiger Zeit auch Kits für eine Länge von 400 bp erhältlich, was die produzierten Datenmengen dementsprechend erhöht. Der Ion Proton™ produziert wesentlich mehr Daten als der Ion PGM™. Dennoch ist auch diese Maschine immer noch relativ klein und wird von der Fa. Life Technologies ebenfalls als Benchtop-Maschine bezeichnet. Im Moment gibt es für den Ion Proton™ nur einen einzigen Sequenzierchip. Dieser kann bis zu 10 Gb an Daten (Read-Länge: 200 bp) produzieren und hat eine Laufzeit von 2–4 h. Ein 2. Chip soll demnächst bis zu 32 Gb und später bis zu 64 Gb an Daten produzieren können. Obwohl dieser Chip schon Anfang 2013 angekündigt wurde, ist er bis jetzt nicht erhältlich.

Zu den größten Vorteilen von Ion Torrent™ gehört, dass die pH-Messung bei den Post-light-Sequenziergeräten die Verwendung von günstigen natürlichen, d. h. nicht fluoreszenzmarkierten, Nukleotiden erlaubt und der Einbau sowie die Detektion hiervon sehr schnell geschehen kann. Dies ermöglicht erstmals Gerätepreise von unter 50.000 €. Einer der momentan deutlichsten Nachteile ist die im Vergleich zu anderen Geräten relativ höhere Fehlerrate der produzierten Daten, besonders in Homopolymerbereichen des Genoms.

SOLiD™

Die SOLiD™-Maschinen stammen genau wie die Ion-Torrent™-Maschinen von der Fa. Life Technologies. Der Name SOLiD™ steht für „sequencing by oligonucleotide ligation and detektion“. Dieser Name sagt schon aus, dass die verwendete Sequenziermethode hier „sequencing by ligation“ ist, d. h. anstelle der üblichen Polymerasereaktion findet die Sequenzierung mittels einer Ligase statt [7, 12]. Die Probenvorbereitung für die SOLiD™-Maschinen ähnelt stark den schon beschriebenen Techniken, sodass Library-Präparation, Emulsions-PCR und Anreicherung hier nicht wiederholt erklärt werden. Für die Probenvorbereitung steht bei SOLiD™ (wie bei den meisten anderen Maschinen übrigens auch) ein eigenes automatisiertes System, das aus 3 Komponenten bestehende sog. EZ Bead™ System, zur Verfügung. Zum anschließenden Sequenzie-

ren werden die Beads auf Glasplatten pipettiert, die bei SOLiD™ als „flow chips“ bezeichnet werden. Das aktuelle SOLiD™-Gerät ist der 5500XL, der als Nachfolger der ursprünglichen Maschinen SOLiD™ 2, SOLiD™ 3, und SOLiD™ 4 fungiert. Pro 5500XL-Gerät kann man 2 „flow chips“ laden, wobei jeder „flow chip“ 6 Spuren („lanes“) enthält, auf die die Beads aufgetragen werden. Pro Sequenzierlauf können also 12 Spuren sequenziert werden. Der eigentliche Sequenziervorgang findet mithilfe von sog. 8-mer-Oligonukleotiden statt. Jedes dieser Oligonukleotide besteht aus 2 spezifischen Nukleotiden, gefolgt von 3 degenerierten Basen und 3 weiteren Basen. Zwischen den degenerierten und den letzten 3 Basen befindet sich eine Schnittstelle zur chemischen Spaltung. Alle 8-mer-Oligomere sind an einen von 4 verschiedenen Fluoreszenzfarbstoffen gekoppelt, wobei die ersten 2 Nukleotide festlegen, welcher Farbstoff angekoppelt wurde. Insgesamt gibt es 4 Farben, 1024 (4⁵) verschiedene Oktamere, und 256 Oktamere pro Farbe. Bei jeder Sequenzierrunde hybridisiert nun ein Oktamer an das zu sequenzierende DNA-Fragment, wobei nur die ersten 2 Nukleotide exakt passen. Die Fluoreszenz des eingebauten Oktamers wird detektiert, und anschließend werden die letzten 3 Basen, an denen auch der Farbstoff hängt, an der Schnittstelle abgetrennt. Dann wird das nächste Oktamer anhybridisiert, und der Prozess beginnt von vorne, bis das Ende des DNA-Fragments erreicht ist. Auf diese Weise werden immer 2 Nukleotide gelesen, dann 3 nicht, dann wieder 2 gelesen, dann wieder 3 nicht gelesen usw. Um die so entstandenen Leselücken zu füllen, beginnt der ganze Prozess wieder von vorne in einer 2. Ligationsrunde, wobei der erste Primer nun um genau ein Nukleotid versetzt bindet. Nach insgesamt 6 Ligationsrunden mit jeweils mehreren aufeinanderfolgenden Primerligationen sind schließlich alle Nukleotide sequenziert. Die Kombination der Fakten, dass das erste Oktamer immer um 1 Nukleotid versetzt beginnt und gleichzeitig immer 2 Nukleotide gelesen werden, führt dazu dass jedes Nukleotid doppelt gelesen wird. Auf diese Weise werden viele technisch bedingte Sequenzierfehler von vornherein vermieden, was i. d. R. zu einer sehr ge-

Tab. 1 Vergleich der vorgestellten Next-Generation-Sequencing-Systeme										
Hersteller	Amplifikation	Sequenziermethode	Gerätetyp	Geschätzte Kosten pro Gb ^a	Anschaffungspreis ^b	Leselänge (Single end-Reads) ^c	Leselänge (Paired end-Reads) ^d	Maximale Anzahl der Reads pro Lauf ^e	Maximale Datenmenge ^f	Zeit pro Sequenzierlauf ^g
Roche 454™	Emulsions-PCR	Polymerase (Pyrosequenzierung)	GS FLX Titanium XL+	Sehr hoch	Sehr hoch	Lang	–	Sehr gering	Sehr gering	Schnell
			GS Junior Titanium	Sehr hoch ^h	Gering	Lang	–	Sehr gering	Sehr gering ⁱ	Schnell
Illumina®	Bridge-PCR	„Sequencing by synthesis“	MiSeq	Hoch	Gering	Lang	Lang	Gering	Gering	Schnell bis mittel
			HiSeq™ 2500	Gering	Sehr hoch	Kurz	Kurz	Sehr hoch	Sehr hoch	Schnell bis mittel
			NextSeq™ 500	Gering	Hoch	Kurz	Kurz	Hoch	Mittel	Schnell bis langsam
			HiSeq™ X Ten	Gering ⁱ	Sehr hoch ^h	Kurz	Kurz	Sehr hoch ^h	Sehr hoch ^h	Langsam
Ion Torrent™	Emulsions-PCR	Polymerase („post light sequencing“)	Ion PGM™ (318)	Sehr hoch	Gering	Lang	–	Gering	Gering	Sehr schnell ^a
			Ion Proton™ (P1)	Gering	Hoch	Kurz	–	Gering	Mittel	Sehr schnell ^h
SOLiD™	Emulsions-PCR	„Sequencing by ligation“	5500xl (12 Spuren)	Hoch	Sehr hoch	Sehr kurz	Sehr kurz	Sehr hoch	Mittel-Hoch	Langsam ⁱ
Pacific Biosciences	Nicht erforderlich	„Single molecule real time“ (SMRT™)	1 SMRT™-Zelle (von 16)	Sehr hoch	Sehr hoch	Sehr lang ^h	–	Sehr gering ⁱ	Sehr gering	Sehr schnell ^h

^aSehr hoch: >100 \$, hoch: 50–100 \$, gering: <50 \$
^bSehr hoch: >500.000 \$, hoch: 150.000–500.000 \$, gering: <150.000 \$
^cSehr lang: >1 kb, lang: >200 bp, kurz: <200 bp, sehr kurz: <100 bp
^dSehr lang: >1 kb, lang: >200 bp; kurz: <200 bp, sehr kurz: <100 bp
^eSehr hoch: >1 Mrd., hoch: 100 Mio. bis 1 Mrd., gering: <100 Mio., sehr gering: <1 Mio
^fSehr hoch: >1 Tb, hoch: 200 Gb bis 1 Tb, mittel: 50–200 Gb, gering 1–50 Gb, sehr gering <1 Gb
^gSehr schnell: <4 h, schnell: <30 h, mittel: <5 Tage, langsam: >5 Tage
^hHöchster Wert
ⁱGeringster Wert

ringen Fehlerrate führt. Die gelesenen Signale entsprechen bei dieser Methode aber immer Farben und nicht Nukleotiden, wie bei anderen Methoden üblich, sodass man hier im Vergleich zur Base-space-Sequenzierung von Color-space-Sequenzierung spricht. Die Abfolge der Farben muss daher erst bioinformatisch wieder in die Nukleotidabfolge umgeschrieben werden, bevor die weitere Sequenzanalyse stattfinden kann.

Der SOLiD™ 5500XL gehört zu den größeren Sequenziermaschinen und wird üblicherweise zur Sequenzierung von ganzen Exomen oder Transkriptomen benutzt. Ein kompletter Lauf benötigt 8–10 Tage und produziert bis zu 200 Gb an Daten bei einer Read-Länge von bis zu 75 bp (plus eventuell 35 bp revers). Zu den Vorteilen der SOLiD™-Technologie gehört

die (theoretisch erreichbare) Qualität der Reads und die Anzahl der in einem Experiment gelesenen Moleküle. Zu den Nachteilen gehört die limitierte Leselänge, die die ligationsbasierte Technologie erlaubt, und die eingeschränkten Weiterentwicklungsmöglichkeiten. So hat sich z. B. die vielversprechende Amplifikationsmethode Wildfire nicht durchsetzen können und die Weiterentwicklung bleibt fraglich, da derselbe Hersteller derzeit v. a. auf die Ion-Torrent™-Technologie baut.

Illumina®/Solexa®

Die Fa. Illumina® ist im Moment im Bereich des NGS einer der Marktführer. Ihre Sequenzier-technologie wurde ursprünglich von der Fa. Solexa® entwickelt, die im Jahr 2007 von Illumina® übernommen

wurde [1, 2]. Das Prinzip, mit dem die Solexa®-Maschinen sequenzieren, hat zunächst einmal Ähnlichkeiten mit den bereits vorgestellten Sequenziermethoden. Allerdings findet die Library-Amplifikation nicht an Beads statt, sondern beruht auf der bereits beschriebenen Clusteramplifikation mittels Bridge-PCR. Bei der Solexa®-/Illumina®-Sequenzier-technologie handelt es sich um ein Sequencing-by-synthesis-Verfahren. Bei der Sequenzierreaktion werden zeitgleich alle 4 Nukleotide über die Glasplatte der Durchflusszelle gespült. Die eingebauten Nukleotide werden mithilfe einer Kamera über den jeweils angekoppelten Farbstoffes detektiert, während die übrigen nichteingebauten Nukleotide weggespült werden. Anschließend werden die terminalen 3'-Blocker, die verhindern dass nach

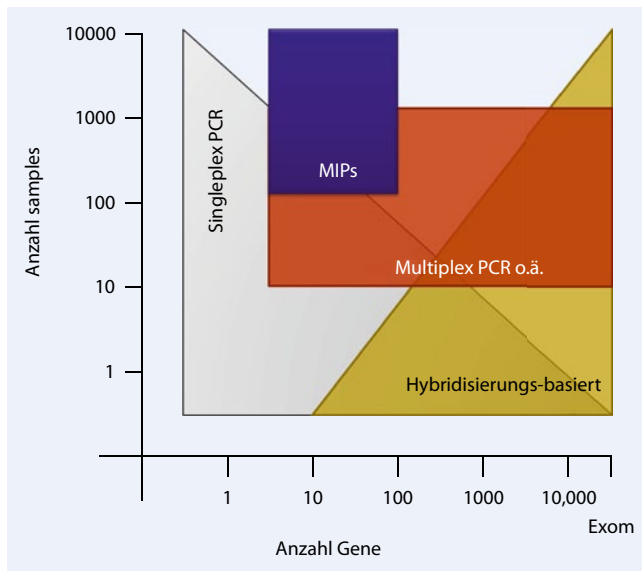


Abb. 1 ▲ Einsatzmöglichkeiten verschiedener Targeted-resequencing-Methoden in Abhängigkeit von Anzahl der zu testenden Gene und der Anzahl zu testender DNA-Proben. Bei der Einschätzung der Effektivität der verschiedenen Technologien berufen sich die Autoren auf eigene Erfahrungen basierend auf der praktischen Durchführbarkeit, Preis-Effizienz sowie dem zeitlichem Rahmen. Diese Abbildung entstand angelehnt an eine Abbildung von Mamanova et al. [5].

einem Einbau sofort weitere Nukleotide eingebaut werden, abgebaut und ein neuer Zyklus kann beginnen.

Die Fa. Illumina® produziert verschiedene Sequenzierplattformen. Das erste Modell war der Genome Analyzer von Solexa®. Mittlerweile sequenzieren viele Labore auf verschiedenen Illumina®-Maschinen, wie z. B. auf dem HiSeq™2000 bzw. HiSeq™2500 oder aber auf dem kleineren MiSeq™ (■ Tab. 1). Erst vor wenigen Wochen hat die Fa. Illumina® 2 neue Plattformen vorgestellt, den NextSeq™500 und den HiSeq™ X Ten. Der NextSeq™500 soll, neben verschiedenen anderen Optionen, in etwa 29 h bei einer Read-Länge von 2×150 bp bis zu 120 Gb an Daten produzieren können. Dadurch soll es z. B. möglich sein, ein komplettes humanes Genom mit einer 30-fachen Abdeckung („coverage“) in einem einzigen Lauf zu sequenzieren. Beim HiSeq™ X Ten dagegen handelt es sich eigentlich um 10 HiSeq™-X-Maschinen, die vom Hersteller nur im 10er-Pack erworben werden können. Laut Fa. Illumina® handelt es sich hierbei um das erste Sequenziergerät, dass das 1000-Dollar-Genom Realität werden lassen kann, allerdings bei Umlage der Amortisationskosten auf eine Jahresproduktion von etwa 18.000 humanen Geno-

men. Mit dem HiSeq™ X Ten wird die Sequenzierung von Genomen auf Populationsniveau ermöglicht, während Illumina® mit der Entwicklung des NextSeq500 die Lücke zwischen den beiden bestehenden Plattformen (MiSeq™ und HiSeq™) füllt. Die großen Vorteile der Instrumente von Illumina® liegen sicherlich in der Robustheit, der Datenqualität und der produzierten Datenmenge pro Sequenzierlauf. Desweiteren ist Illumina® stets um Weiterentwicklung bemüht, sowohl bezüglich der Produktion immer längerer Reads, qualitativ besserer Daten und flexiblerer Maschinen als auch hinsichtlich des Preises (z. B. durch Einschränkung von 4 auf 2 Farben bei dem neuen NextSeq500). Nachteile sind im Moment eher bei den älteren Maschinen zu finden, die im Betrieb vergleichsweise noch relativ teuer und unflexibel waren. Ob die neuen Maschinen halten, was sie versprechen, bleibt abzuwarten.

Gezielte Sequenzianreicherung zur Resequenzierung

Bei DNA-Anwendungen sollte man De-novo-Assemblierung und Resequenzierung unterscheiden. Die De-novo-Assemblierung wird oft durchgeführt, wenn

man z. B. bisher noch unbekannte Genome sequenzieren will. Hierfür sind möglichst lange und hochqualitative Reads von großem Vorteil. Die gezielte Resequenzierung eines bekannten Gens/Genoms kann hingegen bereits ab einer Read-Leselänge von 50 bp möglich sein. Diese Anwendung, die man auch als „targeted resequencing“ bezeichnet, hat für eine Revolutionierung der Humangenetik gesorgt. Sie wird separat in ■ Infobox 1 besprochen.

Anwendungsbasierte Auswahl der Plattform

Die Revolution auf dem Gebiet der Sequenzierung eröffnet zahlreiche Möglichkeiten, Nukleinsäuren aller Art schnell, günstig und in extrem hohem Durchsatz zu sequenzieren. Es gibt dabei grundsätzlich nicht das beste Sequenziergerät für alle Anwendungen, denn die Techniken unterscheiden sich, wie bereits beschrieben, v. a. bezüglich Leselänge, Durchsatz, Genauigkeit, Geschwindigkeit, Preis pro Gerät, Preis pro Experiment bzw. pro Gb sowie bezüglich der unterschiedlichen bioinformatischen Datenformaten (■ Tab. 1). Die möglichen Applikationen der beschriebenen Technologien umfassen dabei u. a. DNA-Sequenzierung, RNA-/Transkriptomsequenzierung, „chromatin immunoprecipitation sequencing“ (ChIP-seq), Chromatinkonformationssequenzierung (3C, 4C, 5C, HiC), und jeweils zahlreiche Unterarten der genannten Applikation. Leider fehlt hier der Raum, um diese Technologien eingehend zu beschreiben. Diese sowie viele neue Anwendungen sind aber durch Shendure und Lieberman Aiden gut zusammengefasst [11].

Immer mehr Labore entschließen sich, nicht mehr selbst in neue Apparate zu investieren. Aus mehreren Gründen bietet sich die Inanspruchnahme von Sequenzierungsdienstleistungen an. Zum einen wird das eigentliche Sequenzieren mehr und mehr zur Routine, zum anderen sind die Entwicklungen nach wie vor rasant, sodass nicht jedes Labor die Mittel für jedes neue Gerät zur Verfügung hat. Daher kann es attraktiv sein, für einige Applikationen die Serviceoption zu wählen. Einige Firmen haben sich sogar komplett

Infobox 1: „Targeted resequencing“

„Targeted resequencing“ fasst mehrere Technologien zusammen, die jeweils eine DNA-Bibliothek für bestimmte Regionen des humanen Genoms anreichern. Je nach Anwendung bieten sich hierzu verschiedene Techniken an (Abb. 1). Die wichtigsten sind sehr gut durch Mamanova et al. [16] zusammengefasst.

Polymerase-Kettenreaktion

Wie bei der Sequenzierung nach Sanger ist die PCR die einfachste Methode, um bestimmte Bereiche des Genoms anzureichern. Die hohe Spezifität und die relativ flexible Größe des anzureichernden Bereichs zählen hierbei sicherlich zu den Vorteilen. Desweiteren ist bei einigen PCR-Methoden die eigentliche Library-Präparation hinfällig, wenn während der PCR bereits die entsprechenden plattformspezifischen NGS-Primer verwendet werden. Die Limitierung auf wenige Regionen/Lozi gehört für die einfache PCR allerdings zu den Nachteilen. Einige Hersteller bieten mittlerweile Einzelmolekül-PCR im Hochdurchsatz an (z. B.: <http://raindancetech.com>. Zugriffen: 21. Mai 2014), und ermöglichen damit bis zu 4000 parallele Amplikonreaktionen [21]. Andere Hersteller setzten auf Multiplex-PCR und ermöglichen mehrere tausend Amplikons pro PCR-Ansatz, sodass sogar alle Exons des humanen Genoms, also das Exom, angereichert werden (z. B.: <http://www.lifetechnologies.com>. Zugriffen: 21. Mai 2014).

Hybridisierungsbasierte Methoden

Neben der PCR haben sich v. a. Technologien durchgesetzt, die auf Hybridisierung der Zielsequenzen basieren (z. B.: <http://www.agilent.com/>. Zugriffen: 21. Mai 2014 oder <http://www.nimblegen.com/>. Zugriffen: 21. Mai 2014). Hierbei boten sich zunächst mikroarraybasierte Hybridisierungsprotokolle an, die allerdings den Nachteil hatten, dass ein Hochdurchsatz nur schwer zu erreichen war. Erst die Umstellung auf eine Hybridisierung in Lösung (Oligonukleotidpools) ermöglichte weitere Automatisierung und erlaubt Zielsequenzen einer Länge von einigen Kb bis über 50 Mb. Dies ermöglicht nun die Anreicherung ganzer Exome, d. h. die Gesamtheit aller Exone des menschlichen Genoms. Gerade letztgenannte Anwendung hat für eine Revolution bei der Erforschung von Mendel-Erkrankungen gesorgt [14, 15, 18], da sie als erste Technologie eine genomweite Suche nach Punktmutationen des menschlichen Genoms in einem Experiment ermöglichte. In einigen Laboren findet diese Technologie bereits in der Diagnostik genetisch heterogener Erkrankungen Anwendung [13, 17, 24].

Von Multiplex-PCR bis „molecular inversion probes“

Bei der Analyse von Erkrankungen, die wenig heterogen sind, werden oft Methoden bevorzugt, die eine kleinere Anzahl von Genen bzw. Exons kostengünstig und sehr zuverlässig anreichern. Die bereits genannte Multiplex-PCR, aber auch weitere Technologien, wie „padlock probes“ [19] oder „molecular inversion probes“ (MIP; [22]) finden hierbei Anwendung [21]. Letztere ermöglichen eine Genanalyse für weniger als 1 € pro DNA-Probe und Gen, und erlauben so ein kostengünstiges Screening von tausenden von Proben [20]. Für einige dieser Technologien gibt es bereits diagnostische Anwendungsbeispiele [23] und in leichter Abwandlung kommerzielle Produkte (z. B. HaloPlex™ von der Fa. Agilent).

auf die Sequenzierung des humanen Genoms spezialisiert und bieten ihre Apparate nicht zum Kauf an, sondern offerieren nur einen Sequenzierservice, wie z. B. die Fa. Complete Genomics. Desweiteren gibt es in nahezu jedem Land inzwischen leistungsfähige Serviceanbieter auf Basis der beschriebenen Technologien.

Neue Technologien und Zukunftsaussichten

Neben den beschriebenen Geräten der sog. 2. Generation sorgen v. a. Sequenziergeräte der 3. Generation für neues Aufsehen. Diese neuen Maschinen haben gemeinsam, dass der Amplifikationsschritt (z. B. Oxford Nanopore) entfällt und be-

reits individuelle Moleküle gelesen werden können. Das erste kommerzielle Single-molecule-real-time (SMRT™)-Sequenziergerät wurde durch Pacific Biosciences (PacBio) entwickelt [3]. Einige andere Geräte konnten sich nach ihrer ersten Markteinführung nicht durchsetzen und sind nicht mehr erhältlich. Weitere Technologien der 3. Generation, die vor kurzem vorgestellt wurden, sind u. a. nanoporebasierte Systeme, wie Oxford Nanopore. Alle Single-molecule-Technologien haben nicht nur den Vorteil, ohne Amplifikationsschritt (und die damit verbundenen Fehler) auszukommen, sondern zeichnen sich v. a. auch dadurch aus, dass sie bislang unerreichte Leselängen erzielen (können). PacBio beschreibt z. B. Le-

selängen von bis zu 20 kb, was die Möglichkeit zur De-novo-Assemblierung von komplexen Genomen bzw. Genombereichen eröffnet [4]. Der größte Nachteil der Single-cell-Maschinen ist bislang die sehr viel höhere Fehlerrate der Sequenzier-Reads, wobei die Konsensusqualität bereits gute Daten produzieren kann. So hat PacBio das erste mittels Single-molecule-Technologie sequenzierte humane Genom publiziert (<http://www.ncbi.nlm.nih.gov/bioproject/PRJNA176729>. Zugriffen: 21. Mai 2014).

Ein Blick in die Zukunft verspricht, dass mit der Entwicklung von neuesten Technologien, wie von Nanoporesequenziergeräten, elektronischen Einzelmolekül-DNA-Kartierungen (z. B. Nabsys) oder rasterelektronenmikroskopbasiertem Lesen von einzelnen DNA-Molekülen, extrem lange Leselängen (>100 kb) erreichbar sein könnten. Die damit verbundenen Möglichkeiten der Sequenzierung des humanen Genoms in wenigen Stunden zu sehr geringem Preis führt hoffentlich zu dem entsprechenden Konkurrenzdruck und Wettbewerb, der auch für die NGS-Revolution der letzten Jahre gesorgt hat.

Neben dieser grundlagentechnikbasierten Weiterentwicklung sieht man auch eine zunehmende Fokussierung auf den diagnostischen Markt. Neben der Entwicklung von Desktop-Maschinen haben bereits weitere Hersteller den Verkauf von NGS-Apparaten angekündigt, die auf diagnostische Anwendungen zugeschnitten sind. Dieser Trend scheint sich auch bei etablierten Geräten durchzusetzen, so z. B. Qiagen den GeneReader™ oder GenapSys den GENIUS™. So hat die Fa. Illumina® eine Lizenzierung durch die Food and Drug Administration (FDA) für den MiSeq™ Dx erhalten. Die Fa. Life Technologies hat dies für den Ion PGM™ für das Jahr 2014 angekündigt.

Die Aussicht, dass weitere Technologien den Markt bereichern, bestätigt die Erwartung der Autoren, dass die NGS-Revolution weiter voranschreitet und dass die Humangenetik eine sehr zentrale Rolle in der Medizin und darüber hinaus spielen wird.

Fazit für die Praxis

- Unter dem Begriff Next Generation Sequencing (NGS) werden verschiedene neue Technologien zur Hochdurchsatzsequenzierung zusammengefasst.
- Verschiedene NGS-Plattformen und Geräte stehen derzeit von mehreren Anbietern für unterschiedliche Applikationen zur Verfügung. Sie unterscheiden sich v. a. bez. zugrunde liegender Technologie, Sequenziergeschwindigkeit, Read-Länge und Probendurchsatz bzw. Kosten pro Sequenzierlauf.
- Die Wahl des geeigneten Anreicherungsverfahrens hängt von der Anzahl der zu untersuchenden DNA-Proben und Gene ab.
- Viele Sequenzierverfahren basieren auf ähnlichen Grundprinzipien; allerdings versprechen Geräte der 3. Generation innovative Ansätze. Zukünftige Entwicklungen sind v. a. durch die Kombination von bereits bestehenden Anwendungen zu erwarten.

Korrespondenzadresse



K. Neveling

Department of Human Genetics, Radboud University Medical Center
Geert Grooteplein 10,
6500HB Nijmegen
Niederlande
kornelia.neveling@
radboudumc.nl



A. Hoischen

Department of Human Genetics, Radboud University Medical Center
Geert Grooteplein 10,
6500HB Nijmegen
Niederlande
alexander.hoischen@
radboudumc.nl

Danksagung. Die Autoren danken Dr. Christian Gilissen und den Arbeitsgruppen Developmental Genomics und Genomic Disorders für deren Hilfe bei der Erstellung dieses Manuskriptes. A.H. wurde durch die Netherlands Organization for Health Research and Development (ZonMW 916.12.095) unterstützt.

Einhaltung ethischer Richtlinien

Interessenkonflikt. K. Neveling und A. Hoischen geben an, dass kein Interessenskonflikt besteht.

Dieser Beitrag beinhaltet keine Studien an Menschen oder Tieren.

Literatur

1. Bentley DR (2006) Whole-genome re-sequencing. *Curr Opin Genet Dev* 16:545–552
2. Bentley DR, Balasubramanian S, Swerdlow HP et al (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456:53–59
3. Eid J, Fehr A, Gray J et al (2009) Real-time DNA sequencing from single polymerase molecules. *Science* 323:133–138
4. Huddleston J, Ranade S, Malig M et al (2014) Reconstructing complex regions of genomes using long-read sequencing technology. *Genome Res* 24:688–696
5. Mamanova L, Coffey AJ, Scott CE et al (2010) Target-enrichment strategies for next-generation sequencing. *Nat Methods* 7:111–118
6. Margulies M, Egholm M, Altman WE et al (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–380
7. McKernan K, Blanchard A, Kotler L, Costa G (2006) Reagents, methods, and libraries, for bead-based sequencing. US patent application 20080003571
8. Metzker ML (2010) Sequencing technologies – the next generation. *Nat Rev Genet* 11:31–46
9. Rothberg JM, Hinz W, Rearick TM et al (2011) An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475:348–352
10. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74:5463–5467
11. Shendure J, Lieberman Aiden E (2012) The expanding scope of DNA sequencing. *Nat Biotechnol* 30:1084–1094
12. Shendure J, Porreca GJ, Reppas NB et al (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309(5741):1728–1732
13. Ligt J de, Willemsen MH, Bon BW van et al (2012) Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* 367:1921–1929
14. Gilissen C, Hoischen A, Brunner HG, Veltman JA (2011) Unlocking Mendelian disease using exome sequencing. *Genome Biol* 12:228
15. Hoischen A, Bon BW van, Gilissen C et al (2010) De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat Genet* 42:483–485
16. Mamanova L, Coffey AJ, Scott CE et al (2010) Target-enrichment strategies for next-generation sequencing. *Nat Methods* 7:111–118
17. Neveling K, Feenstra I, Gilissen C et al (2013) A post-hoc comparison of the utility of sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. *Hum Mutat* 34:1721–1726
18. Ng SB, Buckingham KJ, Lee C et al (2010) Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet* 42:30–35
19. Nilsson M, Malmgren H, Samiotaki M et al (1994) Padlock probes: circularizing oligonucleotides for localized DNA detection. *Science* 265:2085–2088
20. O’Roak BJ, Vives L, Fu W et al (2012) Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 338:1619–1622
21. Schrauwen I, Sommen M, Corneveaux JJ et al (2013) A sensitive and specific diagnostic test for hearing loss using a microdroplet PCR-based approach and next generation sequencing. *Am J Med Genet A* 161A:145–152
22. Turner EH, Lee C, Ng SB et al (2009) Massively parallel exon capture and library-free resequencing across 16 genomes. *Nat Methods* 6:315–316
23. Umbarger MA, Kennedy CJ, Saunders P et al (2014) Next-generation carrier screening. *Genet Med* 16:132–140
24. Yang Y, Muzny DM, Reid JG et al (2013) Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* 369:1502–1511