

# Intel® Select Solutions für die Genomanalyse

**Zugriff auf Leistung, Skalierbarkeit und einfache Installation für Einblicke und Entdeckungen in der Genomik.**



Intel® Select Solutions für die Genomanalyse basieren auf der Referenzarchitektur BIGstack 2.0\*

Fortschritte in der Genomanalyse bieten neue Möglichkeiten, Krankheitsbilder zu verstehen, und sind immer häufiger die Grundlage für innovative Präzisionstherapien. Durch Verarbeiten, Speichern und Analysieren immer umfangreicherer Daten für die Genomsequenzierung werden neue Entdeckungen möglich. 2015 näherte sich die weltweite Sequenzierungsspeicherkapazität einem Petabyte (PB) pro Jahr, und sie verdoppelt sich weiterhin alle sieben Monate.<sup>1,2</sup> Bei dieser Rate wird die Genomsequenzierung in den nächsten fünf Jahren Hunderte von Petabyte pro Jahr erzeugen und könnte bis 2025 fast ein Zetabyte Speicher pro Jahr benötigen.<sup>1,2</sup>

Das Broad Institute of MIT and Harvard ([broadinstitute.org](http://broadinstitute.org)) ist mit täglich etwa 24 Terabyte an neuen Daten einer der weltweit größten Produzenten von Humangenomdaten. Gegenwärtig verwaltet das Broad Institute mehr als 50 PB an Daten.

Die Forscher benötigen die richtigen Tools, um diese enorme Datenmenge zeitnah analysieren zu können und entsprechende Erkenntnisse über Krankheiten und mögliche Behandlungsmethoden zu gewinnen. Sie benötigen Werkzeuge wie das GATK\* (Genome Analysis Toolkit), ein Set an führenden Software-Methoden, die vom Broad Institute entwickelt und von den meisten Genomforschungszentren weltweit geschätzt werden.

Das Broad Institute wird GATK 4.0 als nächste Hauptversion unter einer Open-Source-Lizenz für alle Anwender, auch für kommerzielle Zwecke, veröffentlichen. Eine Open-Source-Lizenz wird GATK einem breiteren Publikum von Wissenschaftlern und Forschern zugänglich machen und dazu beitragen, die Genomanalyse weltweit zu beschleunigen und voranzutreiben.

## Intel-Broad Center for Genomic Data Engineering

Intel und das Broad Institute arbeiten seit Jahren im Bereich der Computerinfrastruktur und Softwareoptimierung zusammen. 2017 starteten sie ein neues Projekt: das Intel-Broad Center for Genomic Data Engineering. Dabei handelt es sich um eine fünfjährige Zusammenarbeit zwischen beiden Partnern zur Vereinfachung und Beschleunigung der Genomik-Workflow-Ausführung mit GATK, dem Burrow-Wheeler Aligner (BWA), Cromwell, der Intel® Genomics Kernel Library (Intel® GKL), der GenomicsDB\* und anderen Tools und Techniken. Gemeinsam entwickeln, optimieren und nutzen Experten vom Broad Institute und Intel Tools und Infrastrukturen, um Wissenschaftler bei der Integration und Verarbeitung genomischer Daten zu unterstützen. Das Ergebnis wird eine wachsende Anzahl von Best Practices für die Bereitstellung von Hard- und Software für die Genomanalyse auf Plattformen mit Intel® Architektur sein, die auf Forschungsdatensätze angewandt werden können, die in privaten Rechenzentren gespeichert sind und sich auf Private, Public und Hybrid Clouds erstrecken werden.

Mit dem rasanten Wachstum der Genomdaten nutzen die Partner die Technologie, um Genomanalysen in großem Maßstab zu ermöglichen. Das Ergebnis resultierte bereits in den Intel® Select Solutions für die Genomanalyse, einer Suite aus optimierter Software und Referenzarchitekturen für die schlüsselfertige Konfiguration, Einrichtung und Installation zur Durchführung von Genomanalysen, die für GATK-Pipelines, Cromwell und GenomicsDB geeignet sind.

## Intel® Select Solutions für die Genomanalyse

Das Intel-Broad Center for Genomic Data Engineering arbeitet daran, GATK auf Intel-Architektur und -Technologien zu optimieren und eine Referenzarchitektur für die Genomanalyse zu definieren. Das Ergebnis sind die Intel Select Solutions für die Genomanalyse, entwickelt von Intel und dem Broad Institute. Sie basieren auf der Referenzarchitektur BIGstack 2.0\* und werden von Intel-Lösungsanbietern geliefert. Die Select Solutions bieten mit GATK 4.0 im Vergleich zu früheren Versionen der Genom-Software eine fünfmal höhere Gesamtleistung und verkürzen die Einrichtungzeit für eine Infrastruktur zur Beschleunigung der Genom-Workflows.<sup>3</sup> Die Leistungssteigerung beinhaltet eine 75-prozentige Beschleunigung für den BWA mit Intel® SSDs und eine zweifache Beschleunigung für HaplotypeCaller\* mit Intel® FPGAs (Field-Programmable Gate Arrays).<sup>3</sup> Die validierten Leistungs- und Qualitätsergebnisse wurden vom Broad Institute zertifiziert.

## Was sind Intel® Select Solutions?

Intel® Select Solutions sind verifizierte Hardware- und Software-Stacks, bei denen Rechen-, Speicher- und Netzwerkressourcen in Hinblick auf spezifische Software-Workloads optimiert sind. Grundlage für die Entwicklung dieser Lösungen sind die langjährigen Erfahrungen von Intel mit Anbietern von Branchenlösungen sowie die umfangreiche Zusammenarbeit mit den weltweit führenden Rechenzentrenbetreibern und Service Providern.

Damit sich ihre Lösung als Intel® Select Solution qualifiziert, müssen Anbieter:

1. die von Intel vorgegebenen Anforderungen an Hardware- und Software-Stacks erfüllen,
2. den Schwellenwert des Referenz-Benchmarks von Intel erreichen oder überbieten, und
3. einen detaillierten Umsetzungsleitfaden veröffentlichen, der den Kunden den Einsatz erleichtert.

Lösungsanbieter können ihre eigenen Optimierungen entwickeln, um den Nutzen ihrer Lösungen zu steigern.

## Intel® Select Solutions für die Genomanalyse

### Anwendung

Vorgefertigte Genom-Anwendungen:



BWA

### Plattform



Cromwell  
Workflow-Ausführung

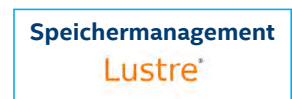


Optimierte Genom-Kernel-Bibliothek\*



GenomicsDB\*  
Großangelegte  
Analysen

### Infrastruktur



### Hardware



Abbildung 1: Detaillierter Überblick über die Konfiguration der Lösung



Bildnachweis: Len Rubenstein & Broad Institute

Leistungsstarke Datenanalyse-Cluster und optimierte Workflows für die Genomanalyse sind komplizierte Hard- und Softwaresysteme. Intel Select Solutions für die Genomanalyse sind End-to-End-optimierte Hardware- und Open-Source-Softwarekonfigurationen, die speziell zur Beschleunigung der Genomanalyse entwickelt wurden - sowohl für die Bereitstellung von Systemen als auch für die Software, die auf ihnen läuft - und verifizierte Stacks für die Einrichtung und Konfiguration dieser komplizierten Genomik-Pipelines zur Verfügung stellen.

Intel Select Solutions für die Genomanalyse sind so konzipiert, dass sie von kleinen bis zu sehr großen geclusterten Supercomputern skaliert werden können. Die kundenspezifischen Systeme können schnell und dynamisch an die jeweiligen Bedürfnisse angepasst werden. Unternehmen können mit zunehmenden Workloads skalieren. Und Intel® Select Solutions für die Genomanalyse umfassen Tools zum Auffinden, Zusammenstellen und Überwachen von Ressourcen mit leistungsstarker, moderner API-basierter Software.

**Tabelle 1:** Basis- und Plus-Konfigurationen für Intel® Select Solutions für die Genomanalyse

| BESTANDTEIL                 | INTEL® SELECT SOLUTIONS FÜR DIE GENOMANALYSE<br>BASIS-KONFIGURATION                            | INTEL® SELECT SOLUTIONS FÜR DIE GENOMANALYSE<br>PLUS-KONFIGURATION                             |
|-----------------------------|--|--|
| <b>HEAD NODES</b>           |  | 1 Head Node  |
| <b>PLATTFORM</b>            |  | Intel® Server-Mainboard S2600WFT   |
| <b>PROZESSOR</b>            |  | Intel® Xeon® Platinum 8180 Prozessor (28 MB Cache, 2,5 GHz)                                    |
| <b>ARBEITSSPEICHER</b>      |  | 128 GB DDR4-2666   |
| <b>HOST-ADAPTER</b>         |  | 100HFA016LS Intel® Omni-Path Host Fabric-Schnittstellenadapter<br>100er-Reihe, 1 Port PCIe x16 |
| <b>ANWENDUNGS-NODES</b>     | 1 Anwendungs-Node (auch Gateway-Node)  | 1 Anwendungs-Node  |
| <b>PLATTFORM</b>            | Intel® Server-Mainboard S2600WFT   | Intel® Server-Mainboard S2600WFT   |
| <b>PROZESSOR</b>            | Intel® Xeon® Platinum 8180 Prozessor (28 Kerne, 2,5 GHz)                                       | Intel® Xeon® Platinum 8180 Prozessor (28 Kerne, 2,5 GHz)                                       |
| <b>ARBEITSSPEICHER</b>      | 128 GB DDR4-2666   | 512 GB DDR4-2666   |
| <b>HOST-ADAPTER</b>         | 100HFA016LS Intel® Omni-Path Host Fabric-Schnittstellenadapter<br>100er-Reihe, 1 Port PCIe x16 | 100HFA016LS Intel® Omni-Path Host Fabric-Schnittstellenadapter<br>100er-Reihe, 1 Port PCIe x16 |
| <b>RECHENFLEXIBLE NODES</b> | 4 rechenflexible Nodes   | 8 rechenflexible Nodes   |
| <b>PLATTFORM</b>            | Intel® Server-Mainboard S2600WFT   | Intel® Server-Mainboard S2600WFT   |
| <b>PROZESSOR</b>            | Intel® Xeon® Platinum 8180 Prozessor (28 Kerne, 2,5 GHz)                                       | Intel® Xeon® Platinum 8180 Prozessor (28 Kerne, 2,5 GHz)                                       |
| <b>ARBEITSSPEICHER</b>      | 256 GB DDR4-2666   | 256 GB DDR4-2666   |

**Solution Brief | Intel® Select Solutions für die Genomanalyse**

|                                       |  |   |
|---------------------------------------|--|---|
| <b>LOKALER DATENSPEICHER</b>          | 7 x 4 TB Intel® SSD der Produktreihe DC P4600, PCIe HHHL   | 7 x 4 TB Intel® SSD der Produktreihe DC P4600, PCIe HHHL  |
| <b>HOST-ADAPTER</b>                   | 100HFA016LS Intel® Omni-Path Host Fabric-Schnittstellenadapter, PCIe x16   | 100HFA016LS Intel® Omni-Path Host Fabric-Schnittstellenadapter, PCIe x16  |
| <b>ERWEITERUNGSKARTEN</b>             | 1 Intel® Arria® 10 FPGA  | 1 Intel® Arria® 10 FPGA   |
| <b>HOCHDICHTE RECHEN-NODES</b>        | 4 hochdichte Rechen-Nodes  | 16 hochdichte Rechen-Nodes  |
| <b>PLATTFORM</b>                      | Intel® Server-Mainboard S2600BPB   | Intel® Server-Mainboard S2600BPB  |
| <b>PROZESSOR</b>                      | Intel® Xeon® Platinum 8176 Prozessor (28 Kerne, 2,1 GHz)   | Intel® Xeon® Platinum 8176 Prozessor (28 Kerne, 2,1 GHz)  |
| <b>ARBEITSSPEICHER</b>                | 256 GB DDR4-2666   | 256 GB DDR4-2666  |
| <b>LOKALER DATENSPEICHER</b>          | 4 TB Intel® SSD der Produktreihe DC S4500, serielle ATA (SATA)<br>2 TB Intel® SSD der Produktreihe DC P4501, PCIe M.2                              | 4 TB Intel® SSD der Produktreihe DC S4500, SATA<br>2 TB Intel® SSD der Produktreihe DC P4501, PCIe M.2  |
| <b>HOST-ADAPTER</b>                   | 100HFA016LS Intel® Omni-Path Host Fabric-Schnittstellenadapter 100er-Reihe, 1 Port PCIe x16  | 100HFA016LS Intel® Omni-Path Host Fabric-Schnittstellenadapter 100er-Reihe, 1 Port PCIe x16   |
| <b>ERWEITERUNGSKARTEN</b>             | 1 Intel® Arria® 10 FPGA  | 1 Intel® Arria® 10 FPGA   |
| <b>NETZWERK-INFRASTRUKTUR</b>         |  |   |
| <b>DATENNETZWERK</b>                  | Intel® Omni-Path Edge Switch der 100er-Serie, 24 Ports   | Intel® Omni-Path Edge Switch der 100er-Serie, 48 Ports  |
| <b>MANAGEMENTNETZWERK</b>             | 1 Gigabit pro Sekunde (Gbps) 48-Port-Switch  | 1 Gbps 48-Port-Switch   |
| <b>NFS SPEICHER-INFRASTRUKTUR</b>     |  |   |
| <b>PLATTFORM</b>                      | Intel® Server-Mainboard S2600WFT   |   |
| <b>PROZESSOR</b>                      | Intel® Xeon® Platinum 8180 Prozessor (28 Kerne, 2,5 GHz)   |   |
| <b>ARBEITSSPEICHER</b>                | 256 GB DDR4-2666   |   |
| <b>FESTPLATTEN</b>                    | 2 x 480 GB Intel® SSD der Produktreihe DC S3520 (gespiegeltes OS)  |   |
| <b>HOST-ADAPTER</b>                   | 12 Gbps Intel® RAID Controller RS3SC008, JBOD-Modus<br>100HFA016LS Intel® Omni-Path Host Fabric-Schnittstellenadapter 100er-Reihe, 1 Port PCIe x16 |   |
| <b>JBOD</b>                           | Intel® Speichersystem JBOD2312S3SP, 12 JBOD-Laufwerke  |   |
| <b>LAUFWERKE</b>                      | 12–24 Seagate*-Enterprise, Kapazität: 6–10 TB  |   |
| <b>RAID</b>                           | ZFS-on-RAID  |   |
| <b>LUSTRE* SPEICHER-INFRASTRUKTUR</b> |  |   |
| <b>LUSTRE MDS*</b>                    |  | 2 Metadaten-Server  |
| <b>PLATTFORM</b>                      |  | Intel® Server-Mainboard S2600WFT  |
| <b>PROZESSOR</b>                      |  | Intel® Xeon® Platinum 8180 Prozessor (28 Kerne, 2,5 GHz)  |
| <b>ARBEITSSPEICHER</b>                |  | 256 GB DDR4-2666  |
| <b>FESTPLATTEN</b>                    |  | 2 x 480 GB Intel® SSD der Produktreihe DC S3520 (gespiegeltes OS)   |
| <b>HOST-ADAPTER</b>                   |  | 12 Gbps Intel® RAID-Controller RS3SC008, JBOD-Modus;<br>100HFA016LS Intel® Omni-Path Host Fabric-Schnittstellenadapter 100er-Reihe, 1 Port PCIe x16 |
| <b>LUSTRE OSS</b>                     |  | 2 Objektspeicher-Server   |
| <b>PLATTFORM</b>                      |  | Intel® Server-Mainboard S2600WFT  |
| <b>PROZESSOR</b>                      |  | Intel® Xeon® Platinum 8180 Prozessor (28 Kerne, 2,5 GHz)  |
| <b>ARBEITSSPEICHER</b>                |  | 256 GB DDR4-2666  |
| <b>FESTPLATTEN</b>                    |  | 2 x 480 GB Intel® SSD der Produktreihe DC S3520 (gespiegeltes OS)   |
| <b>HOST-ADAPTER</b>                   |  | 12 Gbps Intel® RAID-Controller RS3SC008, JBOD-Modus;<br>100HFA016LS Intel® Omni-Path Host Fabric-Schnittstellenadapter 100er-Reihe, 1 Port PCIe x16 |
| <b>LUSTRE OST</b>                     |  | Objektspeicher-Ziel   |
| <b>JBOD</b>                           |  | Colfax CX4270-JBOD* 44 Laufwerke mit Dual-Expander  |
| <b>LAUFWERKE</b>                      |  | 44 Seagate-Enterprise, Kapazität: 6–10 TB   |
| <b>RAID</b>                           |  | ZFS-on-RAID   |
| <b>LUSTRE MDT</b>                     |  | Metadaten-Ziel  |
| <b>JBOD</b>                           |  | Intel® Speichersystem JBOD2224S2DP, 2U, Dual-Port   |

|   |  |  |
|---|--|--|
| SAS-SSDs                                    |  | 4 HGST* SAS 400 GB   |
| <b>SOFTWARE</b>                             | <p>GATK*, BWA und GATK-Workflows für Intel® Technologien optimiert</p> <p>Optimierter Cromwell-Workflow</p> <p>Intel® GKL mit optimierten Routinen zur Beschleunigung von Entwickler-Codes</p> <p>GenomicsDB*, spezialisiert auf groß angelegte Variantenanalyse</p> <p>HTCondor* Auftragsplaner zur Ausführung von geclusterten Analyseaufträgen</p> <p>Docker* für die Ausführung mehrerer Jobs in isolierten Containern über einen Cluster hinweg</p> <p>Apache Spark* für Big-Data-Analytics</p> <p>Lustre*, das Open-Source-Parallel-Dateisystem, für Hochleistungspeicherung</p> | <p>GATK*, BWA und GATK-Workflows für Intel-Technologien optimiert</p> <p>Optimierter Cromwell-Workflow</p> <p>Intel® GKL mit optimierten Routinen zur Beschleunigung von Entwickler-Codes</p> <p>GenomicsDB*, spezialisiert auf groß angelegte Variantenanalyse</p> <p>HTCondor* Auftragsplaner zur Ausführung von geclusterten Analyseaufträgen</p> <p>Docker* für die Ausführung mehrerer Jobs in isolierten Containern über einen Cluster hinweg</p> <p>Apache Spark* für Big-Data-Analysen</p> <p>Lustre*, das Open-Source-Parallel-Dateisystem, für Hochleistungspeicherung</p> |
| <b>FIRMWARE- UND SOFTWARE-OPTIMIERUNGEN</b> | Intel® AVX 512 (Advanced Vector Extensions)  | Intel® AVX-512   |

## Intel® Select Solutions für die Genomanalyse: Software, Firmware und Technologie-Konfiguration

Intel® Select Solutions für die Genomanalyse nutzen die leistungsstarken Funktionen der Intel-Architektur, einschließlich skalierbarer Intel® Xeon® Prozessoren, Intel® SSDs aus der Data-Center-Serie, Intel® Omni-Path-Architektur, Hochleistungs-Fabrics und Intel FPGAs.

Tabelle 1 zeigt die Hard- und Software für die "Basis"- und "Plus"-Konfigurationen von Intel® Select Solutions für die Genomanalyse. Um eine Lösung als Intel® Select Solutions zu bezeichnen, muss ein Serverhersteller oder Anbieter einer Rechenzentrumslösung diese oder bessere Konfigurationen verwenden.

Diese Lösungen können mit 2, 4, 16, 24, 36 oder 48 der angegebenen Rechensysteme und bei Bedarf an lokale und gemeinsam genutzte Speichermedien angepasst werden, um den Anforderungen der jeweiligen Umgebung gerecht zu werden.

### Vereinfachte Code-Entwicklung mit der Intel® GKL (Genomics Kernel Library)

Die Intel® GKL bietet in der Genomik verwendete Codes, die für die Intel-Architektur optimiert sind. Entwickler können diese Routinen nutzen, um ihre Code-Performance zu beschleunigen. Die Bibliothek ermöglicht es Entwicklern, sich auf Funktion und Betrieb ihres Codes zu konzentrieren (anstelle von spezifischen Optimierungen), während Intel® GKL alle Möglichkeiten der Intel-Architektur nutzt.

### Verbesserungen der Skalierbarkeit mit GenomicsDB\*

GenomicsDB ist ein einzigartiger Variantenspeicher, der Hunderttausende von Genomvarianten unterstützt. Er wurde zuerst von Intel Labs entwickelt und für die Anwendungsfälle des Broad Institute angepasst.

GenomicsDB wird mit GATK 4.0 ausgeliefert, was die Arbeitsabläufe erheblich beschleunigen wird. Er bietet einen großen Skalierungsvorteil für den HaplotypeCaller Genomic VCF (GVCF) Workflow. Ohne GenomicsDB benötigte das Broad Institute beispielsweise sechs Wochen, um eine Datenbank

aus 2.300 ganzen Genomen zu erstellen. Mit GenomicsDB war es möglich, Datenbanken mit fünfmal mehr Daten in nur zwei Wochen zu generieren.<sup>4</sup> Das ermöglichte das vom Broad Institute betreute Projekt gnomAD\* (Genomics Aggregation Database), das 15.000 ganze Genome umfasst - eine der größten genomischen Datensammlungen der Welt.<sup>4</sup>

### Fortlaufende Entwicklung

Es gibt große genomische Datenbanken auf der ganzen Welt, die große Vorteile für die weltweite Forschung bringen können. Die laufenden Arbeiten des Intel-Broad Center for Genomic Data Engineering setzen die Entwicklung von Intel® Select Solutions für die Genomanalyse fort, um effizient auf diese Datenbanken für Analysen zuzugreifen. In Zukunft werden integrierte Technologien die Konnektivität, Leistung, Privatsphäre und Sicherheit bieten, die für die Genomik in der Cloud und in gemeinsam genutzten Umgebungen erforderlich sind.

### Vorteile der Zusammenarbeit zwischen Intel und dem Broad Institute

Die Aktivitäten von Intel und dem Broad Institute bieten der Genomik-Gemeinschaft und den Technologen und Geschäftsführern, die sie unterstützen, viele Vorteile. Dazu gehören:

#### Vorteile für Wissenschaftler:

- Unterstützung für optimierte und effiziente Pipelines
- Optimierte, schlüsselfertige Lösungen
- Vorgefertigte WDL-Scripts (Workflow Description Languages)
- Peer-Anwendungsunterstützung
- Low-Touch-IT-Support
- Zugriff auf weitere interne Genomdaten
- Höhere statistische Leistung
- Open-Source-Software
- Flexible Anwendungsarchitektur



Bildnachweis: Len Rubenstein & Broad Institute

#### Vorteile für IT-Abteilungen:

- Einfache Umsetzung
- Skalierbarkeit
- Geringere Rüstzeiten
- Open-Source-Software ohne Lizenzkosten
- Bekannte Referenzarchitektur
- Hersteller- und Lösungssupport
- Optimale Nutzung der Hardware im Verhältnis zur Workload (z. B. vorgefertigte WDLs)

#### Vorteile für Unternehmer:

- Die Möglichkeit, Lösungen an das Budget anzupassen
- Geringe Stromkosten
- Vorkonfigurierte Lösungen zur Reduzierung von Rüstzeiten und Supportkosten
- Maximaler Nutzwert für interne genomische Daten
- Keine Lizenzgebühren
- Open-Source-Anwendungssoftware
- Erweiterbarkeit auf andere Anwendungen

#### OEM-Partner - Vereinfachung der Einrichtung von Genomanalyse-Clustern

Die Einführung von Intel Select Solutions für die Genomanalyse erleichtert die Verarbeitung von Genomik-Workloads. Sie ermöglicht auch die beschleunigte Einrichtung von vorhersagbaren Clustern für die Genomanalyse. Daher haben sich viele Anbieter von Hochleistungssystemen mit Intel zusammengetan und bieten Lösungen an, die den Anforderungen ihrer Kunden in der Genomik-Gemeinschaft gerecht werden.

„Unser Ziel ist es, die Aufgabe von Forschern zu erleichtern, aus immer mehr Genomikdaten immer aussagekräftigere Erkenntnisse zu gewinnen. Für uns hat der Einsatz von GATK 4 auf der Version 1.0 der Intel® Select Solutions für die Genomanalyse sofort einen 5-fachen Leistungszuwachs gebracht. Wir arbeiten mit Intel zusammen, um die GATK Best-Practices-Pipelines noch schneller, in noch größerem Umfang und mit einfacherem Einsatz für die Genomforschung weltweit laufen zu lassen.“

— Geraldine Van der Auwera, Associate Director of Outreach and Communications, Data Sciences Platform Group, Broad Institute

## Zugriff auf Leistung, Skalierbarkeit und einfache Bereitstellung für Genomanalysen

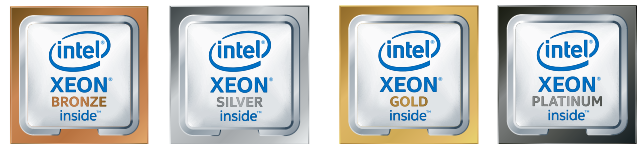
Die Arbeit der Genomwissenschaft ist entscheidend für das Verständnis von Krankheiten und die Schaffung von diagnostischen Instrumenten und sicheren, wirksamen Therapien. Genomdaten und -analysen schreiten schnell voran, da die Forscher die Technologie nutzen, um riesige Genomdatenbanken aufzubauen und die Möglichkeiten dieser Daten zu verstehen. Das Broad Institute ist einer der weltweit größten Verarbeiter von Genomdaten und seine GATK-Software ist das weltweit führende Genomanalyse-Tool für Analytics und Variantenforschung. Das Intel-Broad Center for Genomic Data Engineering vereint Wissenschaft und Technologie zur Optimierung von Genomanalyse-Codes und Workflows und zur Schaffung einer optimierten Infrastruktur - den Intel Select Solutions für die Genomanalyse - zur Ausführung dieser Workloads. Die Ergebnisse ermöglichen schnellere Analysen und kürzere Rüstzeiten für den Einsatz von Hardwarelösungen, die für die Genanalyse maßgeschneidert sind. Mehrere Systemhäuser bieten bereits Dienstleistungen an, um solche Systeme zu installieren, die weitere Entdeckungen in der Genetik ermöglichen.

## Skalierbare Intel® Xeon® Prozessoren

Skalierbare Intel® Xeon® Prozessoren:

- Bieten hohe Skalierbarkeit für Unternehmensrechenzentren
- Bieten Leistungssteigerungen für virtualisierte Infrastrukturen im Vergleich zu Prozessoren der Vorgängergeneration
- Sorgen für außergewöhnliche Ressourcennutzung und Agilität
- Ermöglichen verbesserte Daten- und Workload-Integrität und die Einhaltung gesetzlicher Vorschriften für Rechenzentrumslösungen

Die Produktreihe umfasst Intel® Xeon® Bronze, Silver, Gold und Platinum Prozessoren.



## Weitere Informationen

Das Intel-Broad Center for Genomic Data Engineering: [intel.com/broadinstitute](https://intel.com/broadinstitute)

„Big-Data-Genomik und optimierte Genom-Codes“:

[intel.com/content/www/us/en/healthcare-it/solutions/genomicscode.html](https://intel.com/content/www/us/en/healthcare-it/solutions/genomicscode.html)

Whitepaper von Intel und dem Broad Institute: „Infrastruktur für den Einsatz von GATK Best-Practices-Pipelines“:

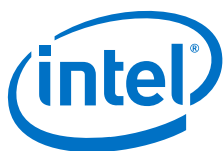
[intel.com/content/dam/www/public/us/en/documents/white-papers/deploying-gatk-best-practices-paper.pdf](https://intel.com/content/dam/www/public/us/en/documents/white-papers/deploying-gatk-best-practices-paper.pdf)

Intel® Select Solutions: [intel.com/selectsolutions](https://intel.com/selectsolutions)

Skalierbare Intel® Xeon® Prozessoren: [intel.com/xeonscalable](https://intel.com/xeonscalable)

Intel® Select Solutions werden unterstützt von Intel® Builders: <http://builders.intel.com>.

Folgen Sie uns auf Twitter: [#IntelBuilders](https://twitter.com/IntelBuilders)



<sup>1</sup> Stephens, Zachary D., et al. „Big Data: Astronomisch oder Genomisch?“ PLOS Biology. Juli 2015. <https://doi.org/10.1371/journal.pbio.1002195>.

<sup>2</sup> Robison, Reid J. „Wie groß ist das menschliche Genom?“ Präzisionsmedizin. Januar 2014. <https://medium.com/precision-medicine/how-big-is-the-human-genome-e90caa3409b0>.

<sup>3</sup> Intel. „Infrastruktur für den Einsatz von GATK Best-Practices-Pipelines“ November 2016. [intel.com/content/dam/www/public/us/en/documents/white-papers/deploying-gatk-best-practices-paper.pdf](https://intel.com/content/dam/www/public/us/en/documents/white-papers/deploying-gatk-best-practices-paper.pdf).

<sup>4</sup> Geraldine Van der Auwera, Ph.D. Broad Institute. Bio-IT World. Mai 2017

Tests dokumentieren die Leistung von Komponenten bei einem bestimmten Test und mit bestimmten Systemen. Unterschiede in der Hardware, der Software oder der Konfiguration des Systems beeinflussen die tatsächliche Leistung. Wer vor dem Kauf die Leistungsfähigkeit bewerten möchte, sollte hierzu andere Informationsquellen heranziehen. Ausführlichere Angaben über die Leistung und Benchmarkergebnisse finden Sie unter [intel.de/benchmarks](https://intel.de/benchmarks).

Die beschriebenen Kostensenkungsszenarien sind als Beispiele dafür gedacht, wie ein bestimmtes Produkt mit Intel-Technik unter den genannten Umständen und in der angegebenen Konfiguration zukünftige Kosten beeinflussen und Einsparungen ermöglichen kann. Die Umstände unterscheiden sich von Fall zu Fall. Intel übernimmt keine Gewähr für Kosten oder Kostensenkungen.

Durch Technologien von Intel ermöglichte Funktionsmerkmale und Vorteile hängen von der Systemkonfiguration ab und können entsprechend geeignete Hardware, Software oder die Aktivierung von Diensten erfordern. Die Leistungsmerkmale variieren je nach Systemkonfiguration. Kein Computersystem bietet absolute Sicherheit. Informieren Sie sich beim Systemhersteller oder Fachhändler oder auf [www.intel.de](https://www.intel.de).

Ausführlichere Angaben über die Leistung und Benchmarkergebnisse finden Sie unter [intel.de/benchmarks](https://intel.de/benchmarks).

Intel schließt sämtliche ausdrücklichen oder impliziten Garantien aus, einschließlich der impliziten Garantie der Marktgängigkeit oder der Eignung für einen bestimmten Zweck sowie der Nichtverletzung von Rechten, sowie Garantien, die aus einer Leistungserbringung, aus dem Handel oder der Verwendung im Handel entstehen.

Intel, das Intel Logo, Intel Optane und Intel Xeon sind Marken der Intel Corporation in den USA und/oder anderen Ländern.

\*Andere Marken oder Produktnamen sind Eigentum der jeweiligen Inhaber.