

Geräteinterne DRAGEN™ Bio-IT Plattform auf den Systemen NextSeq™ 1000 und NextSeq 2000

Genaue und effiziente geräteinterne Lösung für die NGS-Datenanalyse (Next-Generation Sequencing, Sequenzierung der nächsten Generation)

- Bietet einen optimierten Workflow mit übersichtlicher grafischer Benutzeroberfläche für die hochgenaue Sekundäranalyse
- Generiert hochwertige Varianten-Calls direkt auf den Systemen NextSeq 1000 und NextSeq 2000
- Verringert den Bedarf an zusätzlicher Bioinformatikinfrastruktur und erfordert weniger Eingriffe des Anwenders als andere erhältliche DRAGEN-Software

illumina®

Einleitung

In vielen Laboren weltweit geht es darum, die im Genom enthaltenen Informationen zu erschließen. Aktuelle technologische Fortschritte vereinfachen den Zugang zu NGS-Plattformen und -Verfahren, die dazu beitragen können, die Geheimnisse des Genoms zu entschlüsseln. Die Systeme NextSeq 1000 und NextSeq 2000 eignen sich für Anwendungen mit geringem und mittlerem Durchsatz, darunter die Sequenzierung von Exomen, Transkriptomen und kleinen Genomen, die Target-Anreicherung, die Einzelzellsequenzierung und Shotgun-Metagenomik.

Ein wichtiger Vorteil der Systeme NextSeq 1000 und NextSeq 2000 ist die geräteinterne Implementierung der DRAGEN (Dynamic Read Analysis for GENomics) Bio-IT Plattform. Die geräteinterne DRAGEN-Software ermöglicht Laboren eine außergewöhnlich schnelle, genaue und umfassende NGS-Sekundäranalyse, sodass weniger zusätzliche Computerinfrastruktur und Bioinformatikressourcen erforderlich sind. Mit der geräteinternen DRAGEN-Software lassen sich die Sekundäranalyse und der Sequenzierungslauf gleichzeitig konfigurieren, was für einen optimierten Workflow mit weniger Eingriffen durch den Anwender und eine kürzere Durchlaufzeit sorgt. Für Labore, die ein hohes Maß an Konfigurierbarkeit wünschen, bietet Illumina zusätzlich Optionen zur Implementierung von DRAGEN-Software ([Abbildung 1](#)).

Die geräteinterne DRAGEN-Software ist die perfekte Ergänzung zu den vielseitigen und skalierbaren Systemen NextSeq 1000 und NextSeq 2000. Mehrere in die geräteinterne DRAGEN-Software integrierte, herausragende Pipeline-Algorithmen verhindern Engpässe bei der Datenanalyse in der Forschung und generieren schnell genaue Ergebnisse für eine Vielzahl von Anwendungen ([Tabelle 1](#)).

Effizienz

Die geräteinterne DRAGEN-Software auf den Systemen NextSeq 1000 und NextSeq 2000 ermöglicht Laboren die nahtlose Integration von Sequenzierungs- und Analysefunktionen in ihre Laufkonfiguration. Mit der geräteinternen DRAGEN-Software können Labore FASTQ-Dateien und VCF-Dateien (Variant Call Format) für nachgeschaltete Anwendungen unmittelbar nach dem Lauf direkt vom Gerät abrufen, was Zeit bei Dateiübertragung und Analyse einspart. Die geräteinternen DRAGEN-Pipeline-Algorithmen unterstützen Anfänger und erfahrene Anwender bei der Durchführung gängiger Analysefunktionen und verringern somit die Abhängigkeit von externen Informatikern. Die Dateispeicherung per DRAGEN Original Read Archive (ORA) vereinfacht den Umgang mit großen Datendateien. Die Lösung erzielt bei FASTQ-Dateien durchschnittlich eine vierfache verlustfreie Komprimierung. Sie verbraucht 80 % weniger Energie und ist damit nachhaltiger.

 <p>DRAGEN im Gerät (NextSeq 1000/2000)</p>	 <p>DRAGEN auf einem lokalen Server</p>	 <p>DRAGEN auf BaseSpace Sequence Hub</p>	 <p>DRAGEN auf Illumina Connected Analytics</p>	 <p>DRAGEN Multi-Cloud mit eigener Lizenz (Amazon Web Services, Azure)</p>
<ul style="list-style-type: none"> • Hochwertige Varianten-Calls direkt aus dem Sequenziersystem und damit geringerer Bedarf an zusätzlicher Bioinformatikinfrastruktur • Intuitive grafische Benutzeroberfläche • Optimierter Workflow mit weniger Eingriffen durch den Anwender • Kurze Durchlaufzeit 	<ul style="list-style-type: none"> • Skalierbare Hochdurchsatzanalysen auf lokalem Enterprise-Server • Vielseitige Befehlszeilenschnittstelle • Voll konfigurierbar • Regelmäßige Pipeline-Updates 	<ul style="list-style-type: none"> • Anwenderfreundliche Sekundäranalyseanwendungen • Intuitive grafische Benutzeroberfläche • Einfache Integration in andere Tools für die nachgeschaltete Analyse • Zuverlässige und sichere Cloudplattform, die den Bedarf an zusätzlicher Infrastruktur reduziert 	<ul style="list-style-type: none"> • Hochgradig skalierbar und konfigurierbar für Kunden mit höherem Durchsatz • Vielseitige Befehlszeilenschnittstelle • Einfache Integration in andere Tools für die nachgeschaltete Analyse 	<ul style="list-style-type: none"> • Nutzung von Illumina-Lizenz und -Software beim Cloudanbieter Ihrer Wahl • Vielseitige Befehlszeilenschnittstelle

Abbildung 1: Optionen für die Implementierung von DRAGEN Bio-IT-Pipelines passend zu den Anforderungen an die NGS-Analyse aller Labore.

Tabelle 1: Geräteinterne DRAGEN-Softwareanwendungen auf den Systemen NextSeq 1000 und NextSeq 2000

Anwendung ^a	Beschreibung
BCL-Konvertierung	Konvertiert von Illumina-Sequenziersystemen generierte BCL-Dateien in FASTQ-Dateien.
DRAGEN ORA-Kompression	Führt die verlustfreie, referenzbasierte Komprimierung von FASTQ-Dateien durch.
DRAGEN FASTQ + MultiQC	Liefert mithilfe von Hardwarebeschleunigung ohne zusätzliche Laufzeit FastQC-Metriken.
Genom	Führt das Mapping, Alignment und Calling kleiner Varianten für das Humangenom durch. Nur Keimbahn.
Anreicherung (einschließlich Exom)	Führt das Calling kleiner Varianten in Keimbahnproben oder Calling seltener Varianten in somatischen Proben durch.
DNA-Amplikon	Analysiert die genetische Variation in bestimmten genomischen Regionen. Verwendet die DRAGEN-DNA-Pipeline mit einem zusätzlichen Schritt zum Soft-Clipping von Primern und zum Neuschreiben von Alignments, um sicherzustellen, dass Primer-Sequenzen für Varianten-Calls verwendet werden.
RNA	Bietet einen (spleißstellensensitiven) RNA-Seq-Aligner mit optionaler rRNA-Filterung während des Alignments, was die Laufzeit und die Dateigröße reduziert.
Einzelzell-RNA	Verarbeitet eine Vielzahl von Einzelzell-RNA-Seq-Datensätzen, von Reads bis hin zu zellspezifischen Genexpressionsmatrizen.
Differenzial-expression	Führt den DESeq2-Algorithmus für RNA-Quantifizierungsdaten aus, die von der DRAGEN-RNA-Pipeline generiert wurden. Gibt Gene und Transkripte aus, die bei zwei Probengruppen eine jeweils andere Expression aufweisen.
NanoString GeoMx NGS	Optimiert die Analyse für Kunden, die sowohl GeoMx NGS- als auch Illumina-Geräte für Spatial-Genomics-Workflows verwenden.
Methylierung	Verarbeitet mit Bisulfit-Sequenzierung und TAPS (TET-Assisted Pyridine Sequencing, TET-gestützte Pyridin-Sequenzierung) generierte Methylierungsdaten.

a. Zusätzliche Anwendungspipelines sind unter BaseSpace Sequence Hub, Illumina Connected Analytics, DRAGEN Multi-Cloud und lokalen DRAGEN-Servern verfügbar.

Genauigkeit

DRAGEN-Analysepipelines liefern außergewöhnlich genaue Ergebnisse. Bei der Precision FDA Truth Challenge V2 (PrecisionFDA V2) 2020 erzielte DRAGEN v3.7 die genauesten Ergebnisse in den Kategorien „Alle Benchmark-Regionen“ und „Schwer zu mappende Regionen“ für Illumina-Sequenzierungsdaten. Weitere Innovationen bei Graph Genomes und Illumina Machine Learning mit DRAGEN v3.10 setzten über alle Sequenzierungstechnologien hinweg einen neuen Maßstab für die Datengenauigkeit im Datensatz „Alle Benchmark-Regionen“. Der erreichte F1-Score (der errechnete Anteil der richtig positiven und richtig negativen Ergebnisse an den Gesamtergebnissen) liegt bei 99,83 % (Abbildung 2). DRAGEN 3.10, Graph und Illumina Machine Learning erzielten im Vergleich zu allen PrecisionFDA V2-Einreichungen in der Region MHC (Major Histocompatibility Complex, Haupthistokompatibilitätskomplex) ebenfalls den höchsten F1-Score für das genaueste Calling.^{1,2}

Die außergewöhnlich effiziente DRAGEN Bio-IT Plattform zeichnet sich dank FPGA-Technologie (Field-Programmable Gate Array, feldprogrammierbarer Gate-Array) durch kurze Analysezeiten aus. Die exklusiv in die Systeme NextSeq 1000 und NextSeq 2000 integrierte FPGA-Karte ermöglicht den Einsatz hardwarebeschleunigter Sekundäranalyse-Informatikpipelines für zahlreiche Anwendungen. Diese DRAGEN-Pipelines werden kontinuierlich verbessert. Außerdem werden zusätzliche Pipelines veröffentlicht, die für bestmögliche Funktionalität, Genauigkeit und Geschwindigkeit sorgen. Die Plattform ist auf die Analyseanforderungen in der Forschung zugeschnitten und ermöglicht die schnelle und einfache Ausführung von Pipelines auf Produktionsebene, sodass sich Forscher umfassend auf die Ergebnisse konzentrieren können.

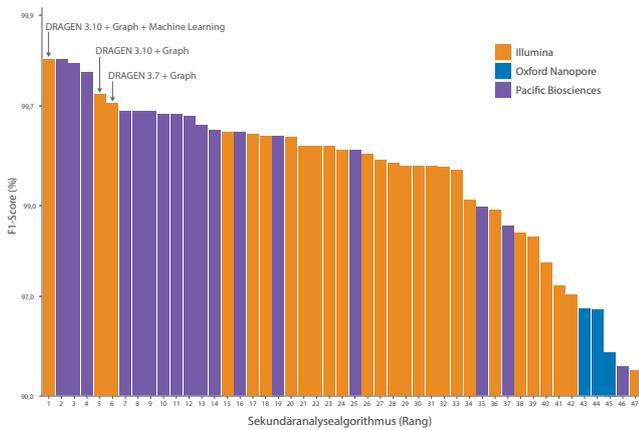


Abbildung 2: Genauigkeit von DRAGEN 3.10, Graph und Machine Learning beim Calling kleiner Varianten im Vergleich zu den Einreichungen bei der PrecisionFDA Truth Challenge v2 im Datensatz für die Kategorie „Alle Benchmark-Regionen“: Das Analyseverfahren DRAGEN 3.10 + Graph + Machine Learning wurde gemeinsam mit der Read-Einreichung von Pacific Biosciences mit dem ersten Platz ausgezeichnet. DRAGEN 3.10 + Graph ist DRAGEN 3.7 + Graph aufgrund von Verbesserungen bei der Handhabung von Graphen und Referenz/alternativem Contig überlegen.

Breite Auswahl an Anwendungen

Die geräteinterne DRAGEN-Software der Systeme NextSeq 1000 und NextSeq 2000 umfasst vielseitige Analysepipelines und kann verschiedene Ausgabedateien in verschiedenen Workflow-Phasen generieren (Abbildung 3).

Die enthaltene Auswahl an Analysepipelines eignet sich für eine Vielzahl von Analysetypen, darunter Einzelzell-, Exom- und RNA-Analysen (Tabelle 1).

Geräteinterne DRAGEN Single-Cell RNA-Pipelines

Die DRAGEN Single-Cell RNA-Pipelines sind schnelle und skalierbare Lösungen zur Verarbeitung zahlreicher Einzelzell-RNA-Seq-Datensätze mit anwenderfreundlichen Ausgabeformaten, einschließlich Zelltyp-Clustering-Diagrammen auf Basis von RNA-Expressionsaktivitäten (Abbildung 4). Darüber hinaus sind nützliche Erweiterungen für die Verarbeitung von multiplexierten Datensätzen, die aus mehreren Proben bestehen (z. B. mit Genotyp-Demultiplexing oder Zell-Hashing), und die Zählung der Expression von Zelloberflächenproteinen vorhanden. Zusätzlich werden zahlreiche Bibliotheksvorbereitungstypen unterstützt, sodass die Kompatibilität mit nachgeschalteten Analysetools gewährleistet ist.

Die geräteinterne Qualitätssicherung von Einzelzell-Expressionsbibliotheken und die Einzelzell-Analysepipeline werden schneller ausgeführt als bei cloudbasierten Lösungen und zudem ohne Genauigkeitsverluste. Die Einzelzell-Pipeline deckt auf den Systemen NextSeq 1000 und NextSeq 2000 mit einem einzigen Eingriff des Anwenders den gesamten Prozess von der Laufkonfiguration bis hin zur quantifizierten Expression je Zelle ab, wodurch weniger zusätzliche Rechenressourcen erforderlich sind.

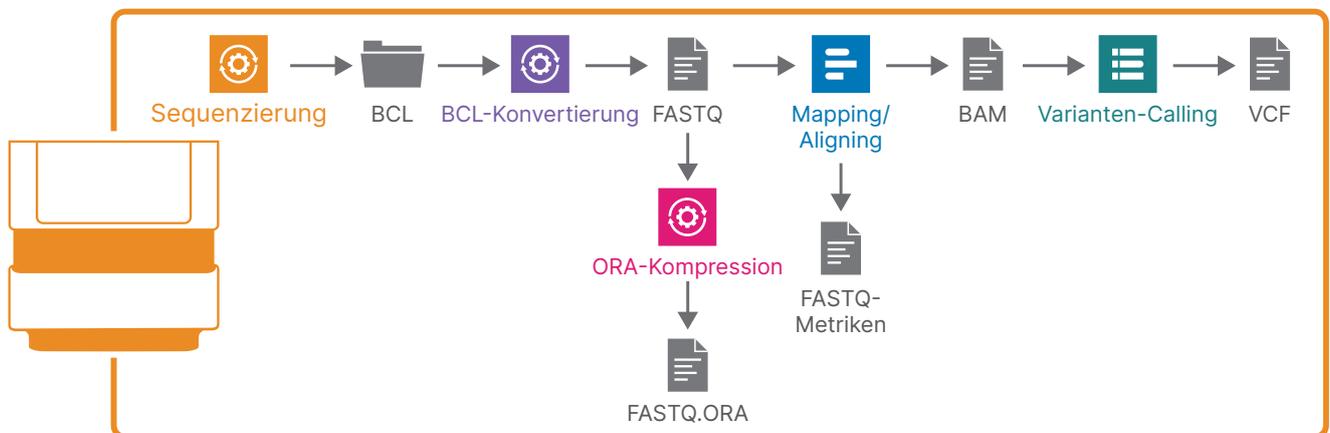


Abbildung 3: Die geräteinterne DRAGEN-Software auf den Systemen NextSeq 1000 und NextSeq 2000 bietet einen optimierten Datenverarbeitungsworkflow und ermöglicht die Sekundäranalyse. Die integrierte geräteinterne DRAGEN-Softwarelösung ist ein hocheffizientes Tool zur Generierung der Ergebnisse in den Formaten, die Labore in nachgeschalteten Anwendungen benötigen.

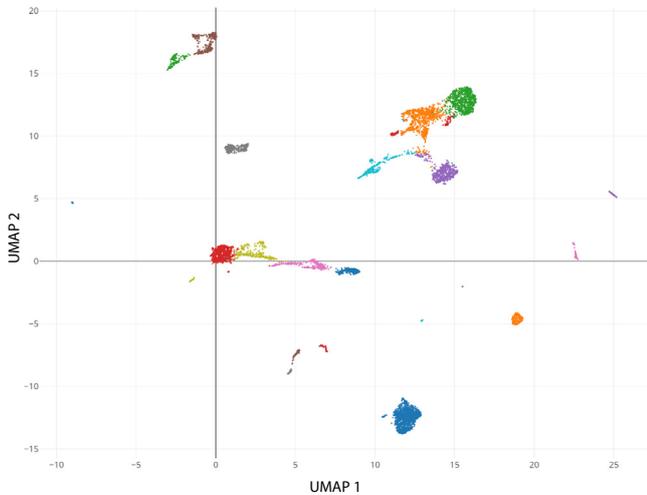


Abbildung 4: Beispiel für ein Zelltyp-Clustering-Diagramm aus der DRAGEN Single Cell RNA-Pipeline: Die UMAP-Analyse (Uniform Manifold Approximation and Projection, einheitliche Approximation und Projektion von Mannigfaltigkeiten) ermöglicht die Visualisierung einzelner Zellen nach Typ oder Funktionsstatus auf der Grundlage komplexer, mehrdimensionaler Expressionsprofile.⁴

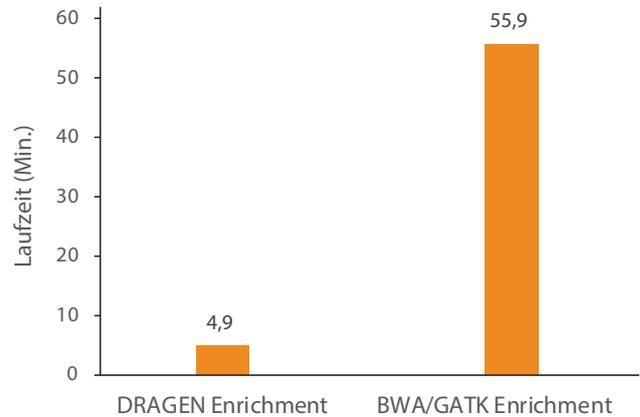


Abbildung 5: Durchschnittliche Laufzeit je Probe für die DRAGEN Enrichment-Pipeline und die BWA Enrichment-Anwendung auf BaseSpace™ Sequence Hub: Bei der Keimbahnreicherungsanalyse von 24 NGS-Replikaten zeigen sich kürzere Analysezeiten je Probe für die DRAGEN Enrichment-Pipeline als bei der vergleichbaren Analyse mit BWA-Anreicherung, gefolgt von Varianten-Calls mit dem GATK-Varianten-Caller.

Geräteinterne DRAGEN Enrichment-Pipeline

Die geräteinterne DRAGEN Enrichment-Pipeline ermöglicht eine schnelle Analyse, einschließlich einer fortschrittlichen Fehlermodellkalibrierung, und zeichnet sich durch eine höhere Genauigkeit bei Anreicherungs- und auf Hybridisierung basierenden Exom-Panels sowie Target-zugabebasierte Metriken aus. Die Software kann im Keimbahn- oder im somatischen Modus ausgeführt werden, wobei das Calling von kleinen Keimbahnvarianten oder von somatischen Varianten (mit niedriger Häufigkeit) erfolgen kann. Der Keimbahnmodus ermöglicht das Calling von Kopienzahlvarianten (CNV, Copy Number Variation) sowie von strukturellen Varianten. Die DRAGEN Enrichment-Pipeline ist schneller und genauer als die Analyse mit dem Burrows-Wheeler Aligner (BWA) und dem Varianten-Caller des Genome Analysis Toolkit (GATK), insbesondere bei der Bestimmung von Indels (Tabelle 2, Abbildung 5). Die Verarbeitung einer Fließzelle auf den Systemen NextSeq 1000 und NextSeq 2000 dauert in der Regel weniger als zwei Stunden.

Geräteinterne DRAGEN RNA-Pipeline

Die geräteinterne DRAGEN RNA-Pipeline führt die Sekundäranalyse von RNA-Transkripten durch. Die RNA-Pipeline bietet mehrere Betriebsmodi, einschließlich Nur-Referenz-Alignment und annotationsgestützten Alignments mit Genfusionserkennung. Mit einem optionalen Filter für ribosomale RNA (rRNA) lassen sich die Laufzeit der Analyse und die Dateigröße reduzieren. Das Genfusionsmodul nutzt den DRAGEN RNA-Spleißaligner für Split-Read-Analysen von ergänzenden (chimären) Alignments. Dadurch werden potenzielle Bruchstellen ermittelt und die Dauer der Gesamtanalyse auf ein Minimum reduziert. Die Funktionen der geräteinternen DRAGEN-Software ermöglichen die Ermittlung von Genfusionen, die Quantifizierung von Transkripten und das Varianten-Calling.

Tabelle 2: Ergebnisse zur Genauigkeit der DRAGEN Enrichment-Pipeline sowie von BWA/GATK

	SNV			Indel		
	Präzision (%)	Recall (%)	F1	Präzision (%)	Recall (%)	F1
DRAGEN Enrichment	99,7	94,64	97,11	96,54	83,83	89,74
BWA/GATK Enrichment	99,82	92,77	96,17	96,71	71,32	82,1

Zusammenfassung

Die DRAGEN Bio-IT-Software auf den Systemen NextSeq 1000 und NextSeq 2000 zeichnet sich durch einfach einzurichtende Bioinformatiklösungen und eine vereinfachte Benutzererfahrung in einer lokalen Implementierung aus. Die Sekundäranalyse der auf dem Gerät generierten NGS-Daten mit der Software erfolgt schneller als mit den DRAGEN-Pipelines auf BaseSpace™ Sequence Hub. Geräteinterne Pipelines, beispielsweise für Einzelzellen, Anreicherung und RNA, bieten gegenüber aktuellen Angeboten zahlreiche Geschwindigkeits- und Genauigkeitsvorteile.

Bestellinformationen

Produkt	Katalog-Nr.
NextSeq 1000 Sequencing System	20038898
NextSeq 2000 Sequencing System	20038897
NextSeq 1000 to NextSeq 2000 upgrade	20047256

Weitere Informationen

Systeme NextSeq 1000 und NextSeq 2000, illumina.com/systems/sequencing-platforms/nextseq-1000-2000.html

DRAGEN Bio-IT Platform, illumina.com/products/by-type/informatics-products/dragen-bio-it-platform.html

DRAGEN support, support.illumina.com/sequencing/sequencing_software/dragen-bio-it-platform.html

illumina®

+1.800.809.4566 (USA, gebührenfrei) | +1.858.202.4566 (Tel. außerhalb der USA)
techsupport@illumina.com | www.illumina.com

© 2022 Illumina, Inc. Alle Rechte vorbehalten. Alle Marken sind Eigentum von Illumina, Inc. bzw. der jeweiligen Eigentümer. Spezifische Informationen zu Marken finden Sie unter www.illumina.com/company/legal.html.
M-GL-00729 v1.0 DEU

Quellen

1. Food and Drug Administration. Truth Challenge V2: Calling Variants from Short and Long Reads in Difficult-to-Map Regions. precision.fda.gov/challenges/10. Aufgerufen am 14. März 2022.
2. Illumina. DRAGEN Sets New Standard for Data Accuracy in PrecisionFDA Benchmark Data. Optimizing Variant Calling Performance with Illumina Machine Learning and DRAGEN Graph. illumina.com/science/genomics-research/articles/dragen-shines-again-precisionfda-truth-challenge-v2.html. Aufgerufen am 14. März 2022.
3. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43(5):491-498. doi:10.1038/ng.806
4. McInnes, L, Healy J, and Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*. Überarbeitet am 18. September 2020. Aufgerufen am 5. Mai 2022. doi.org/10.48550/arXiv.1802.03426