



Dieses Dokument und dessen Inhalt sind Eigentum von Illumina, Inc. und deren Partner-/Tochterunternehmen („Illumina“) und ausschließlich für den bestimmungsgemäßen Gebrauch durch den Kunden in Verbindung mit dem Gebrauch des hier beschriebenen Produkts (der hier beschriebenen Produkte) und für keinen anderen Bestimmungszweck ausgelegt. Dieses Handbuch und dessen Inhalt dürfen ohne schriftliches Einverständnis von Illumina nicht verwendet und zu keinem anderen Zweck verteilt bzw. anderweitig übermittelt, offengelegt oder auf irgendeine Weise reproduziert werden. Illumina überträgt mit diesem Dokument keine Lizenzen unter seinem Patent, Markenzeichen, Urheberrecht oder bürgerlichen Recht bzw. ähnlichen Rechten an Dritte.

Die Anweisungen in diesem Dokument müssen von qualifiziertem und entsprechend ausgebildetem Personal genau befolgt werden, damit die in diesem Dokument beschriebene Anwendung der Produkte sicher und ordnungsgemäß erfolgt. Vor der Verwendung dieser Produkte muss der Inhalt dieses Dokuments vollständig gelesen und verstanden worden sein.

FALLS NICHT ALLE HIERIN AUFGEFÜHRTE ANWEISUNGEN VOLLSTÄNDIG GELESEN UND BEFOLGT WERDEN, KÖNNEN PRODUKTSCHÄDEN, VERLETZUNGEN DER BENUTZER UND ANDERER PERSONEN SOWIE ANDERWEITIGER SACHSCHADEN EINTRETEN, WAS ZU EINEM ERLÖSCHEN DER PRODUKTGARANTIE FÜHRT.

ILLUMINA ÜBERNIMMT KEINERLEI HAFTUNG FÜR SCHÄDEN, DIE AUS DER UNSACHGEMÄSSEN VERWENDUNG DER HIERIN BESCHRIEBENEN PRODUKTE (EINSCHLIESSLICH TEILEN HIERVON ODER DER SOFTWARE) ENTSTEHEN.

© 2020 Illumina, Inc. Alle Rechte vorbehalten.

Alle Marken sind das Eigentum von Illumina, Inc. oder ihrer jeweiligen Inhaber. Spezifische Informationen zu Marken finden Sie unter [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html).

# Versionshistorie

Dokument-Nr.	Datum	Beschreibung der Änderung
Dokument-Nr. 15038356 v03	April 2020	Adresse der autorisierten europäischen Vertretung aktualisiert. Adresse der australischen Niederlassung aktualisiert.
Dokument-Nr. 15038356 v02	September 2017	Regulierungsetiketten wurden aktualisiert.
Dokument-Nr. 15038356 v01	Dezember 2016	Informationen in der Tabelle „Targets“ für Universal Kit 1.0 wurden korrigiert, indem in der Zeile „Deletions“ (Deletionen) Insertionen in Deletionen geändert wurde. Chrome wurde aus der Liste der unterstützten Browser für MiSeq Reporter bei Verwendung außerhalb des Geräts entfernt. Formatierungsfehler behoben.
Teile-Nr. 15038356 Rev. A	März 2014	Erste Version

# Inhaltsverzeichnis

Versionshistorie .....	iii
Inhaltsverzeichnis .....	iv
<b>Kapitel 1 Überblick .....</b>	<b>1</b>
Einleitung .....	2
Anzeigen von MiSeq Reporter .....	3
MiSeq Reporter-Konzepte .....	4
MiSeq Reporter-Benutzeroberfläche .....	5
Analyse erneut in die Warteschlange stellen .....	13
Analysekennzahlen .....	14
Analyseverfahren .....	16
MiSeqAnalysis-Ordner .....	17
<b>Kapitel 2 Datendarstellung .....</b>	<b>19</b>
Einleitung .....	20
Anforderungen für die Eingabedatei .....	21
Custom Amplicon-Workflow .....	22
Analyse-Ausgabedateien für die CF-Assays .....	34
<b>Kapitel 3 Installation und Fehlerbehebung .....</b>	<b>37</b>
Anforderungen für MiSeq Reporter außerhalb des Geräts .....	38
Installieren von MiSeq Reporter außerhalb des Geräts .....	39
Verwenden von MiSeq Reporter außerhalb des Geräts .....	41
Fehlerbehebung bei MiSeq Reporter .....	42
<b>Anhang A Universal Kit 1.0-Analyse-Ausgabedateien .....</b>	<b>45</b>
Arten von Analyse-Ausgabedateien .....	46
BAM-Dateiformat .....	47
VCF-Dateiformat .....	48
Amplicon-Abdeckungsdatei .....	51
Ergänzende Ausgabedateien .....	52
<b>Index .....</b>	<b>53</b>
<b>Technische Unterstützung .....</b>	<b>55</b>

[Diese Seite wurde absichtlich leer gelassen.]

# Überblick

Einleitung .....	2
Anzeigen von MiSeq Reporter .....	3
MiSeq Reporter-Konzepte .....	4
MiSeq Reporter-Benutzeroberfläche .....	5
Analyse erneut in die Warteschlange stellen .....	13
Analysekennzahlen .....	14
Analyseverfahren .....	16
MiSeqAnalysis-Ordner .....	17



## Einleitung

Das MiSeqDx™-Gerät umfasst drei Softwareanwendungen, die der Reihe nach ausgeführt werden, um Bilder der Cluster auf der Fließzelle zu erstellen, um die Bildanalyse und Base-Calling durchzuführen und um eine sekundäre Analyse auf dem Gerät durchzuführen.

- ▶ Während des Laufs erfasst die MiSeq Operating Software (MOS) Bilder von Clustern auf der Fließzelle für die Bildanalyse, steuert den Fließzellentisch, gibt Befehle zum Zuführen von Reagenzien und ändert die Temperaturen der Fließzelle.
- ▶ Die integrierte Primäranalysesoftware für die Echtzeitanalyse (RTA) führt die Bildanalyse und das Base-Calling durch und weist während des Laufs jeder Base für jeden Zyklus einen Qualitäts-Score zu. Nach dem Abschluss der Primäranalyse der RTA wird MiSeq Reporter zum Starten der Sekundäranalyse initiiert.
- ▶ MiSeq Reporter führt auf dem Gerät eine Sekundäranalyse zu Base-Calls und Qualitäts-Scores durch, die während des Sequenzierungslaufs von RTA generiert werden. MiSeq Reporter wird als Windows-Dienst ausgeführt und kann in einem Webbrowser angezeigt werden. Die Software kann auch außerhalb des Geräts auf einem Computer installiert werden. Weitere Informationen hierzu finden Sie unter *Installieren von MiSeq Reporter außerhalb des Geräts* auf Seite 39.

## Allgemeines zu Windows-Dienstanwendungen

Windows-Dienstanwendungen führen spezifische Funktionen ohne Benutzereingriff durch und werden fortlaufend im Hintergrund ausgeführt, solange Windows ausgeführt wird. Da MiSeq Reporter als ein Windows-Dienst ausgeführt wird, beginnt er automatisch mit der Sekundäranalyse, wenn die Primäranalyse abgeschlossen ist.

## Sequenzierung während der Analyse

Die Datenverarbeitungsressourcen des MiSeqDx-Geräts werden entweder für die Sequenzierung oder die Analyse verwendet. Wenn auf dem MiSeqDx-Gerät ein neuer Sequenzierungslauf gestartet wird, bevor die Sekundäranalyse eines vorherigen Laufs abgeschlossen ist, wird ein Bestätigungsdialogfeld angezeigt. Nach der Bestätigung des Sequenzierungslaufs wird die Sekundäranalyse beendet.

Verwenden Sie zum erneuten Starten einer Sekundäranalyse die Funktion **Requeue** (Wieder in die Warteschlange stellen) auf der Benutzeroberfläche von MiSeq Reporter, sobald der neue Sequenzierungslauf abgeschlossen ist. Zu diesem Zeitpunkt beginnt die Sekundäranalyse wieder von vorne.

## Anzeigen von MiSeq Reporter

Die MiSeq Reporter-Benutzeroberfläche kann nur in einem Webbrowser angezeigt werden. Um die MiSeq Reporter-Benutzeroberfläche während der Analyse anzuzeigen, öffnen Sie einen Webbrowser auf einem Computer, der Zugriff auf dasselbe Netzwerk wie das MiSeqDx-Gerät hat. Stellen Sie unter Verwendung einer der aufgeführten Methoden eine Verbindung zum HTTP-Dienst auf Port **8042** her:

- ▶ Verbinden unter Verwendung der Geräte-IP-Adresse, gefolgt von 8042.

IP-Adresse	HTTP-Service-Port	HTTP-Adresse
10.10.10.10, z. B.	8042	10.10.10.10:8042

- ▶ Verbinden mithilfe des Netzwerknamens für MiSeqDx, gefolgt von 8042.

Netzwerkname	HTTP-Service-Port	HTTP-Adresse
beispielsweise MiSeqDx01	8042	MiSeqDx01:8042

Bei Installationen von MiSeq Reporter außerhalb des Geräts Verbindung mithilfe der Methode für lokal installierte Dienstanwendungen herstellen, **localhost**, gefolgt von 8042.

Außerhalb des Geräts	HTTP-Service-Port	HTTP-Adresse
localhost	8042	localhost:8042

Weitere Informationen hierzu finden Sie unter *Installieren von MiSeq Reporter außerhalb des Geräts* auf Seite 39.

## MiSeq Reporter-Konzepte

Die folgenden Konzepte und Begriffe gelten für MiSeq Reporter.

Konzept	Beschreibung
Manifest	Die Datei, die ein Referenzgenom und Ziel-Referenzregionen angibt, die im Alignment-Schritt zu verwenden sind. Die von Cystic Fibrosis-Assays verwendete Manifest-Datei ist auf MiSeqDx vorgeladen.
Repository	Ein Ordner, in dem die bei Sequenzierungsläufen generierten Daten gespeichert werden. Jeder Laufordner ist ein Unterordner im Repository.
Laufordner	Die von der primären RTA-Analysesoftware gefüllte Ordnerstruktur (MiSeqOutput-Ordner) oder der von MiSeq Reporter (MiSeqAnalysis) gefüllte Ordner.
Probenblatt	Eine kommagetrennte Wertedatei (*.csv), die die zum Einrichten und Analysieren eines Sequenzierungslaufs erforderlichen Informationen enthält, darunter eine Liste von Proben und deren Indexsequenzen. Sie wird außerhalb des Geräts mithilfe von Illumina Worklist Manager erstellt. Das Probenblatt muss während der Laufeinrichtungsschritte auf MiSeqDx vorhanden sein. Nach dem Start des Laufs wird das Probenblatt automatisch in SampleSheet.csv umbenannt und in die Laufordner kopiert: MiSeqOutput und MiSeqAnalysis.
Workflow	Ein sekundäres Analyseverfahren, das von MiSeq Reporter ausgeführt wird. Der Workflow für jeden Lauf wird in den Probenblattdaten angegeben.

## MiSeq Reporter-Benutzeroberfläche

Wenn MiSeq Reporter im Browser geöffnet wird, wird der Hauptbildschirm mit einem Bild des Geräts in der Mitte angezeigt. Die Symbole „Settings“ (Einstellungen) und „Help“ (Hilfe) befinden sich in der oberen rechten Ecke und die Registerkarte „Analyses“ (Analysen) befindet sich in der oberen linken Ecke.

- ▶ **MiSeq Reporter Help** (MiSeq Reporter-Hilfe) – Wählen Sie das Symbol „Help“ (Hilfe), um die MiSeq Reporter-Dokumentation im Browserfenster zu öffnen.
- ▶ **Settings** (Einstellungen) – Wählen Sie das Symbol „Settings“  aus, um die Server-URL und den Repository-Pfad zu ändern.
- ▶ **Registerkarte „Analyses“** (Analysen) – Wählen Sie „Analyses“ (Analysen), um die Registerkarte zu erweitern. Die Registerkarte „Analyses“ zeigt eine Liste der Analyseläufe, die entweder abgeschlossen sind, zur Analyse in die Warteschlange gestellt wurden oder aktuell verarbeitet werden.

Abbildung 1 MiSeq Reporter-Hauptbildschirm

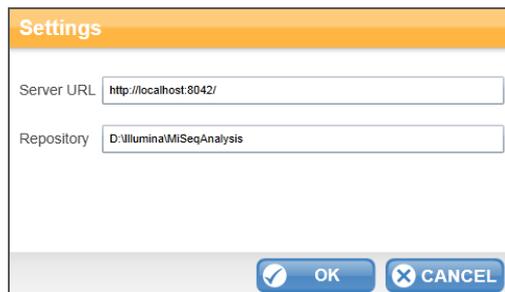


## Server-URL- oder Repository-Einstellungen

Verwenden Sie die Funktion „Settings“ (Einstellungen)  zum Ändern der Server-URL und des Repository-Pfads:

- ▶ **Server URL** (Server-URL) – Der Server, auf dem MiSeq Reporter ausgeführt wird.
- ▶ **Repository path** (Repository-Pfad) – Speicherort des Analyseordners, in dem die Ausgabedateien gespeichert werden.

Abbildung 2 Einstellungen für Server-URL und Repository


 The image shows a "Settings" dialog box with an orange header. It contains two text input fields: "Server URL" with the value "http://localhost:8042/" and "Repository" with the value "D:\Illumina\MiSeqAnalysis". At the bottom, there are two buttons: "OK" and "CANCEL".

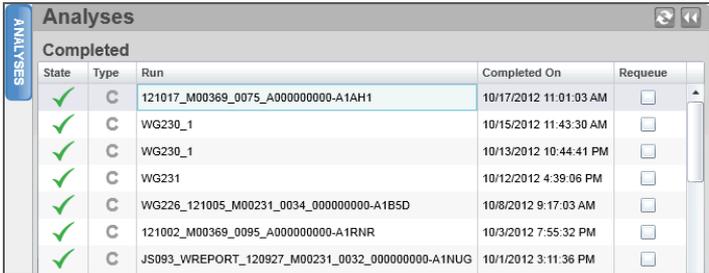
In der Regel müssen diese Einstellungen nicht geändert werden, es sei denn, MiSeq Reporter wird außerhalb des Geräts ausgeführt. Stellen Sie in diesem Fall für den Repository-Pfad den Speicherort des MiSeqOutput-Ordners im Netzwerk ein. Weitere Informationen hierzu finden Sie unter *Verwenden von MiSeq Reporter außerhalb des Geräts* auf Seite 41.

## Registerkarte „Analyses“ (Analysen)

Auf der Registerkarte „Analyses“ (Analysen) werden alle Sequenzierungsläufe aufgelistet, die sich im angegebenen Repository befinden. Auf dieser Registerkarte können die Ergebnisse aus einem der aufgelisteten Läufe geöffnet werden oder ein ausgewählter Lauf kann zur Analyse erneut in die Warteschlange gestellt werden.

Klicken Sie in der oberen rechten Ecke auf das Symbol **Refresh Analysis List** (Analyse-Liste aktualisieren) , um die Liste zu aktualisieren.

Abbildung 3 Erweiterte Registerkarte „Analyses“ (Analysen)



State	Type	Run	Completed On	Requeue
✓	C	121017_M00369_0075_A000000000-A1AH1	10/17/2012 11:01:03 AM	<input type="checkbox"/>
✓	C	WG230_1	10/15/2012 11:43:30 AM	<input type="checkbox"/>
✓	C	WG230_1	10/13/2012 10:44:41 PM	<input type="checkbox"/>
✓	C	WG231	10/12/2012 4:39:06 PM	<input type="checkbox"/>
✓	C	WG226_121005_M00231_0034_000000000-A1BSD	10/8/2012 9:17:03 AM	<input type="checkbox"/>
✓	C	121002_M00369_0095_A000000000-A1RNR	10/3/2012 7:55:32 PM	<input type="checkbox"/>
✓	C	JS093_WREPORT_120927_M00231_0032_000000000-A1NUG	10/1/2012 3:11:36 PM	<input type="checkbox"/>

Die Spalten der Registerkarte „Analyses“ (Analysen) sind „State“ (Status), „Type“ (Typ), „Run“ (Lauf), „Completed On“ (Abgeschlossen am) und „Requeue“ (Erneut in die Warteschlange stellen):

- ▶ **State** (Status) – Zeigt den aktuellen Status der Analyse anhand eines von drei Statussymbolen an.

Tabelle 1 Symbole für den Status der Analyse

Symbol	Beschreibung
	Gibt an, dass die Sekundäranalyse erfolgreich abgeschlossen wurde.
	Gibt an, dass die Sekundäranalyse noch durchgeführt wird.
	Gibt an, dass Fehler aufgetreten sind und die Sekundäranalyse daher fehlgeschlagen ist.

- ▶ **Type** (Typ) – Listet mittels eines Buchstabenbezeichners den Analyse-Workflow auf, der mit dem jeweiligen Lauf verbunden ist. Für CF-Assays und Universal Kit 1.0 lautet der Buchstabenbezeichner C.
- ▶ **Run** (Lauf) – Der Name des Laufordners im MiSeqOutput- und MiSeqAnalysis-Ordner.
- ▶ **Completed On** (Abgeschlossen am) – Das Datum, an dem die Sekundäranalyse abgeschlossen wurde.
- ▶ **Requeue** (Erneut in die Warteschlange stellen) – Aktivieren Sie dieses Kontrollkästchen, um einen Auftrag zur Analyse erneut in die Warteschlange zu stellen. Die Schaltfläche „Requeue“ (Erneut in die Warteschlange stellen) wird eingblendet. Weitere Informationen hierzu finden Sie unter *Analyse erneut in die Warteschlange stellen* auf Seite 13.

Wenn die Analyse in die Warteschlange gestellt wird, erscheint der Lauf im unteren Teil der Registerkarte „Analyses“ (Analysen) mit dem Symbol, das angibt, dass die Analyse noch durchgeführt wird .

Abbildung 4 Ein in die Warteschlange gestellter Lauf auf der Registerkarte „Analyses“ (Analysen)

	C	JS092_120921_M00369_0092_A000000000-A1AHY	9/22/2012 9:20:38 PM	<input type="checkbox"/>
	C	WG223	9/20/2012 10:15:33 PM	<input type="checkbox"/>
<b>Queued</b>				
State	Type	Run	Started/Queued On	
	C	121017_M00231_0020_A000000000-A1B5F	10/17/2012 12:11:33 PM	

## Registerkarten „Analysis Information“ (Informationen zur Analyse) und „Results“ (Ergebnisse)

Nach Auswahl eines Laufs auf der Registerkarte „Analyses“ (Analysen) werden Informationen und Ergebnisse für diesen Lauf auf einer Reihe von Registerkarten auf der MiSeq Reporter-Benutzeroberfläche angezeigt: Summary (Zusammenfassung), Details, Analysis Info (Informationen zur Analyse), Sample Sheet (Probenblatt), Logs (Protokolle) und Errors (Fehler). Anfänglich erscheinen die Informationen auf den Registerkarten „Analysis Info“ (Informationen zur Analyse) und „Sample Sheet“ (Probenblatt). Nach Abschluss der Analyse werden alle Registerkarten mit den entsprechenden Daten gefüllt.

Abbildung 5 Registerkarten „Information“ (Informationen) und „Results“ (Ergebnisse)

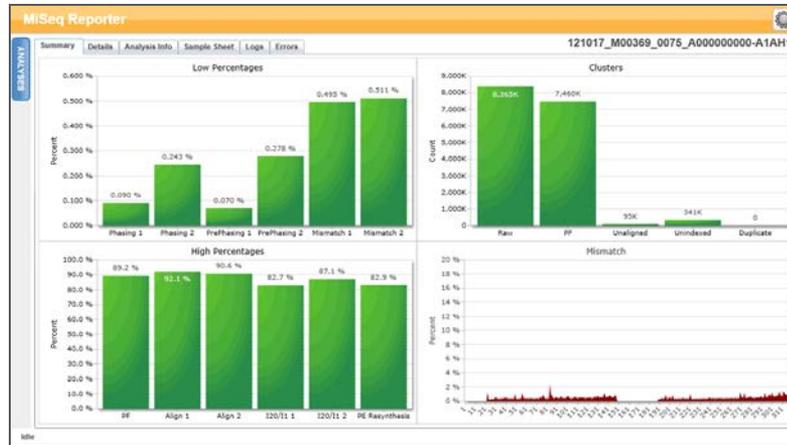


## Registerkarte „Summary“ (Zusammenfassung)

Die Registerkarte „Summary“ (Zusammenfassung) enthält eine Zusammenfassung der Analyseergebnisse. Auf der Registerkarte „Summary“ (Zusammenfassung) werden vier Diagramme angezeigt:

- ▶ **Diagramm „Low Percentages“** (Niedrige Prozentwerte) – Zeigt Phasierung, Vorphasierung und Nichtübereinstimmungen in Prozent an. Niedrige Prozentwerte deuten auf gute Laufstatistiken hin. Weitere Informationen finden Sie unter *Phasierung und Vorphasierung* auf Seite 14.
- ▶ **Diagramm „High Percentages“** (Hohe Prozentwerte) – Zeigt Cluster nach Filterung, Alignment auf der Basis einer Referenz und Intensitäten in Prozent an. Hohe Prozentwerte deuten auf gute Laufstatistiken hin.
- ▶ **Diagramm „Clusters“** (Cluster) – Zeigt die Anzahl der rohen Cluster, der Cluster nach Filterung, der Cluster ohne Alignment, der Cluster, die keinem Index zugeordnet sind, und der Duplikate an.
- ▶ **Diagramm „Mismatch“** (Nichtübereinstimmung) – Zeigt Nichtübereinstimmungen pro Zyklus an. Eine Nichtübereinstimmung bezieht sich auf eine Nichtübereinstimmung zwischen dem Sequenzierungs-Read und einem Referenzgenom nach dem Alignment.

Abbildung 6 Registerkarte „Summary“ (Zusammenfassung)



### Registerkarte „Details“

Die Registerkarte „Details“ enthält Einzelheiten zu den Analyseergebnissen. Die folgenden Tabellen können je nach verwendetem Assay oder Kit auf der Registerkarte „Details“ angezeigt werden:

- ▶ **Tabelle „Samples“** (Proben) – Fasst die Sequenzierungsergebnisse für jede Probe zusammen.
- ▶ **Tabelle „Targets“** – Zeigt die Statistik für die Zielregionen einer ausgewählten Probe an. (Nur Universal Kit 1.0)
- ▶ **Tabelle „Variants“** (Varianten) – Zeigt die Unterschiede zwischen Proben-DNA und der Referenz an.
- ▶ **Diagramm „Coverage“** (Abdeckung) – Zeigt an, wie tief die Probe sequenziert wurde, indem die Anzahl der in der Probensequenz vorhandenen Basen für jede Position der Referenz gemessen wird.
- ▶ **Diagramm „Qscore“** – Zeigt den durchschnittlichen Qualitäts-Score an, der die geschätzte Wahrscheinlichkeit eines Base-Call-Fehlers darstellt. Weitere Informationen hierzu finden Sie unter *Diagramm „Qscore“* auf Seite 32.
- ▶ **Diagramm „Variant Score“** (Varianten-Score) – Zeigt die Position von SNVs und Indels.

Abbildung 7 Registerkarte „Details“ für den CF 139-Varianten-Assay, Beispiel

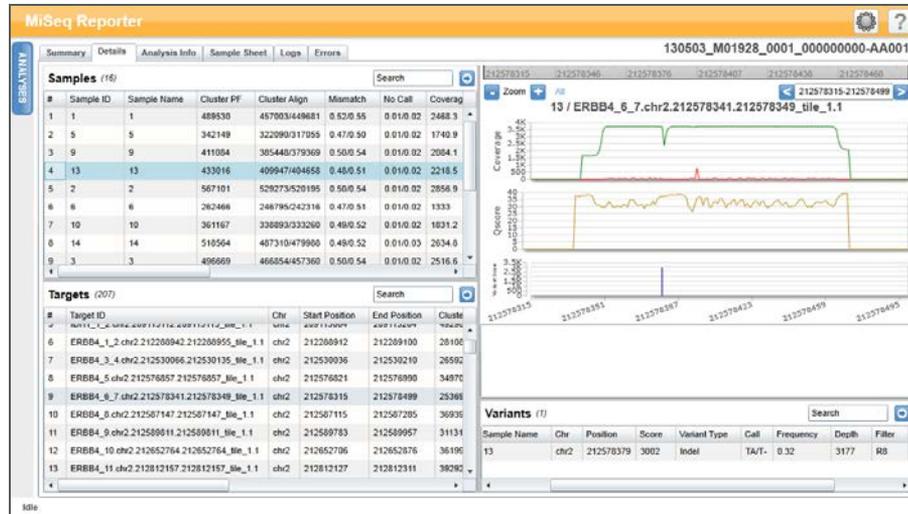
The screenshot shows the MiSeq Reporter Details page for sample 121017\_M00369\_0075\_A000000000-A1AH1. It displays two tables:

#	Sample ID	Sample Name	Call Rate	Performance	Control	Comment
1	NA01445	NA01445	100.00	Pass		
2	NA02828	NA02828	100.00	Pass		
3	NA04330	NA04330	100.00	Pass		
4	NA07339	NA07339	100.00	Pass		
5	NA07381	NA07381	100.00	Pass		
6	NA07441	NA07441	100.00	Pass		
7	NA07732	NA07732	100.00	Pass		
8	NA07854	NA07854	100.00	Pass		
9	NA07857	NA07857	100.00	Pass		

#	Sample ID	Sample Name	Mutation Name	Type	dbSNP rsID	CFTR Gene Region	Genomic Location	cDNA Name (HGVS)	Protein Name (HGVS)	Result	Col
1	NA01445	NA01445	PolyTG/PolyT	PolyTGPolyT	N/A	Intron 9	117188661-117188689	c.1210-12T[5_9]	N/A	(TG)10(T)7(TG)10(T)9	53
2	NA01445	NA01445	F595del	DEL	rs113893960	Exon 11	117199645	c.1521_1523delCTT	p.Phe508del	HET	301
3	NA01445	NA01445	W1282X	SNV	rs77010898	Exon 23	117282620	c.3846G>A	p.Tyr1282X	HET	301

Abbildung 8 Registerkarte „Details“ für das Universal Kit 1.0, Beispiel



Die Ergebnisse in den Tabellen „Samples“ (Proben), „Targets“ oder „Variants“ (Varianten) können durch Klicken auf das Symbol **Export table data to text file** (Tabellendaten in Textdatei exportieren) einzeln in eine Textdatei exportiert werden. Dieser Export ändert nicht die Analyseberichtsdatei.



Bei CF-Assays können Ergebnisse durch Klicken auf das Symbol **Export data to CF report** (Daten in den CF-Analysebericht exportieren) in den CF-Analysebericht exportiert werden. Weitere Informationen finden Sie unter *Analyse-Ausgabedateien für die CF-Assays* auf Seite 34.

## Registerkarte „Analysis Info“ (Informationen zur Analyse)

Die Registerkarte „Analysis Info“ (Informationen zur Analyse) enthält logistische Informationen über den Lauf und die Analyse.

Abbildung 9 Registerkarte „Analysis Info“ (Informationen zur Analyse)

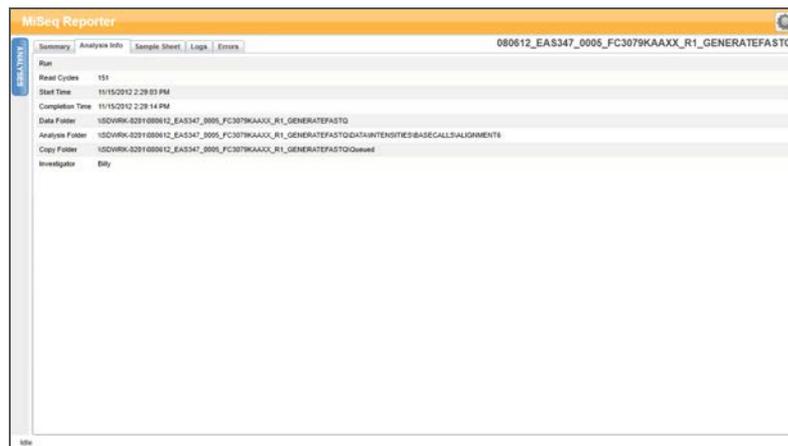


Tabelle 2 Registerkarte „Analysis Info“ (Informationen zur Analyse)

Zeile	Beschreibung
Read Cycles (Read-Zyklen)	Eine Darstellung der Anzahl der Zyklen bei jedem Read, einschließlich der Notation für alle Index-Reads. Wenn bei einem Lauf beispielsweise 151, 8 (I), 8 (I), 151 angegeben ist, bedeutet dies, dass der erste Read 151 Zyklen, der erste Index-Read 8 Zyklen, der zweite Index-Read 8 Zyklen und der letzte Read wieder 151 Zyklen umfasst.
Start Time (Gestartet um)	Die Uhrzeit, zu der die Sekundäranalyse gestartet wurde.
Completion Time (Abgeschlossen um)	Die Uhrzeit, zu der die Sekundäranalyse abgeschlossen wurde.
Data Folder (Datenordner)	Die Stammebene des von der RTA-Primäranalyse-Software angelegten Ausgabeordners (MiSeqOutput), der die Ausgabedaten aller Primär- und Sekundäranalysen für den Lauf enthält.
Analysis Folder (Analyseordner)	Der vollständige Pfad des Ordners „Alignment“ im MiSeqAnalysis-Ordner (Data \ Intensities \ BaseCalls \ Alignment).
Copy Folder (Kopien-Ordner)	Der vollständige Pfad des Queued-Unterordners im MiSeqAnalysis-Ordner.

### Registerkarte „Sample Sheet“ (Probenblatt)

Die Registerkarte „Sample Sheet“ (Probenblatt) enthält Laufparameter aus dem Probenblatt und stellt Tools zum Bearbeiten des Probenblatts und zum erneuten Platzieren des Laufs in der Warteschlange zur Verfügung.

Abbildung 10 Registerkarte „Sample Sheet“ (Probenblatt), Beispiel Universal Kit 1.0

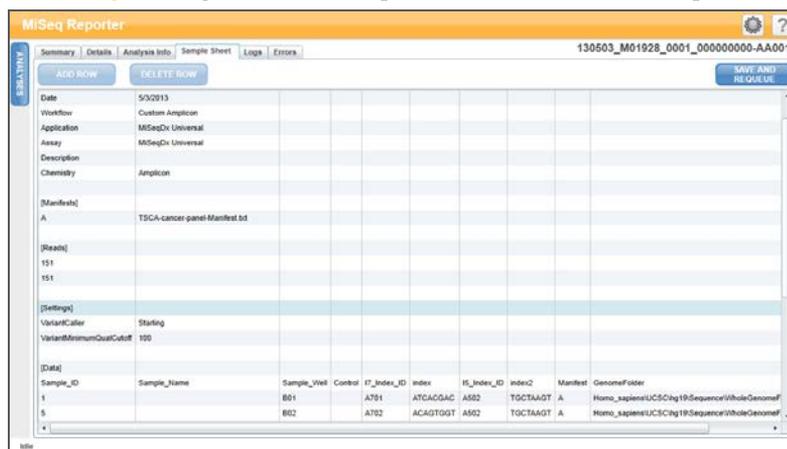


Tabelle 3 Inhalt der Registerkarte „Sample Sheet“ (Probenblatt)

Zeile	Beschreibung
Date (Datum)	Das Datum, an dem der Sequenzierungslauf durchgeführt wurde.

Tabelle 3 Inhalt der Registerkarte „Sample Sheet“ (Probenblatt)

Zeile	Beschreibung
Workflow	Der Analyse-Workflow für den Lauf. Für die CF Assays und das Universal Kit 1.0 lautet der Workflow-Name „Custom Amplicon“.
Application (Anwendung)	Der Anwendungsname. Dieses Feld wird von der Illumina Worklist Manager-Software verwendet und gibt an, welcher Assay oder welches Kit für den Lauf verwendet wird.
Assay	Der Name des Assays oder Kits.
Chemistry (Chemie)	Der Chemistry-Name gibt die Rezepturfragmente an, die zum Erstellen der laufspezifischen Rezeptur verwendet werden. Für MiSeqDx-Läufe lautet der Chemie-Name „Amplicon“.
Manifests (Manifestdateien)	Der Name der Manifestdatei, die ein Referenzgenom und Ziel-Referenzregionen angibt, die im Alignment-Schritt zu verwenden sind.
Reads	Die Anzahl der Zyklen, die in Read 1 und Read 2 durchgeführt werden. Index-Reads sind in diesem Abschnitt nicht enthalten.
Settings (Einstellungen)	Optionale Laufparameter.
Data (Daten)	Die Proben-ID, der Probenname, die Indexsequenzen und der Pfad zum Genomordner. Die Anforderungen variieren je nach Workflow.

### Registerkarte „Logs“ (Protokolle)

Auf der Registerkarte „Logs“ (Protokolle) sind alle Schritte aufgeführt, die während der Analyse durchgeführt werden. Diese Schritte werden in Protokolldateien aufgezeichnet, die sich im Protokollordner befinden. Eine Zusammenfassung wird in der Datei „AnalysisLog.txt“, einer wichtigen Datei für die Fehlerbehebung, festgehalten.

### Registerkarte „Errors“ (Fehler)

Auf der Registerkarte „Errors“ (Fehler) werden alle Fehler aufgeführt, die während der Analyse aufgetreten sind. Eine Zusammenfassung wird in der Datei „AnalysisError.txt“, einer wichtigen Datei für die Fehlerbehebung, festgehalten.

## Bearbeiten des Probenblatts in MiSeq Reporter

Probenblattdaten können von der Registerkarte „Probenblatt“ der Web-Oberfläche von MiSeq Reporter aus für einen bestimmten Lauf bearbeitet werden. Eine Maus und eine Tastatur sind zum Bearbeiten des Probenblatts erforderlich.



#### VORSICHT

Das Bearbeiten der Probenblattinformationen sollte mit äußerster Sorgfalt und Vorsicht erfolgen. Die Probenverfolgung kann verändert werden und möglicherweise zu falschen Ergebnisberichten führen.

- ▶ Um eine Zeile im Probenblatt zu bearbeiten, klicken Sie auf ein Feld und nehmen Sie die gewünschten Änderungen vor.

- ▶ Um eine Zeile zum Probenblatt hinzuzufügen, klicken Sie auf die Zeile und wählen Sie **Add Row** (Zeile hinzufügen). Die neue Zeile erscheint unterhalb der ausgewählten Zeile.

ADD ROW

- ▶ Um eine Zeile aus dem Probenblatt zu löschen, klicken Sie auf die Zeile und wählen Sie **Delete Row** (Zeile löschen).

DELETE ROW

- ▶ Nachdem Sie alle Änderungen am Probenblatt vorgenommen haben, wählen Sie **Save and Requeue** (Speichern und erneut in die Warteschlange stellen). Damit werden die Änderungen gespeichert und es wird eine Sekundäranalyse unter Verwendung des bearbeiteten Probenblatts gestartet.

SAVE AND  
REQUEUE

- ▶ Falls versehentlich eine nicht gewünschte Änderung vorgenommen wurde, klicken Sie auf eine benachbarte Registerkarte, bevor Sie die Änderungen speichern. Eine Warnmeldung erscheint, die besagt, dass die Änderungen nicht gespeichert wurden. Klicken Sie auf **Discard** (Verwerfen), um die Änderungen rückgängig zu machen.

DISCARD

## Speichern von Diagrammen als Bilder

MiSeq Reporter bietet eine Option zum Speichern eines Bilds der Diagramme, die für einen Lauf generiert wurden. Klicken Sie mit der rechten Maustaste in der Registerkarte „Summary“ (Zusammenfassung) auf eine beliebige Position oder in der Registerkarte „Details“ auf die Position der Diagramme und wählen Sie anschließend **Save Image As** (Bild speichern unter). Geben Sie der Datei bei Aufforderung einen Namen und navigieren Sie zu dem Speicherort, an dem die Datei gespeichert werden soll.

Alle Bilder werden im JPG-Format gespeichert. Alle auf der Registerkarte angezeigten Diagramme werden zusammen als einzelnes Diagramm exportiert. Zur Verwendung dieser Option ist eine Maus erforderlich.

## Analyse erneut in die Warteschlange stellen

Über die MiSeq Reporter-Web-Oberfläche ist es möglich, die Analyse erneut in die Warteschlange zu stellen. Überprüfen Sie vor dem Fortfahren, dass kein Sequenzierungslauf ausgeführt wird.

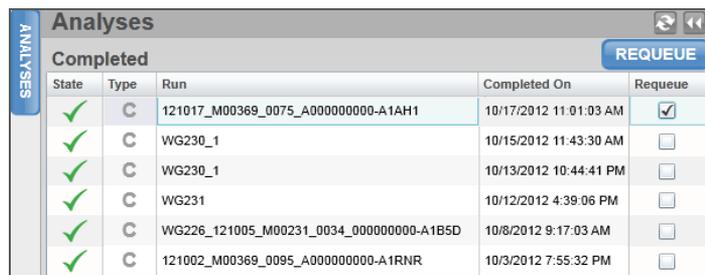
Jedes Mal, wenn die Analyse erneut in die Warteschlange gestellt wird, wird im MiSeqAnalysis-Ordner ein neuer Alignment-Ordner mit einer fortlaufenden Zahl erstellt, die an den Ordnernamen angehängt wird. Beispielsweise Alignment, Alignment1, Alignment2.

MiSeqAnalysis \<RunFolderName> \Data \Intensities \BaseCalls \Alignment2

- 1 Klicken Sie in der MiSeq Reporter-Web-Oberfläche auf **Analyses** (Analysen).
- 2 Navigieren Sie zu dem Lauf in der Liste der verfügbaren Läufe und klicken Sie neben dem Namen des Laufs auf das Kontrollkästchen „Requeue“ (Erneut in die Warteschlange stellen).

Wenn der Lauf nicht aufgeführt ist, ändern Sie das angegebene Repository in den korrekten Speicherort. Weitere Informationen finden Sie unter *Server-URL- oder Repository-Einstellungen* auf Seite 5.

Abbildung 11 Analyse erneut in die Warteschlange stellen



State	Type	Run	Completed On	Requeue
✓	C	121017_M00369_0075_A000000000-A1AH1	10/17/2012 11:01:03 AM	<input checked="" type="checkbox"/>
✓	C	WG230_1	10/15/2012 11:43:30 AM	<input type="checkbox"/>
✓	C	WG230_1	10/13/2012 10:44:41 PM	<input type="checkbox"/>
✓	C	WG231	10/12/2012 4:39:06 PM	<input type="checkbox"/>
✓	C	WG226_121005_M00231_0034_000000000-A1B5D	10/8/2012 9:17:03 AM	<input type="checkbox"/>
✓	C	121002_M00369_0095_A000000000-A1RNR	10/3/2012 7:55:32 PM	<input type="checkbox"/>

- 3 Klicken Sie auf **Requeue** (Erneut in die Warteschlange stellen). Das Statussymbol links neben dem Laufnamen ändert sich und zeigt an, dass die Analyse läuft 🔄.



### HINWEIS

Falls die Analyse nicht gestartet wird, stellen Sie sicher, dass die folgenden Eingabedateien im Analyselaufordner vorhanden sind: SampleSheet.csv, RTAComplete.txt und RunInfo.xml.

## Analysekennzahlen

Während des Sequenzierungslaufs generiert die Echtzeitanalyse (RTA) Datendateien, die Analysekennzahlen enthalten, die von MiSeq Reporter für die Sekundäranalyse verwendet werden. Kennzahlen, die in den Berichten der Sekundäranalyse enthalten sind, sind Cluster nach Filterung, Base-Call-Qualitäts-Scores sowie Phasierungs- und Vorphasierungswerte.

### Cluster nach Filterung

Cluster nach Filterung ist eine Messung der Clusterqualität. Dieser Filter entfernt die am wenigsten zuverlässigen Daten, indem die Rohdaten gefiltert werden, um alle Reads zu entfernen, die den Qualitätsansprüchen nicht genügen. Cluster nach Filterung werden in den Analyseberichten mit „PF“ bezeichnet.

### Qualitäts-Scores

Ein Qualitäts-Score oder Q-Score ist eine Prognose über die Wahrscheinlichkeit eines fehlerhaften Base-Calls. Während des Sequenzierungslaufs werden Base-Call-Qualitäts-Scores für jeden Zyklus aufgezeichnet. Während der Analyse werden Qualitäts-Scores in FASTQ-Dateien in einem ASCII-kodierten Format aufgezeichnet.

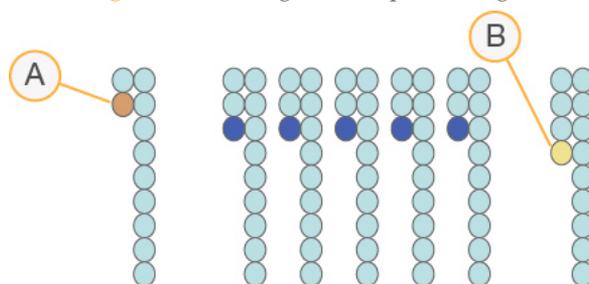
Die folgende Tabelle zeigt die Beziehung zwischen dem Qualitäts-Score und der Fehlerwahrscheinlichkeit.

Qualitäts-Score	Fehlerwahrscheinlichkeit
Q40	0,0001 (1 von 10.000)
Q30	0,001 (1 von 1.000)
Q20	0,01 (1 von 100)
Q10	0,1 (1 von 10)

### Phasierung und Vorphasierung

Während der Sequenzierungsreaktion erweitert sich jeder DNA-Strang in einem Cluster um eine Base pro Zyklus. Ein kleiner Anteil an Strängen kann mit dem aktuellen Inkorporationszyklus außer Phase geraten, entweder durch Zurückfallen hinter eine Base (Phasierung) oder durch Vorspringen um eine Base (Vorphasierung). Die Raten für die Phasierung und Vorphasierung sind eine Schätzung des Bruchteils der Moleküle, die in jedem Zyklus phasiert bzw. vorphasiert wurden.

Abbildung 12 Phasierung und Vorphasierung



- A Read mit einer phasierenden Base
- B Read mit einer vorphasierenden Base

Die Anzahl der in einem Read ausgeführten Zyklen ist um einen Zyklus höher als die Anzahl der analysierten Zyklen. Beispiel: Bei einem Paired-End-Lauf mit 150 Zyklen werden zwei Reads mit 151 Zyklen ( $2 \times 151$ ) ausgeführt, sodass sich eine Gesamtanzahl von 302 Zyklen ergibt. Am Ende des Laufs werden  $2 \times 150$  Zyklen analysiert. Der eine Extrazyklus für Read 1 und Read 2 wird für Vorphasierungsberechnungen benötigt.

## Analyseverfahren

MiSeq Reporter führt die Sekundäranalyse mithilfe einer Reihe von Analyseverfahren durch, zu denen die Demultiplexierung, die Generierung der FASTQ-Dateien, das Alignment und das Varianten-Calling gehören.

### Demultiplexierung

Die Demultiplexierung ist der erste Schritt der Analyse, wenn im Probenblatt mehrere Proben aufgelistet sind und der Lauf über Index-Reads verfügt.

Die Demultiplexierung trennt Daten von gepoolten Proben auf der Basis kurzer Indexsequenzen, die Proben aus verschiedenen Bibliotheken markieren. Jede Index-Read-Sequenz wird mit den im Probenblatt angegebenen Index-Sequenzen verglichen. In diesem Schritt werden keine Qualitätswerte berücksichtigt.

### Generieren von FASTQ-Dateien

Nach der Demultiplexierung werden mit diesem Verfahren temporäre Dateien im FASTQ-Dateiformat, dem Textformat für die Darstellung von Sequenzen, generiert. FASTQ-Dateien sind die primären Eingabedateien für den Alignment-Schritt. FASTQ-Dateien enthalten die Reads für jede Probe sowie die Qualitäts-Scores (ausgenommen Reads von Clustern, die den Filter nicht passiert haben).

### Alignment

Beim Alignment werden Sequenzen mit der Referenz verglichen, um eine Beziehung zwischen den Sequenzen zu identifizieren. Außerdem wird ein Score basierend auf Ähnlichkeitsregionen zugewiesen. Aligierte Reads werden in Dateien im BAM-Format gespeichert.

Für Daten, die auf MiSeq Reporter generiert werden, verwendet MiSeqDx einen beschränkten („banded“) Smith-Waterman-Algorithmus, der lokale Sequenz-Alignments durchführt, um ähnliche Regionen zwischen zwei Sequenzen festzustellen. Statt die gesamte Sequenz zu betrachten, vergleicht der Smith-Waterman-Algorithmus Segmente aller möglichen Längen. Lokale Alignments sind für unterschiedliche Sequenzen nützlich, bei denen vermutet wird, dass sie ähnliche Regionen innerhalb der größeren Sequenz enthalten.

### Varianten-Calling

Beim Varianten-Calling werden Einzelnukleotid-Polymorphismen (Single Nucleotide Polymorphisms, SNPs), Insertionen und Deletionen (Indels) und andere strukturelle Varianten aufgezeichnet.

Für die auf dem MiSeqDx-Gerät generierten Daten wird das Varianten-Calling vom Starling-Varianten-Caller in MiSeq Reporter durchgeführt. Starling führt Calls von SNPs und kleinen Indels durch und gibt die Tiefe sowie die Fehlerwahrscheinlichkeit für jede Stelle im Genom an. Bei jedem SNP- oder Indel-Call wird die Wahrscheinlichkeit eines Fehlers als ein Varianten-Qualitäts-Score angegeben.

Am Ende des Vorgangs erstellt Starling Berichte über SNPs und Indels im HTML-Format sowie tabulatorgetrennte Textdateien, die Varianten im Varianten-Call-Format (VCF) enthalten. Weitere Informationen finden Sie unter *VCF-Dateiformat* auf Seite 48.

## MiSeqAnalysis-Ordner

Der MiSeqAnalysis-Ordner ist der Hauptlaufordner für MiSeq Reporter. Die Beziehung zwischen dem MiSeqOutput- und MiSeqAnalysis-Laufordner wird wie folgt zusammengefasst:

- ▶ Während der Sequenzierung stellt die Echtzeitanalyse (RTA) die während der Primäranalyse generierten Dateien in den MiSeqOutput-Ordner.
- ▶ Abgesehen von Fokus- und Miniaturbildern stellt die Echtzeitanalyse (RTA) Kopien der Dateien in Echtzeit in den MiSeqAnalysis-Ordner. Nach Abschluss der Primäranalyse legt die Echtzeitanalyse (RTA) die Datei RTAComplete.xml in beiden Laufordnern an.
- ▶ MiSeq Reporter überwacht den MiSeqAnalysis-Ordner und startet die Sekundäranalyse, sobald die Datei RTAComplete.xml vorhanden ist.
- ▶ Noch während der Ausführung der Sekundäranalyse erstellt MiSeq Reporter Analyse-Ausgabedateien im MiSeqAnalysis-Ordner und kopiert sie anschließend in den MiSeqOutput-Ordner.

[Diese Seite wurde absichtlich leer gelassen.]

# Datendarstellung

Einleitung .....	20
Anforderungen für die Eingabedatei .....	21
Custom Amplicon-Workflow .....	22
Analyse-Ausgabedateien für die CF-Assays .....	34



## Einleitung

MiSeq Reporter führt eine Sekundäranalyse durch und generiert verschiedene Informationstypen, die nach Abschluss der Analyse spezifisch für den Assay sind. Ergebnisse werden für jeden Lauf in der MiSeq Reporter-Web-Oberfläche in Form von Diagrammen und Tabellen angezeigt. MiSeqDx-Produkte umfassen die in der Tabelle unten aufgelisteten Produkte:

Produkt	Beschreibung
139-Varianten-Assay für zystische Fibrose	Erkennt 139 klinisch relevante Varianten im CFTR-Gen bei einem Maximum von 48 Proben.
Klinischer Sequenzierungs-Assay für zystische Fibrose	Erkennt Mutationen in den Protein-kodierenden Regionen einschließlich Intron/Exon-Grenzen, zwei großen Deletionen und zweier tiefer intronischer Mutationen im CFTR-Gen bei einem Maximum von acht Proben.
Universal Kit 1.0	Satz an Reagenzien und Verbrauchsmaterialien, die zusammen mit einem vom Benutzer bereitzustellenden anwendungsspezifischen Oligo verwendet werden, um eine gezielte Resequenzierung spezifischer genomischer Regionen von Interesse durchzuführen.

## Anforderungen für die Eingabedatei

MiSeq Reporter benötigt die folgenden Primäranalysedateien, die während des Sequenzierungslaufs generiert werden, um eine Sekundäranalyse durchzuführen oder die Analyse erneut in die Warteschlange zu stellen. Primäranalysedateien, z. B. \*.bcl, \*.filter und \*.locs, werden zum Durchführen der Analyse benötigt.

Es ist nicht erforderlich, Dateien vor dem Beginn der Analyse an einen anderen Speicherort zu verschieben oder zu kopieren. Erforderliche Dateien werden während des Sequenzierungsprozesses automatisch in den MiSeqAnalysis-Ordner kopiert.

Dateiname	Beschreibung
RTAComplete.txt	Eine Markerdatei, die angibt, dass die RTA-Verarbeitung abgeschlossen ist. Das Vorhandensein dieser Datei veranlasst MiSeq Reporter dazu, die Analyse in die Warteschlange zu stellen.
SampleSheet.csv	Bietet Parameter für den Lauf und die nachfolgende Analyse. Zu Beginn des Laufs wird das Probenblatt in den Stammordner des Laufordners kopiert und in SampleSheet.csv umbenannt.
RunInfo.xml	Enthält Laufinformationen der höchsten Ebene, z. B. die Anzahl der Reads und Zyklen im Sequenzierungslauf und die Angabe, ob ein Read indiziert ist.

## Vorinstallierte Datenbanken und Genome

MiSeqDx umfasst vorinstallierte Datenbanken und Genome.

Vorinstalliert	Beschreibung
Datenbanken	dbSNP für Menschen, Version 131 refGene für Menschen
Genome	Mensch ( <i>Homo sapiens</i> ) Build hg19

## Custom Amplicon-Workflow

Der für CF-Assays und Universal Kit 1.0 verwendete Custom Amplicon-Workflow wertet kurze Regionen amplifizierter DNA oder Amplikons bei Varianten aus. Das fokussierte Sequenzieren von Amplikons ermöglicht eine hohe Abdeckung bestimmter Regionen über eine große Anzahl von Proben hinweg.

Nach der Demultiplexierung und der Generierung der FASTQ-Dateien werden im Rahmen des Workflows die folgenden Schritte durchgeführt:

- ▶ **Alignment** – Cluster von jeder Probe werden an den in der Manifestdatei angegebenen Amplikon-Sequenzen ausgerichtet.
  - Im Falle von Paired-End-Daten wird jeder Read zunächst in Bezug auf sein Alignment mit den relevanten Probensequenzen für diesen Read ausgewertet. Read 1 wird anhand des umgekehrten Komplements der stromabwärts gelegenen lokusspezifischen Oligos (DLSO, downstream locus specific oligos) und Read 2 anhand der stromaufwärts gelegenen lokusspezifischen Oligos (ULSO, upstream locus specific oligos) ausgewertet. Wenn der Beginn einer Read-Sequenz einer Probensequenz mit maximal einer Nichtübereinstimmung entspricht, wird die volle Länge des Reads an der Amplikon-Zielsequenz für diese Probensequenz ausgerichtet. Dieses Alignment wird mithilfe eines „banded“ Smith-Waterman-Alignments längs der Amplikon-Zielsequenzen durchgeführt.
  - Indels innerhalb der DLSO und ULSO werden aufgrund der Assay-Chemie nicht beobachtet.
- ▶ **Paired-End-Auswertung** – Im Falle von Paired-End-Läufen wird das Alignment mit dem höchsten Score für jeden Read berücksichtigt. Falls das Alignment bei einem der beiden Reads fehlgeschlagen oder mit anderen Chromosomen erfolgt ist, werden die Reads als ungelöstes Paar markiert. Wenn darüber hinaus die beiden Alignments aus unterschiedlichen Amplikonen stammen (d. h., unterschiedliche Reihen im Abschnitt „Targets“ der Manifestdatei), werden die Reads als ungelöstes Paar markiert.
- ▶ **Behälter/Sortieren** – Reads werden nach Probe und Chromosom gruppiert und anschließend nach Chromosomposition sortiert. Die Ergebnisse werden in einer BAM-Datei pro Probe festgehalten.
- ▶ **Varianten-Calling** – Mutationen werden vom Varianten-Caller identifiziert. Weitere Informationen hierzu finden Sie unter *Varianten-Calling* auf Seite 16.
- ▶ **Varianteanalyse und Annotation** – Unter Verwendung einer vorinstallierten SNP-Datenbank (dbsnp.txt) werden alle bekannten Mutationen in der Analysenberichtsdatei markiert.
- ▶ **Statistikberichte** – Statistiken werden zusammengefasst und entsprechende Berichte generiert.

### Registerkarte „Summary“ (Zusammenfassung)

Die Informationen, die auf der Registerkarte „Summary“ (Zusammenfassung) angezeigt werden, umfassen ein Diagramm mit niedrigen Prozentwerten, ein Diagramm mit hohen Prozentwerten, ein Clusterdiagramm und ein Diagramm mit Nichtübereinstimmungen.

Abbildung 13 Registerkarte „Summary“ (Zusammenfassung), Beispiel



Diagramm „Low Percentages“ (Niedrige Prozentwerte)

Y-Achse	X-Achse	Beschreibung
Percent (Prozent)	Phasing 1 (Phasierung 1)	Der prozentuale Anteil von Molekülen in einem Cluster, die hinter dem aktuellen Zyklus innerhalb von Read 1 zurückbleiben.
	Phasing 2 (Phasierung 2)	Der prozentuale Anteil von Molekülen in einem Cluster, die hinter dem aktuellen Zyklus innerhalb von Read 2 zurückbleiben.
	PrePhasing 1 (Vorphasierung 1)	Der prozentuale Anteil von Molekülen in einem Cluster, die dem aktuellen Zyklus innerhalb von Read 1 vorauslaufen.
	PrePhasing 2 (Vorphasierung 2)	Der prozentuale Anteil von Molekülen in einem Cluster, die dem aktuellen Zyklus innerhalb von Read 2 vorauslaufen.
	Mismatch 1 (Nichtübereinstimmung 1)	Der durchschnittliche prozentuale Anteil von Nichtübereinstimmungen für Read 1 über alle Zyklen hinweg.
	Mismatch 2 (Nichtübereinstimmung 2)	Der durchschnittliche prozentuale Anteil von Nichtübereinstimmungen für Read 2 über alle Zyklen hinweg.

Diagramm „High Percentages“ (Hohe Prozentwerte)

Y-Achse	X-Achse	Beschreibung
Percent (Prozent)	PF	Der prozentuale Anteil von Clustern nach Filterung.
	Align 1 (Ausrichtung 1)	Der Prozentsatz der Cluster, die auf die Referenz in Read 1 ausgerichtet wurden.
	Align 2 (Ausrichtung 2)	Der Prozentsatz der Cluster, die auf die Referenz in Read 2 ausgerichtet wurden.
	I20 / I1 1	Das Verhältnis der Intensitäten bei Zyklus 20 zu den Intensitäten bei Zyklus 1 für Read 1.
	I20 / I1 2	Das Verhältnis der Intensitäten bei Zyklus 20 zu den Intensitäten bei Zyklus 1 für Read 2.
	PE Resynthesis (PE-Resynthese)	Das Verhältnis der Erstzyklenintensitäten für Read 1 zu den Erstzyklenintensitäten für Read 2.

## Diagramm „Clusters“ (Cluster)

Y-Achse	X-Achse	Beschreibung
Clusters (Cluster)	Raw (Roh)	Die Gesamtzahl der im Lauf erkannten Cluster.
	PF	Die Gesamtzahl der Cluster nach Filterung in diesem Lauf.
	Unaligned (Nicht ausgerichtet)	Die Gesamtzahl der Cluster nach Filterung, die ggf. kein Alignment mit dem Referenzgenom aufweisen. Cluster, die nicht indiziert sind, werden nicht in die Anzahl der nicht ausgerichteten Cluster aufgenommen.
	Unindexed (Nicht indiziert)	Die Gesamtzahl der Cluster nach Filterung, die keiner Indexsequenz im Lauf zugeordnet waren.
	Duplicate (Doppelt)	Dieser Wert gilt nicht für CF-Assays oder das Universal Kit 1.0 und wird daher immer 0 sein.

## Diagramm „Mismatch“ (Nichtübereinstimmung)

Y-Achse	X-Achse	Beschreibung
Percent (Prozent)	Cycle (Zyklus)	Stellt den prozentualen Anteil der Nichtübereinstimmungen für alle Cluster in einem Lauf nach Zyklen dar.

## Registerkarte „Details“ für den CF 139-Varianten-Assay

Die auf der Registerkarte „Details“ für den CF 139-Varianten-Assay angezeigten Informationen umfassen eine Probentabelle und eine Variantentabelle.

Abbildung 14 Registerkarte „Details“ für den CF 139-Varianten-Assay, Beispiel

**MiSeq Reporter** 121017\_M00369\_0075\_A000000000-A1AH1

Summary Details Analysis Info Sample Sheet Logs Errors

**Samples (48)**

#	Sample ID	Sample Name	Call Rate	Performance	Control	Comment
1	NA01445	NA01445	100.00	Pass		
2	NA02828	NA02828	100.00	Pass		
3	NA04330	NA04330	100.00	Pass		
4	NA07339	NA07339	100.00	Pass		
5	NA07381	NA07381	100.00	Pass		
6	NA07441	NA07441	100.00	Pass		
7	NA07732	NA07732	100.00	Pass		
8	NA07854	NA07854	100.00	Pass		
9	NA07857	NA07857	100.00	Pass		

**Variants (3)**

#	Sample ID	Sample Name	Mutation Name	Type	dbSNP rsID	CFTR Gene Region	Genomic Location	cDNA Name (HGVS)	Protein Name (HGVS)	Result	Cov
1	NA01445	NA01445	PolyTG/PolyT	Poly/TG/PolyT	N/A	Intron 9	117188661-117188689	c.1210-12T[S_S]	N/A	(TG)10(TT)7(TG)10(TT)9	53
2	NA01445	NA01445	F508del	DEL	rs113993860	Exon 11	117199645	c.1521_1523delCTT	p.Phe508del	HET	301
3	NA01445	NA01445	W1282X	SNV	rs77010898	Exon 23	117282820	c.3846G>A	p.Trp1282X	HET	301

## Probentabelle für den CF 139-Varianten-Assay

Spalte	Beschreibung
#	Eine Ordinalzahl in der Tabelle, die der Identifikation dient.

Spalte	Beschreibung
Sample ID (Proben-ID)	Die Proben-ID aus dem Probenblatt. Die Proben-ID muss immer ein eindeutiger Wert sein.
Sample Name (Probenname)	Der Probenname aus dem Probenblatt.
Call Rate (Call-Rate)	Die Anzahl der Mutationspositionen, die den vordefinierten Schwellenwert des Zuverlässigkeitswerts erfüllen, geteilt durch die Gesamtzahl der analysierten Mutationspositionen. Die Call-Rate wird auf einer Pro-Proben-Basis und als Prozentsatz angegeben, der wie folgt berechnet wird: 1 minus [Anzahl der Positionen mit unvollständigen Calls geteilt durch die Gesamtzahl der sequenzierten Positionen].
Performance (Leistung)	„Pass“- oder „Fail“-Bewertung auf Basis der Call-Rate. Für eine positive Kontrollprobe: <ul style="list-style-type: none"> <li>• PASS – mit einer Call-Rate <math>\geq 99</math> %</li> <li>• FAIL – mit einer Call-Rate <math>&lt; 99</math> %</li> </ul> Für eine negative Kontrollprobe: <ul style="list-style-type: none"> <li>• PASS – mit einer Call-Rate <math>\leq 10</math> %</li> <li>• FAIL – mit einer Call-Rate <math>&gt; 10</math> %</li> </ul> Für eine Probe, die nicht als positive oder negative Kontrolle gekennzeichnet ist: <ul style="list-style-type: none"> <li>• PASS – mit einer Call-Rate <math>\geq 99</math> %</li> <li>• FAIL – mit einer Call-Rate <math>&lt; 99</math> %</li> </ul>
Control (Kontrolle)	Die Art der Kontrolle, wie im Probenblatt aufgelistet. Die Werte sind positiv oder negativ. Ein leeres Feld bedeutet, dass es sich um eine reine Probe handelt.
Comment (Anmerkung)	Ein optionales Textfeld für Kommentare. In dieses Feld eingegebene Kommentare werden in der Analyseberichtsdatei MiSeqDxCf139VariantAssay.txt gespeichert. Wenn die Analyse erneut in die Warteschlange gestellt wird, wird eine neue Berichtsdatei erstellt. Kommentare aus dem vorherigen Analyselauf werden nicht in den nächsten Analyselauf übernommen.

### Variantentabelle für den CF 139-Varianten-Assay

Spalte	Beschreibung
#	Eine Ordinalzahl in der Tabelle, die der Identifikation dient.
Sample ID (Proben-ID)	Die Proben-ID aus dem Probenblatt. Die Proben-ID muss immer ein eindeutiger Wert sein.
Sample Name (Probenname)	Der Probenname aus dem Probenblatt.

Spalte	Beschreibung
Mutations (Common Name) (Mutationen (Allgemeiner Name))	Allgemeiner Name der Mukoviszidosevariante, wie in der CFTR2-Datenbank beschrieben.
Mutation Type (Mutationstyp)	Der Typ der Variante. <ul style="list-style-type: none"> <li>• SNV – Single Nucleotide Variant (einzelne Nukleotidvariante)</li> <li>• DIV – Deletion Insertion Variant (Deletions-/Insertionsvariante)</li> <li>• DEL – Große Deletion</li> <li>• PolyTGPolyT – PolyTG/PolyT-Genotyp im ZF-Gen</li> </ul>
dbSNP rsID	dbSNP rsID der Variante, sofern zutreffend.
CFTR Gene Region (CFTR-Genregion)	CFTR-Genregion (Exon-Nummer oder Intron-Nummer), in der die Variante vorhanden ist.
Genomic Location (Genomischer Ort)	Genomischer Ort der Variante.
cDNA Name (cDNA-Name) (HGVS)	Beschreibung einer Variante auf DNA-Ebene, die die coding DNA (cDNA)-Sequenz-Nomenklatur verwendet, die von der Human Genome Variation Society (HGVS) empfohlen wird.
Protein Name (Protein-Name) (HGVS)	Beschreibung einer Variante auf Proteinebene, die die Proteinsequenz-Nomenklatur verwendet, die von der Human Genome Variation Society (HGVS) empfohlen wird.
Result (Ergebnis)	Variant-Genotyp. Für SNVs, DIVs und DELs: <ul style="list-style-type: none"> <li>• HET – Heterozygot</li> <li>• HOM – Homozygot</li> </ul> Für die PolyTGPolyT-Variante wird der tatsächliche Genotyp gemeldet. <b>HINWEIS:</b> PolyTGPolyT wird nur gemeldet, wenn die R117H-Variante entdeckt wird.

## Registerkarte „Details“ für den klinischen CF-Sequenzierungs-Assay

Die auf der Registerkarte „Details“ für den klinischen CF-Sequenzierungs-Assay angezeigten Informationen umfassen eine Proben-tabelle, eine Variantentabelle, ein Abdeckungsdiagramm, ein Qscore-Diagramm und ein Diagramm mit Varianten-Score.

Abbildung 15 Registerkarte „Details“ für den klinischen CF-Sequenzierungs-Assay, Beispiel

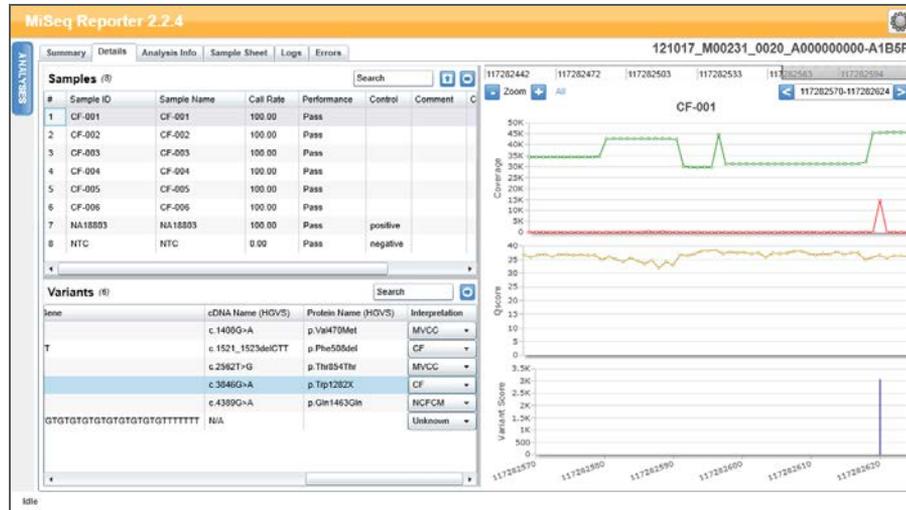


Tabelle „Samples“ (Proben) für den klinischen CF-Sequenzierungs-Assay

Spalte	Beschreibung
#	Eine Ordinalzahl in der Tabelle, die der Identifikation dient.
Sample ID (Proben-ID)	Die Proben-ID aus dem Probenblatt. Die Proben-ID muss immer ein eindeutiger Wert sein.
Sample Name (Probenname)	Der Probenname aus dem Probenblatt.
Call Rate (Call-Rate)	Die Anzahl der Basen, die den Schwellenwert für den Qualitäts-Score erfüllen, geteilt durch die Gesamtzahl der analysierten Basen. Die Call-Rate wird auf einer Pro-Proben-Basis und als Prozentsatz angegeben, der wie folgt berechnet wird: 1 minus [Anzahl der Positionen mit unvollständigen Calls geteilt durch die Gesamtzahl der sequenzierten Basen/Positionen].
Performance (Leistung)	„Pass“- oder „Fail“-Bewertung auf Basis der Call-Rate. Für eine positive Kontrollprobe: <ul style="list-style-type: none"> <li>• PASS – mit einer Call-Rate <math>\geq 99</math> %</li> <li>• FAIL – mit einer Call-Rate <math>&lt; 99</math> %</li> </ul> Für eine negative Kontrollprobe: <ul style="list-style-type: none"> <li>• PASS – mit einer Call-Rate <math>\leq 10</math> %</li> <li>• FAIL – mit einer Call-Rate <math>&gt; 10</math> %</li> </ul> Für eine Probe, die nicht als positive oder negative Kontrolle gekennzeichnet ist: <ul style="list-style-type: none"> <li>• PASS – mit einer Call-Rate <math>\geq 99</math> %</li> <li>• FAIL – mit einer Call-Rate <math>&lt; 99</math> %</li> </ul>

Spalte	Beschreibung
Control (Kontrolle)	Die Art der Kontrolle, wie im Probenblatt aufgelistet. Die Werte sind positiv oder negativ. Ein leeres Feld bedeutet, dass es sich um eine reine Probe handelt.
Comment (Anmerkung)	Ein optionales Textfeld für Kommentare. In dieses Feld eingegebene Kommentare werden in der Analyseberichtsdatei MiSeqDxCFClinicalSequencing.txt gespeichert. Wenn die Analyse erneut in die Warteschlange gestellt wird, wird eine neue Berichtsdatei erstellt. Kommentare aus dem vorherigen Analyselauf werden nicht in den nächsten Analyselauf übernommen.
Coordinates Not Called (Koordinaten nicht aufgerufen)	Genomkoordinaten innerhalb der Zielregion, in der ein Call aufgrund niedriger Zuverlässigkeitswerte nicht gemeldet wurde.

### Variantentabelle für den klinischen CF-Sequenzierungs-Assay

Spalte	Beschreibung
#	Eine Ordinalzahl in der Tabelle, die der Identifikation dient.
Sample ID (Proben-ID)	Die Proben-ID aus dem Probenblatt. Die Proben-ID muss immer ein eindeutiger Wert sein.
Sample Name (Probenname)	Der Probenname aus dem Probenblatt.
Chr	Das Referenzziel oder der Name des Chromosoms.
Position	Die Position, an der die Variante gefunden wurde.
Variant Type (Variantentyp)	Der Typ der Variante. <ul style="list-style-type: none"> <li>• SNV – Single Nucleotide Variant (einzelne Nukleotidvariante)</li> <li>• DIV – Deletion Insertion Variant (Deletions-/Insertionsvariante)</li> <li>• DEL – Große Deletion</li> <li>• PolyTGPolyT – PolyTG/PolyT-Genotyp im ZF-Gen</li> </ul>
Call	Eine Zeichenfolge, die angibt, wie sich die Base oder die Basen an dieser Stelle in der Referenz geändert hat bzw. haben.
Frequency (Frequenz)	Der Teil der Reads für die Probe, der die Variante enthält. Wenn beispielsweise die Referenzbasis an einer bestimmten Position A ist und Probe 1 60 A-Reads und 40 T-Reads hat, dann hat SNV eine Variantenhäufigkeit von 0,4.
Depth (Tiefe)	Die Anzahl der Reads für eine Probe, die eine bestimmte Position abdeckt.
Filter	Die Kriterien für eine gefilterte Variante.
dbSNP ID	Der dbSNP-Name der Variante.

Spalte	Beschreibung
RefGene	Das Gen gemäß RefGene, in dem diese Variante vorkommt.
cDNA Name (cDNA-Name) (HGVS)	Beschreibung einer Variante auf DNA-Ebene, die die coding DNA (cDNA)-Sequenz-Nomenklatur verwendet, die von der Human Genome Variation Society (HGVS) empfohlen wird.
Protein Name (Protein-Name) (HGVS)	Beschreibung einer Variante auf Proteinebene, die die Proteinsequenz-Nomenklatur verwendet, die von der Human Genome Variation Society (HGVS) empfohlen wird.
Interpretation	<p>Dieses Feld ermöglicht es dem Medizingenetiker, eine klinische Interpretation der Mutation für jede Probe anzubieten. Die folgenden Optionen sind in der Dropdown-Liste für jede Probe verfügbar:</p> <ul style="list-style-type: none"> <li>• CF – CF causing (ZF verursachend)</li> <li>• MVCC – Mutation of Varying Clinical Consequence (Mutation mit variierender klinischer Konsequenz)</li> <li>• MOUS – Mutation of Unknown Significance (Mutation mit unbekannter Signifikanz)</li> <li>• NCFCM – Non CF Causing Mutation (Mutation, die nicht ZF verursacht)</li> <li>• Unknown (Unbekannt)</li> </ul> <p>Anhand des Symbols kann ein neuer Bericht generiert werden.</p>

### Spalte „Interpretation“ der Varianten-Tabelle

Die Spalte „Interpretation“ bietet eine Auswahl, die es dem Medizingenetiker ermöglicht, eine klinische Interpretation der Mutation für jede Probe anzubieten. Die folgenden Optionen sind in der Dropdown-Liste „Interpretation“ verfügbar:

- **CF** – CF causing (ZF verursachend)
- **MVCC** – Mutation of Varying Clinical Consequence (Mutation mit variierender klinischer Konsequenz)
- **MOUS** – Mutation of Unknown Significance (Mutation mit unbekannter Signifikanz)
- **NCFCM** – Non CF Causing Mutation (Mutation, die nicht ZF verursacht)
- **Unknown** (Unbekannt)

Abbildung 16 Spalte „Interpretation“



Die Ergebnisse in den Varianten-Tabellen können durch Klicken auf das Symbol **Export table data to text file** (Tabellendaten in Textdatei exportieren) einzeln in eine Textdatei exportiert werden. Dieser Export ändert nicht die Analyseberichtsdatei.



Nachdem der Medizingenetiker die Variantensignifikanz ermittelt hat, können die Interpretationseinstellungen im Analysebericht gespeichert werden. Der Dateiname des ursprünglichen Analyseberichts wird automatisch mit einem Zeit-/Datumsstempel versehen.

## Diagramm „Coverage“ (Abdeckung) für den klinischen CF-Sequenzierungs-Assay

Y-Achse	X-Achse	Beschreibung
Coverage (Abdeckung)	Position	Die grüne Kurve stellt die Anzahl der ausgerichteten Reads dar, die jede Position in der Referenz abdecken. Die rote Kurve stellt die Anzahl der ausgerichteten Reads dar, die an dieser Position in der Referenz einen Miscall aufweisen. SNVs und andere Varianten werden als Spitzen in der roten Kurve dargestellt.

## Diagramm „Qscore“

Y-Achse	X-Achse	Beschreibung
Qscore	Position	Der durchschnittliche Qualitäts-Score der Basen an der gegebenen Position der Referenz.

## Varianten-Score-Diagramm für den klinischen CF-Sequenzierungs-Assay

Y-Achse	X-Achse	Beschreibung
Score	Position	Stellt grafisch den Qualitäts-Score und die Position von SNVs und Indels dar.

## Registerkarte „Details“ für das Universal Kit 1.0

Die auf der Registerkarte „Details“ für das Universal Kit 1.0 angezeigten Informationen umfassen eine Probentabelle, eine Targets-Tabelle, ein Abdeckungsdiagramm, ein Qscore-Diagramm, ein Varianten-Score-Diagramm und eine Variantentabelle.

Abbildung 17 Registerkarte „Details“ für das Universal Kit 1.0, Beispiel

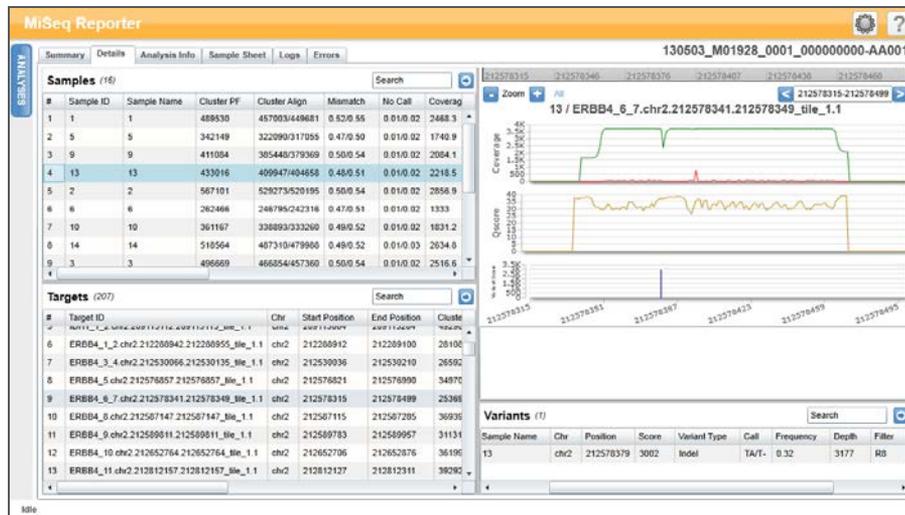


Tabelle „Samples“ (Proben) für das Universal Kit 1.0

Spalte	Beschreibung
#	Eine Ordinalzahl in der Tabelle, die der Identifikation dient.
Sample ID (Proben-ID)	Die Proben-ID aus dem Probenblatt. Die Proben-ID muss immer ein eindeutiger Wert sein.
Sample Name (Probenname)	Der Probenname aus dem Probenblatt.
Cluster PF (Cluster nach Filterung)	Die Anzahl der Cluster nach Filterung in der Probe.
Cluster Align (Cluster-Ausrichtung)	Die Gesamtzahl der ausgerichteten PF-Cluster der Probe (Read 1/Read 2).
Mismatch (Nichtübereinstimmung)	Die prozentuale Nichtübereinstimmung mit der Referenz, gemittelt über Zyklen pro Read (Read 1/Read 2).
No Call (Kein Call)	Der Prozentsatz an Basen der Probe, die nicht aufgerufen werden konnten (No-Call), gemittelt über Zyklen pro Read (Read 1/Read 2).
Coverage (Abdeckung)	Die mittlere Abdeckung (Anzahl an Basen, die an einer bestimmten Referenzposition ausgerichtet sind), gemittelt über alle Positionen.
Het SNPs (Het-SNPs)	Die Anzahl der entdeckten heterozygoten SNPs in der Probe.
Hom SNPs (Hom-SNPs)	Die Anzahl der entdeckten homozygoten SNPs in der Probe.
Insertions (Insertionen)	Die Anzahl der entdeckten Insertionen in der Probe.
Deletions (Deletionen)	Die Anzahl der entdeckten Deletionen in der Probe.
Manifest	Die Datei, die ein Referenzgenom und Ziel-Referenzregionen angibt, die im Alignment-Schritt zu verwenden sind.
Genome (Genom)	Der Name des Referenzgenoms.

Tabelle „Targets“ für das Universal Kit 1.0

Spalte	Beschreibung
#	Eine Ordinalzahl in der Tabelle, die der Identifikation dient.
Target ID (Target-ID)	Der Name des Targets im Manifest.
Chr	Das Referenzziel oder der Name des Chromosoms.
Start Position (Startposition)	Die Startposition der Target-Region.
End Position (Endposition)	Die Endposition der Target-Region.

Spalte	Beschreibung
Cluster PF (Cluster nach Filterung)	Die Anzahl der Cluster nach Filterung für das Target, angezeigt pro Read (Read 1/Read 2).
Mismatch (Nichtübereinstimmung)	Der Prozentsatz der nicht mit dem Target übereinstimmenden Basen, gemittelt über alle Zyklen, angezeigt pro Read. Nichtübereinstimmung = [mittlere (Fehleranzahl in Zyklen) / Cluster PF] * 100.
No Call (Kein Call)	Der Prozentsatz der No-Call-Basen für das Target, gemittelt über Zyklen, angezeigt pro Read.
Het SNPs (Het-SNPs)	Die Anzahl der für das Target entdeckten heterozygoten SNPs in allen Proben.
Hom SNPs (Hom-SNPs)	Die Anzahl der für das Target entdeckten homozygoten SNPs in allen Proben.
Insertions (Insertionen)	Die Anzahl der für das Target erkannten Insertionen in allen Proben.
Deletions (Deletionen)	Die Anzahl der für das Target erkannten Deletionen in allen Proben.
Manifest	Die Datei, die ein Referenzgenom und Ziel-Referenzregionen angibt, die im Alignment-Schritt zu verwenden sind.

### Diagramm „Coverage“ (Abdeckung) für das Universal Kit 1.0

Y-Achse	X-Achse	Beschreibung
Coverage (Abdeckung)	Position	Die grüne Kurve stellt die Anzahl der ausgerichteten Reads dar, die jede Position in der Referenz abdecken. Die rote Kurve stellt die Anzahl der ausgerichteten Reads dar, die an dieser Position in der Referenz einen Miscall aufweisen. SNPs und andere Varianten werden als Spitzen in der roten Kurve dargestellt.

### Diagramm „Qscore“

Y-Achse	X-Achse	Beschreibung
Qscore	Position	Der durchschnittliche Qualitäts-Score der Basen an der gegebenen Position der Referenz.

### Diagramm „Variant Score“ (Variant-Score) für das Universal Kit 1.0

Y-Achse	X-Achse	Beschreibung
Score	Position	Stellt grafisch den Varianten-Qualitäts-Score und die Position von SNPs und Indels dar.

Tabelle „Variants“ (Varianten) für das Universal Kit 1.0

Spalte	Beschreibung
#	Eine Ordinalzahl in der Tabelle, die der Identifikation dient.
Sample ID (Proben-ID)	Die Proben-ID aus dem Probenblatt. Die Proben-ID muss immer ein eindeutiger Wert sein.
Sample Name (Probenname)	Der Probenname aus dem Probenblatt.
Chr	Das Referenzziel oder der Name des Chromosoms.
Position	Die Position, an der die Variante gefunden wurde.
Score	Der Varianten-Qualitäts-Score für diese Variante.
Variant Type (Variantentyp)	Der Variantentyp, entweder SNP oder Indel.
Call	<p>Eine Darstellung davon, wie sich die Base oder die Basen an dieser Stelle in der Referenz geändert hat bzw. haben.</p> <ul style="list-style-type: none"> <li>• SNPs werden im Format „Referenz &gt; AllelA/AllelB“ aufgeführt.</li> <li>• Insertionen werden im Format „Referenz/Insertion“ aufgeführt. „G-/GA“ gibt die Insertion von A an.</li> <li>• Deletionen werden im Format „Referenz/Deletion“ aufgeführt. „AGG/A--“ gibt die Deletion von GG an.</li> </ul>
Frequency (Frequenz)	Der Teil der Reads für die Probe, der die Variante enthält. Wenn beispielsweise die Referenzbasis A ist und Probe 1 60 A-Reads und 40 T-Reads hat, dann hat der SNP eine Variantenhäufigkeit von 0,4.
Depth (Tiefe)	Die Anzahl der Reads für eine Probe, die eine bestimmte Position abdeckt.
Filter (Nach Filterung)	Die Kriterien für eine gefilterte Variante. Wenn alle Filter passiert werden, wird „PASS“ in die Filterspalte geschrieben. Weitere Informationen finden Sie unter <i>Überschriften und Anmerkungen in der VCF-Datei</i> auf Seite 49.
dbSNP	Der dbSNP-Name der Variante, sofern zutreffend.
RefGene	Das Gen gemäß RefGene, in dem diese Variante vorkommt.

## Analyse-Ausgabedateien für die CF-Assays

Die Analyseergebnisse für die CF-Assays werden auf der Registerkarte „Details“ angezeigt.

Abbildung 18 Registerkarte „Details“ für den CF 139-Varianten-Assay, Beispiel

#	Sample ID	Sample Name	Call Rate	Performance	Control	Comment
1	NA01445	NA01445	100.00	Pass		
2	NA02828	NA02828	100.00	Pass		
3	NA04330	NA04330	100.00	Pass		
4	NA07339	NA07339	100.00	Pass		
5	NA07381	NA07381	100.00	Pass		
6	NA07441	NA07441	100.00	Pass		
7	NA07732	NA07732	100.00	Pass		
8	NA07854	NA07854	100.00	Pass		
9	NA07857	NA07857	100.00	Pass		

#	Sample ID	Sample Name	Mutation Name	Type	dbSNP rsID	CFTR Gene Region	Genomic Location	cDNA Name (HGVS)	Protein Name (HGVS)	Result	Cor
1	NA01445	NA01445	PolyTG/PolyT	PolyTG/PolyT	N/A	Intron 9	117188661-117188689	c.1210-12T[S_S]	N/A	(TG)10(T)/(TG)10(T)9	53
2	NA01445	NA01445	F508del	DEL	rs113993960	Exon 11	117199645	c.1521_1523delCTT	p.Phe508del	HET	30
3	NA01445	NA01445	W1282X	SNV	rs77010898	Exon 23	117282620	c.3846G>A	p.Tyr1282X	HET	30



Die Ergebnisse in den Varianten-Tabellen können durch Klicken auf das Symbol **Export table data to text file** (Tabellendaten in Textdatei exportieren) einzeln in eine Textdatei exportiert werden. Dieser Export ändert nicht die Analyseberichtsdatei.



Nachdem der Medizingenetiker die Variantensignifikanz ermittelt hat, können die Interpretationseinstellungen im Analysebericht gespeichert werden. Der Dateiname des ursprünglichen Analyseberichts wird automatisch mit einem Zeit-/Datumsstempel versehen.

Die Ausgabedateien für die CF-Assays werden auch in einer durch Tabulatoren getrennten Textdatei zusammengefasst, die nach dem für den Lauf verwendeten Assay benannt wird. Diese Ergebnisse sind mit den Ergebnissen identisch, die auf der Registerkarte „Details“ angezeigt werden.

- ▶ Für den CF 139-Varianten-Assay lautet der Dateiname MiSeqDxCf139VariantAssay.txt.
- ▶ Für den klinischen CF-Sequenzierungs-Assay lautet der Dateiname MiSeqDxCfClinicalSequencingAssay.txt.

Nach Abschluss der Analyse wird die Ausgabedatei im Alignment-Ordner des Laufs gespeichert. Beispiel:

MiSeqAnalysis \<RunFolderName> \Data \Intensities \BaseCalls \Alignment

Wenn die Analyse wiederholt oder erneut in die Warteschlange gestellt wurde, wird eine neue Berichtsdatei generiert und im Alignment-Ordner des Analyselaufs gespeichert. Weitere Informationen hierzu finden Sie unter *Analyse erneut in die Warteschlange stellen* auf Seite 13.

Die Ausgabedatei enthält einen Kopfbereich mit folgenden Informationen über den Lauf:

Überschrift	Beschreibung
Test	Beschreibt den Test, der durchgeführt wurde.
Run ID (Lauf-ID)	Dies ist die Lauf-ID, die zu Beginn des Sequenzierungslaufs von MOS generiert wurde.

Überschrift	Beschreibung
Run Date (Laufdatum)	Das Datum (TTMMJJ), an dem der Sequenzierungslauf in MOS gestartet wurde.
Analysis Version (Analyseversion)	Dies ist die Version von MiSeq Reporter, die für die Analyse verwendet wurde.

**Abbildung 19** Kopfbereich für die Ausgabedatei des CF 139-Varianten-Assays, Beispiel

```
Test CF 139-Variant Assay
For In Vitro Diagnostic Use.
Run ID 140212_M01018_0071_0000000000-A2618
Run Date 140212
Analysis Version 2.2.31.1
```

Nach dem Kopfbereich folgt ein Abschnitt mit einer Zusammenfassung für jede Proben-ID, der Spalten für jeden gemeldeten Wert enthält. Beschreibungen der Spalten finden Sie unter Registerkarte „Details“ für den CF 139-Varianten-Assay auf Seite 24 und Registerkarte „Details“ für den klinischen CF-Sequenzierungs-Assay auf Seite 26.



#### HINWEIS

Die Analyse-Pipeline, die Ausgabedateien generiert, ist zwischen den CF-Assays und Universal Kit 1.0 nicht identisch. Die generierten Ausgabedateien für Universal Kit 1.0 sind \*.bam-Dateien, \*.vcf-Dateien und AmpliconCoverage\_M#.tsv-Dateien. Weitere Informationen zu Ausgabedateien für Universal Kit 1.0 finden Sie unter Anhang A Universal Kit 1.0-Analyse-Ausgabedateien.

[Diese Seite wurde absichtlich leer gelassen.]

# Installation und Fehlerbehebung

Anforderungen für MiSeq Reporter außerhalb des Geräts .....	38
Installieren von MiSeq Reporter außerhalb des Geräts .....	39
Verwenden von MiSeq Reporter außerhalb des Geräts .....	41
Fehlerbehebung bei MiSeq Reporter .....	42



## Anforderungen für MiSeq Reporter außerhalb des Geräts

Durch die Installation von MiSeq Reporter auf einem Windows-Computer außerhalb des Geräts wird die Sekundäranalyse von Sequenzierungsdaten ermöglicht, während das MiSeqDx einen nachfolgenden Sequenzierungslauf durchführt.

Weitere Informationen hierzu finden Sie unter *Installieren von MiSeq Reporter außerhalb des Geräts* auf Seite 39.

### IT-Anforderungen

Die MiSeq Reporter-Software benötigt die folgenden IT-Komponenten:

- ▶ 64-Bit-Version eines Windows-Betriebssystems (Vista, Windows 7, Windows Server 2008 64-Bit)
- ▶  $\geq$  Mindestens 8 GB RAM;  $\geq$  16 GB RAM empfohlen
- ▶  $\geq$  1 TB Speicherplatz
- ▶ Quad Core-Prozessor (2,8 GHz oder höher)
- ▶ Microsoft .NET 4

### Unterstützte Browser

MiSeq Reporter kann in den folgenden Browsern angezeigt werden:

- ▶ Safari 5.1.7 oder höher
- ▶ Firefox 13.0.1 oder höher
- ▶ Internet Explorer 8 oder höher

### Herunterladen und Lizenzierung

- 1 Laden Sie eine zweite Kopie der MiSeq Reporter-Software von der Illumina-Website herunter. Hierfür müssen Sie sich bei MyIllumina anmelden.
- 2 Stimmen Sie der Endbenutzer-Lizenzvereinbarung (EULA) zu, wenn Sie während des Installationsvorgangs dazu aufgefordert werden. Sie benötigen keinen Lizenzschlüssel, da diese zusätzliche Kopie kostenlos ist.

## Installieren von MiSeq Reporter außerhalb des Geräts

Um MiSeq Reporter auf einem Windows-Computer außerhalb des Geräts zu installieren, richten Sie zuerst die Berechtigung **Log on as a service** (Als Dienst anmelden) ein und führen Sie dann den Installationsassistenten aus. Konfigurieren Sie anschließend die Software so, dass sie auf das entsprechende Repository und den entsprechenden GenomePath verweist.

### Einrichten von Benutzer- oder Gruppenkonten auf Windows 7

Wenn Sie das Benutzer- oder Gruppenkonto zum Aktivieren der Berechtigung **Log on as a service** (Als Dienst anmelden) konfigurieren möchten, benötigen Sie Administratorrechte. Wenden Sie sich an den vor Ort zuständigen Administrator Ihres Unternehmens, falls Sie Hilfe benötigen.

- 1 Wählen Sie im Windows-Menü **Start Control Panel** (Systemsteuerung) und klicken Sie dann auf **System and Security** (System und Sicherheit).
- 2 Klicken Sie auf **Administrative Tools** (Verwaltung) und dann doppelt auf **Local Security Policy** (Lokale Sicherheitsrichtlinie).
- 3 Doppelklicken Sie links in der Verzeichnisstruktur „Security Settings“ (Sicherheitseinstellungen) auf **Local Policies** (Lokale Richtlinien) und klicken Sie anschließend auf **User Rights Assignments** (Zuweisen von Benutzerrechten).
- 4 Doppelklicken Sie im Detailfenster rechts auf **Log on as a service** (Als Dienst anmelden).
- 5 Klicken Sie im Windows-Dialogfeld auf **Add User or Group** (Benutzer oder Gruppe hinzufügen).
- 6 Geben Sie den Namen des Benutzer- oder Gruppenkontos für diesen Computer ein. Klicken Sie auf **Check Names** (Namen überprüfen), um das Konto zu validieren.
- 7 Klicken Sie dann nacheinander in den geöffneten Dialogfeldern auf **OK** und schließen Sie anschließend die Systemsteuerung.

Weitere Informationen finden Sie unter [technet.microsoft.com/en-us/library/cc739424\(WS.10\).aspx](http://technet.microsoft.com/en-us/library/cc739424(WS.10).aspx) auf der Microsoft-Website.

### Ausführen des MiSeq Reporter-Installationsassistenten

- 1 Laden Sie das Installationspaket von MiSeq Reporter von der Illumina-Website herunter und entpacken Sie es.
- 2 Doppelklicken Sie auf die Datei Setup.exe.
- 3 Klicken Sie bei den Aufforderungen im Installationsassistenten auf **Next** (Weiter).
- 4 Wenn Sie dazu aufgefordert werden, geben Sie den Benutzernamen und das Kennwort eines Kontos mit der Berechtigung **Log on as a service** (Als Dienst anmelden) an, wie im vorherigen Schritt eingerichtet.
- 5 Fahren Sie mit den verbleibenden Aufforderungen fort.

## Konfigurieren von MiSeq Reporter

Um MiSeq Reporter zur Lokalisierung des Laufordners und des Referenzgenomordners zu konfigurieren, bearbeiten Sie die Konfigurationsdatei in einem Texteditor wie Notepad.

- 1 Navigieren Sie zum Installationsordner (standardmäßig C:\Illumina\MiSeq Reporter) und öffnen Sie die Datei MiSeq Reporter.exe.config in einem Texteditor.
- 2 Suchen Sie das Tag **Repository** und ändern Sie den **value** (Wert) in den Standardspeicherort für Daten auf dem Computer außerhalb des Geräts.

Beispiel:

```
<add key="Repository" value="E:\Data\Repository" />
```

Alternativ kann dieser Speicherort ein Speicherort im Netzwerk sein, auf den der (nicht im Gerät integrierte) Computer zugreifen kann.

- 3 Suchen Sie das Tag **GenomePath** und ändern Sie den **value** (Wert) in den Speicherort des Ordners mit den Referenzgenomdateien im FASTA-Format.

Beispiel:

```
<add key="GenomePath" value="E:\MyGenomes\FASTA" />
```

## Starten des MiSeq Reporter-Diensts

Nach Abschluss der Installation startet der MiSeq Reporter-Dienst automatisch. Falls der Dienst nicht gestartet wird, starten Sie ihn mithilfe der folgenden Anweisungen manuell oder führen Sie einen Neustart des Computers durch.

- 1 Klicken Sie im Windows-Menü **Start** mit der rechten Maustaste auf **Computer** und wählen Sie **Manage** (Verwalten) aus.
- 2 Doppelklicken Sie links in der Verzeichnisstruktur „Computer Management“ (Computerverwaltung) auf **Services and Applications** (Dienste und Anwendungen) und klicken Sie anschließend auf **Services** (Dienste).
- 3 Klicken Sie mit der rechten Maustaste auf **MiSeq Reporter** und wählen Sie **Properties** (Eigenschaften) aus.
- 4 Stellen Sie sicher, dass die Option **Startup Type** (Starttyp) auf der Registerkarte „General“ (Allgemein) auf **Automatic** (Automatisch) eingestellt ist, und klicken Sie anschließend auf **Start**.
- 5 Legen Sie auf der Registerkarte „Log On“ (Anmelden) **user name** (Benutzername) und **password** (Kennwort) für ein Konto vom Typ „Services“ (Dienste) fest, das Schreibberechtigungen für den Server hat. Illumina empfiehlt für die meisten Benutzer die Verwendung des Kontos **Local System** (Lokales System). Wenden Sie sich an den vor Ort zuständigen Administrator Ihres Unternehmens, falls Sie Hilfe zu standortspezifischen Netzwerkanforderungen benötigen.
- 6 Klicken Sie dann nacheinander in den geöffneten Dialogfeldern auf **OK** und schließen Sie anschließend das Fenster „Computer Management“ (Computerverwaltung).
- 7 Stellen Sie nach dem Start des MiSeq Reporter-Diensts eine lokale Verbindung zur Software über localhost:8042 in einem Web-Browser her.

## Verwenden von MiSeq Reporter außerhalb des Geräts

Um MiSeq Reporter außerhalb des Geräts zu verwenden, muss der Zugriff auf Ordner mit Laufdaten und Referenzgenomen möglich sein.

- 1 Wenn kein Netzwerkspeicherort für die Sequenzierungsdaten und Referenzgenome verwendet wird, kopieren Sie die folgenden Ordner auf den lokalen Computer:
  - Kopieren Sie Laufdaten vom MiSeqDx-Computer nach D:\MiSeqOutput\<<Laufordner>.
  - Kopieren Sie Referenzgenome vom MiSeqDx-Computer nach C:\Illumina\MiSeq Reporter\Genomes.
- 2 Öffnen Sie einen Web-Browser mit der Adresse <http://localhost:8042>, wodurch die MiSeq Reporter-Web-Oberfläche geöffnet wird.
- 3 Ändern Sie den Pfad zum Repository mithilfe des Symbols **Settings** (Einstellungen)  in der rechten oberen Ecke der Web-Oberfläche.



#### HINWEIS

Der Repository-Pfad wird nur temporär in den Einstellungen angegeben. Beim nächsten Neustart des Computers verweist der Pfad standardmäßig wieder auf den Repository-Speicherort, der in MiSeq Reporter.exe.config angegeben ist.

- 4 Wählen Sie auf der linken Seite der Web-Oberfläche **Analyses** (Analysen), um die am angegebenen Repository-Speicherort verfügbaren Läufe anzuzeigen.
- 5 Bevor die Analyse eines Laufs mit einer Installation von MiSeq Reporter außerhalb des Geräts erneut in die Warteschlange gestellt werden kann, muss der GenomeFolder-Pfad im Probenblatt aktualisiert werden. Dieser Vorgang kann auf der Registerkarte „Sample Sheet“ (Probenblatt) durchgeführt werden. Klicken Sie nach dem Aktualisieren des GenomeFolder-Pfads auf **Save and Requeue** (Speichern und erneut in die Warteschlange stellen).

Weitere Informationen finden Sie unter *Bearbeiten des Probenblatts in MiSeq Reporter* auf Seite 11.

## Fehlerbehebung bei MiSeq Reporter

MiSeq Reporter wird als Windows-Dienstanwendung ausgeführt. Benutzerkonten müssen konfiguriert werden, um die Berechtigungen **Log on as a service** (Anmelden als Dienst) zu aktivieren, bevor MiSeq Reporter installiert wird. Weitere Informationen finden Sie unter *Einrichten von Benutzer- oder Gruppenkonten auf Windows 7* auf Seite 39.

Weitere Informationen finden Sie unter [msdn.microsoft.com/en-us/library/ms189964.aspx](http://msdn.microsoft.com/en-us/library/ms189964.aspx).

### Starten des Diensts fehlgeschlagen

Wenn der Dienst nicht gestartet werden kann, überprüfen Sie im Windows-Ereignisprotokoll die Details der Fehlermeldung.

- 1 Öffnen Sie **Control Panel** (Systemsteuerung) und wählen Sie **Administrative Tools** (Verwaltung) aus.
- 2 Wählen Sie **Event Viewer** (Ereignisanzeige) aus.
- 3 Wählen Sie im Fenster „Event Viewer“ (Ereignisanzeige) **Windows Logs | Application** (Windows-Protokolle | Anwendungen) aus. Der im Ereignisprotokoll aufgelistete Fehler bezieht sich auf Syntaxfehler in MiSeq Reporter.exe.config. Eine fehlerhafte Syntax in der Datei MiSeq Reporter.exe.config kann dazu führen, dass der Dienst fehlschlägt.

### Dateien können nicht kopiert werden

Wenn das Kopieren von Dateien an den beabsichtigten Speicherort fehlschlägt, überprüfen Sie die folgenden Einstellungen:

- 1 Überprüfen Sie den Pfad zum angegebenen Repository-Ordner oder zum MiSeqOutput-Ordner:
  - Prüfen Sie für Installationen außerhalb des Geräts den Repository-Speicherort mithilfe von Settings (Einstellungen)  auf der MiSeq Reporter-Web-Oberfläche.
  - Prüfen Sie für Installationen auf dem Gerät den Speicherort des Ordners MiSeqOutput auf der MOS-Anzeige „Run Options“ (Laufoptionen) auf der Registerkarte „Folder Settings“ (Ordneureinstellungen).

Es muss der vollständige UNC-Pfad (z. B. \\server1\Runs) angegeben werden.

Da MiSeq Reporter als ein Windows-Dienst ausgeführt wird, erkennt er keine vom Benutzer zugeordneten Laufwerke (z. B.: Z:\Runs).

- 2 Vergewissern Sie sich, dass es Schreibzugriff auf den Ausgabeordner gibt. Wenden Sie sich an den vor Ort zuständigen Administrator Ihres Unternehmens, falls Sie Hilfe benötigen.
- 3 Stellen Sie sicher, dass das Kopieren in MiSeq Reporter.exe.config nicht deaktiviert ist. Diese Einstellung befindet sich im Abschnitt <appSettings> und der Wert sollte auf **1** eingestellt sein.

```
<add key="CopyToRTAOutputPath" value="1"/>
```

### Anzeigen von Protokolldateien für einen fehlgeschlagenen Lauf

Der Inhalt der Protokolldateien kann beim Identifizieren spezifischer Fehler für Fehlerbehebungszwecke hilfreich sein.

- 1 Um die Protokolldateien mithilfe der Webbrowser-Oberfläche von MiSeq Reporter anzuzeigen, wählen Sie den Lauf in der Registerkarte „Analyses“ (Analysen) aus.

- 2 Wählen Sie die Registerkarte „Logs“ (Protokolle), um eine Liste aller Schritte anzuzeigen, die während der Analyse durchgeführt wurden. Die Protokollinformationen werden in der Datei AnalysisLog.txt aufgezeichnet, die sich auf der Stammebene des MiSeqAnalysis-Ordners befindet.
- 3 Wählen Sie die Registerkarte „Errors“ (Fehler), um eine Liste der Fehler anzuzeigen, die während der Analyse aufgetreten sind. Die Fehlerinformationen werden in der Datei AnalysisError.txt aufgezeichnet, die sich auf der Stammebene des MiSeqAnalysis-Ordners befindet.

[Diese Seite wurde absichtlich leer gelassen.]

# Universal Kit 1.0-Analyse- Ausgabedateien

Arten von Analyse-Ausgabedateien .....	46
BAM-Dateiformat .....	47
VCF-Dateiformat .....	48
Amplikon-Abdeckungsdatei .....	51
Ergänzende Ausgabedateien .....	52



## Arten von Analyse-Ausgabedateien

Die folgende Tabelle beschreibt die generierten Ausgabedateien für das Universal Kit 1.0. Sie enthalten Analyseergebnisse für das Alignment, Varianten-Calling und die Abdeckung.

Dateiname	Beschreibung
*.bam-Dateien	Enthalten ausgerichtete Reads für eine bestimmte Probe. Befinden sich in Data\Intensities\BaseCalls\Alignment.
*.vcf-Dateien	Enthalten Informationen über Varianten, die an bestimmten Positionen in einem Referenzgenom gefunden wurden. Befinden sich in Data\Intensities\BaseCalls\Alignment.
AmpliconCoverage_M#.tsv	Enthält Details zur resultierenden Abdeckung pro Amplikon pro Probe. M# steht für die Manifestnummer. Befinden sich in Data\Intensities\BaseCalls\Alignment.



### HINWEIS

Die Analyse-Pipeline, die diese Ausgabedateien generiert, ist zwischen den CF-Assays und Universal Kit 1.0 nicht identisch. In diesem Abschnitt werden nur die Analyse-Ausgabedateien für Universal Kit 1.0 beschrieben.

## BAM-Dateiformat

Eine BAM-Datei (\*.bam) ist die komprimierte Binärversion einer SAM-Datei, die für ausgerichtete Sequenzen verwendet wird. SAM- und BAM-Formate werden ausführlich auf der Website von SAMtools beschrieben: [samtools.sourceforge.net](http://samtools.sourceforge.net).

BAM-Dateien werden im Alignment-Ordner unter `Data\Intensities\BaseCalls\Alignment` in dem Dateinamenformat `Probename_S#.bam` gespeichert, wobei # für die Probennummer steht, die sich von der Reihenfolge ableitet, in der die Proben im Probenblatt aufgeführt sind.

BAM-Dateien enthalten einen Kopfbereich und einen Alignment-Bereich:

- ▶ **Kopfbereich** – Enthält Informationen über die gesamte Datei, z. B. den Probennamen und die Probenlänge. Alignments im Alignment-Bereich sind mit spezifischen Informationen im Kopfbereich verbunden.
- ▶ **Alignments** – Enthält den Read-Namen, die Read-Sequenz, die Read-Qualität und anwendungsspezifische Tags.

**Abbildung 20** Beispiel, Alignment-Bereich einer BAM-Datei

```
GA23_40:8:1:10271:11781 64 chr22 17552189 8 35M * 0 0
TACAGACATCCACCACCACACCCAGCTAATTTTTG
IIIII>FA?C: :B=:GGGB>GGGEGIIIIHI3EEE#
BC:Z:ATCACG XD:Z:55 SM:I:8
```

Der Read-Name enthält das Chromosom und die Startkoordinate (**chr22 17552189**), die Alignment-Qualität (**8**) und den Übereinstimmungsdeskriptor (**35M \* 0 0**).

BAM-Dateien können mit einem externen Viewer, z. B. IGV oder UCSC Genome Browser, betrachtet werden.

## VCF-Dateiformat

VCF ist ein gängiges Dateiformat, das von der Genomik-Wissenschaftsgemeinde entwickelt wurde und Informationen über Varianten enthält, die an spezifischen Positionen in einem Referenzgenom gefunden wurden.

VCF-Dateien verwenden das Dateinamensformat `Probenname_S#.vcf`, wobei # für die Probennummer steht, die sich von der Reihenfolge ableitet, in der die Proben im Probenblatt aufgeführt sind.

- ▶ **Kopfbereich der VCF-Datei** – Enthält die Version des VCF-Dateiformats und die Version des Varianten-Callers. Außerdem enthält der Kopfbereich die Anmerkungen, die im Rest der Datei verwendet werden. Die letzte Zeile im Kopfbereich enthält die Spaltenüberschriften für die Datenzeilen. Weitere Informationen finden Sie unter *Überschriften und Anmerkungen in der VCF-Datei* auf Seite 49.

Abbildung 21 Beispiel, Kopfbereich der VCF-Datei

```
##fileformat=VCFv4.1
##FORMAT=<ID=GQX,Number=1,Type=Integer,Description="Minimum of
  {Genotype quality assuming variant position,Genotype quality
  assuming non-variant position}">
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">
##FORMAT=<ID=GQ,Number=1,Type=Float,Description="Genotype
  Quality">
##FORMAT=<ID=AD,Number=.,Type=Integer,Description="Allelic depths
  for the ref and alt alleles in the order listed">
##FORMAT=<ID=VF,Number=1,Type=Float,Description="Variant
  Frequency, the ratio of the sum of the called variant depth to
  the total depth">
##INFO=<ID=TI,Number=.,Type=String,Description="Transcript ID">
##INFO=<ID=GI,Number=.,Type=String,Description="Gene ID">
##INFO=<ID=EXON,Number=0,Type=Flag,Description="Exon Region">
##INFO=<ID=FC,Number=.,Type=String,Description="Functional
  Consequence">
##INFO=<ID=AC,Number=A,Type=Integer,Description="Allele count in
  genotypes, for each ALT allele, in the same order as listed">
##INFO=<ID=AF,Number=A,Type=Float,Description="Allele Frequency,
  for each ALT allele, in the same order as listed">
##INFO=<ID=AN,Number=1,Type=Integer,Description="Total number of
  alleles in called genotypes">
##INFO=<ID=DP,Number=1,Type=Integer,Description="Approximate read
  depth; some reads may have been filtered">
##FILTER=<ID=LowVariantFreq,Description="Low variant frequency <
  0.20">
##FILTER=<ID=LowGQ,Description="GQ below < 20.00">
##FILTER=<ID=LowQual,Description="QUAL below < 100.00">
##FILTER=<ID=R8,Description="IndelRepeatLength is greater than
  8">
##fileDate=20130506
##source=Starling 0.3
##phasing=none
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT
```

- **Datenzeilen der VCF-Datei** – Jede Zeile enthält Informationen zu einer einzelnen Variante. Die Datenzeilen befinden sich unter den Spaltenüberschriften im Kopfbereich.

## Überschriften und Anmerkungen in der VCF-Datei

Das VCF-Dateiformat ist flexibel und erweiterbar. In den folgenden Tabellen sind die Überschriften und Anmerkungen in der VCF-Datei beschrieben, die von MiSeq Reporter generiert werden.

### Überschriften der VCF-Datei

Überschrift	Beschreibung
CHROM	Das Chromosom des Referenzgenoms. Chromosomen werden in derselben Reihenfolge wie in der FASTA-Referenzdatei aufgeführt.
POS	Die Einzelbasenposition der Variante im Referenzchromosom. Bei SNPs ist diese Position die Referenzbase mit der Variante. Bei Indels oder Deletionen ist diese Position die Referenzbase unmittelbar vor der Variante.
ID	Die rs-Nummer für den SNP aus dbSNP.txt, falls vorhanden. Wenn es mehrere rs-Nummern an dieser Position gibt, wird die Liste durch Semikola getrennt. Wenn an dieser Position kein dbSNP-Eintrag existiert, wird eine Kennung für einen fehlenden Wert ('.') verwendet.
REF	Der Referenz-Genotyp. Beispielsweise wird die Deletion eines einzelnen T als Referenz-TT und alternatives T dargestellt.
ALT	Die Allele, die sich vom Referenz-Read unterscheiden. Beispielsweise wird die Insertion eines einzelnen T als Referenz-A und alternatives AT dargestellt.
QUAL	Ein vom Varianten-Caller zugewiesener Phred-skaliertes Qualitäts-Score. Höhere Scores weisen auf eine höhere Zuverlässigkeit der Variante und eine geringere Fehlerwahrscheinlichkeit hin. Bei einem Qualitäts-Score von Q beträgt die geschätzte Fehlerwahrscheinlichkeit $10^{-(Q/10)}$ . Beispielsweise hat die Reihe der Q30-Calls eine Fehlerrate von 0,1 %. Viele Varianten-Caller weisen Qualitäts-Scores auf Basis ihrer statistischen Modelle zu, die relativ zur beobachteten Fehlerrate hoch sind.

### Anmerkungen in der VCF-Datei

Überschrift	Beschreibung
FILTER	<p>Wenn alle Filter passiert werden, wird „PASS“ in die Filterspalte geschrieben.</p> <ul style="list-style-type: none"> <li>• <b>LowDP</b> – Angewendet auf Positionen, bei denen die Abdeckungstiefe unter dem Cutoff liegt.</li> <li>• <b>LowGQ</b> – Die Genotypisierungsqualität (GQ) liegt unter dem Cutoff.</li> <li>• <b>LowQual</b> – Die Variantenqualität (QUAL) liegt unter dem Cutoff.</li> <li>• <b>LowVariantFreq</b> – Die Variantenhäufigkeit liegt unter dem Schwellenwert.</li> <li>• <b>R8</b> – Bei einem Indel ist die Anzahl der benachbarten Wiederholungen (1-Base oder 2-Base) in der Referenz größer als 8.</li> </ul>

Überschrift	Beschreibung
INFO	<ul style="list-style-type: none"> <li>• <b>AC</b> – Allelzahl in Genotypen für jedes ALT-Allel, in derselben Reihenfolge wie aufgeführt.</li> <li>• <b>AF</b> – Allelhäufigkeit für jedes ALT-Allel, in derselben Reihenfolge wie aufgeführt.</li> <li>• <b>AN</b> – Die Gesamtzahl der Allele in aufgerufenen Genotypen.</li> <li>• <b>Exon</b> – Eine kommasetrennte Liste der Exon-Regionen, die aus dem RefGene gelesen wurden.</li> <li>• <b>FC</b> – Functional Consequence (funktionale Konsequenz).</li> <li>• <b>GI</b> – Eine kommasetrennte Liste der Gen-IDs, die aus dem RefGene gelesen wurden.</li> <li>• <b>TI</b> – Eine kommasetrennte Liste der Transkript-IDs, die aus dem RefGene gelesen wurden.</li> </ul>
FORMAT	<ul style="list-style-type: none"> <li>• <b>AD</b> – Eintrag im Format X,Y, wobei X für die Anzahl der Referenz-Calls und Y für die Anzahl der alternativen Calls steht.</li> <li>• <b>DP</b> – Ungefähre Read-Tiefe; Reads mit MQ = 255 oder mit fehlerhaften „Mates“ werden herausgefiltert.</li> <li>• <b>GQ</b> – Genotypqualität.</li> <li>• <b>GQX</b> – Genotypqualität. GQX ist das Minimum des GQ-Werts und der QUAL-Spalte. In der Regel sind diese Werte ähnlich. Als Minimum ist GQX der konservativere Messwert für die Genotypqualität.</li> <li>• <b>GT</b> – Genotyp. 0 entspricht der Referenzbase, 1 entspricht dem ersten Eintrag in der ALT-Spalte usw. Der Schrägstrich (/) gibt an, dass keine Phasierungsinformationen vorhanden sind.</li> <li>• <b>VF</b> – Variant frequency (Variantenhäufigkeit); der Prozentsatz der Reads, die das alternative Allel unterstützen.</li> </ul>
SAMPLE (PROBE)	Die Probenspalte enthält die Werte, die in der Spalte FORMAT angegeben sind.

## Amplikon-Abdeckungsdatei

Für jedes Manifest wird eine Amplikon-Abdeckungsdatei generiert. Das „M#“ im Dateinamen steht für die Manifestnummer, wie sie im Probenblatt aufgeführt ist.

Jede Datei beginnt mit einem Kopfbereich, der die mit dem Manifest verbundenen Proben-IDs enthält. Die erste Spalte enthält die Target-IDs. Jede weitere Spalte enthält die Abdeckungstiefe für die zugehörige Proben-ID.

## Ergänzende Ausgabedateien

Die folgenden Ausgabedateien bieten ergänzende Informationen oder eine Zusammenfassung von Laufergebnissen und Analysefehlern. Zwar sind diese Dateien nicht für die Beurteilung der Analyseergebnisse erforderlich, können aber für die Fehlerbehebung verwendet werden.

Dateiname	Beschreibung
AnalysisLog.txt	Verarbeitungsprotokoll, in dem jeder Schritt während der Analyse des aktuellen Laufordners beschrieben ist. Diese Datei enthält keine Fehlermeldungen. Befindet sich auf der Stammebene des Laufordners.
AnalysisError.txt	Verarbeitungsprotokoll, in dem alle während der Analyse aufgetretenen Fehler aufgeführt sind. Diese Datei ist nur vorhanden, wenn Fehler aufgetreten sind. Befindet sich auf der Stammebene des Laufordners.
AmpliconRunStatistics.xml	Enthält eine laufspezifische Zusammenfassungsverstatistik. Befindet sich auf der Stammebene des Laufordners.
CompletedJobInfo.xml	Wird nach Abschluss der Analyse erstellt, enthält Informationen über den Lauf, z. B. Datum, Fließzellen-ID, Softwareversion und andere Parameter. Befindet sich auf der Stammebene des Laufordners.
DemultiplexSummaryF1L1.txt	Gibt die Demultiplexing-Ergebnisse in einer Tabelle mit einer Zeile pro Platte und einer Spalte pro Probe an. Befindet sich in Data\Intensities\BaseCalls\Alignment.
ErrorsAndNoCallsByLaneTileReadCycle.csv	Eine Datei mit kommagetrennten Werten, die den Prozentsatz der Fehler und No-Calls pro Platte, pro Read und pro Zyklus enthält. Befindet sich in Data\Intensities\BaseCalls\Alignment.
Mismatch.htm	Enthält Histogramme von Nichtübereinstimmungen pro Zyklus und von No-Calls pro Zyklus für jede Platte. Befindet sich in Data\Intensities\BaseCalls\Alignment.
Summary.xml	Enthält eine Zusammenfassung der Nichtübereinstimmungsraten und anderer Base-Calling-Ergebnisse. Befindet sich in Data\Intensities\BaseCalls\Alignment.
Summary.htm	Enthält eine aus Summary.xml generierte Zusammenfassungsverwebseite. Befindet sich in Data\Intensities\BaseCalls\Alignment.

\*

\*.bam 47  
\*.bam.bai 47  
\*.vcf 48

## A

Abdeckungsdiagramm, Qscore-Diagramm 8  
Alignment 16  
Als Dienst anmelden 39  
Analyse  
während der Sequenzierung 2  
Analyse erneut in die Warteschlange stellen 6, 11, 13  
Analyseordner 9  
Analyseordner, Laufordner  
Beziehung 17  
AnalysisError.txt 42  
AnalysisLog.txt 42  
Anmelden als Dienst 42  
Anzeigen von MiSeq Reporter 3  
Assays 20

## B

BAM-Dateien  
Dateiformat 47  
BAM-Indexdateien 47  
Bearbeiten des Probenblatts 11

## C

CF 139-Varianten-Assay 20  
Cluster nach Filterung 14  
Custom Amplicon-Workflow 22

## D

Dateien können nicht kopiert werden 42  
Datenbanken, vorinstalliert 21  
Datenordner 9  
dbSNP-Datenbank 21  
Demultiplexierung 16  
Diagramm „Clusters“ (Cluster) 7  
Diagramm „High Percentages“ (Hohe Prozentwerte) 7  
Diagramm „Low Percentages“ (Niedrige Prozentwerte) 7  
Diagramm „Mismatch“ (Nichtübereinstimmung) 7  
Diagramm mit Varianten-Score 8  
DLSO 22  
Dokumentation 55

## E

Eingabedateien 21

## F

FASTQ-Dateien 16  
Fehlerbehebung  
Dateien können nicht kopiert werden 42  
Protokolldateien 42  
Starten des Diensts  
fehlgeschlagen 42  
Fehlerwahrscheinlichkeit 14

## G

Genompfad 40  
GI Gen-ID 49  
GT Genotyp 49

## H

Hilfe, technisch 55

## I

Installation, außerhalb des Geräts 39  
IP-Adresse, MiSeq Reporter 3  
IT-Anforderungen 38

## K

Kit 20  
Klinischer CF-Sequenzierungs-Assay 20  
Kopien-Ordner 9  
Kundendienst 55

## L

Laufordner  
Info 4  
Lizenz (EULA) 38  
localhost 3  
Lokale Sicherheitsrichtlinie 39  
Lokales Systemkonto 40  
LowDP 49  
LowGQ 49  
LowVariantFreq 49

## M

Manifest-Datei 4  
Manifestdatei 10  
MiSeqAnalysis-Ordner 17  
MiSeqDxCf139VariantAssay.txt 34  
MiSeqDxCfClinicalSequencingAssay.txt 34  
MiSeqOutput-Ordner 17

## P

Passing Filter (PF) 14  
Phasierung, Vorphasierung 14

- Probenblatt
  - bearbeiten 11
  - Info 4
- Probentabelle 8
- Protokolldateien 42
- Q
- Qualitäts-Scores,Q-Scores 14
- R
- r8 49
- Read-Zyklen 9
- Referenz-Genome, vorinstalliert 21
- RefGene-Datenbank 21
- Repository-Pfad 5, 40
- RTAComplete.txt 21
- RunInfo.xml 21
- S
- SAM-Tools 47
- SampleSheet.csv 21
- Server-URL 5
- Smith-Waterman 16, 22
- Starten des Diensts fehlgeschlagen 42
- Symbole, Analysestatus 6
- T
- Technische Unterstützung 55
- TI Transkript-ID 49
- U
- ULSO 22
- Universal Kit 1.0 20
- V
- Variantentabelle 8
- VCF-Dateien
  - Anmerkungen 49
  - Dateiformat 48
- VF Variantenhäufigkeit 49
- W
- Windows-Dienst
  - Info 2
  - Log on as service 42
- Workflow
  - Custom Amplicon 22
  - MiSeqDx-Software 2

## Technische Unterstützung

Wenn Sie technische Unterstützung benötigen, wenden Sie sich bitte an den technischen Support von Illumina.

**Tabelle 4** Allgemeine Kontaktinformationen für Illumina

Website	www.illumina.com
E-Mail	techsupport@illumina.com

**Tabelle 5** Allgemeine Kontaktinformationen für Illumina

Website	www.illumina.com
E-Mail	techsupport@illumina.com

**Tabelle 6** Telefonnummern des Illumina-Kundendienstes

Region	Telefonnummer	Region	Telefonnummer
Nordamerika	1.800.809.4566	Japan	0800.111.5011
Australien	1.800.775.688	Neuseeland	0800.451.650
Belgien	0800.81102	Niederlande	0800.0223859
China	400.635.9898	Norwegen	800.16836
Dänemark	80882346	Österreich	0800.296575
Deutschland	0800.180.8994	Schweden	020790181
Finnland	0800.918363	Schweiz	0800.563118
Frankreich	0800.911850	Singapur	1.800.579.2745
Großbritannien	0800.917.0041	Spanien	900.812168
Hongkong	800960230	Taiwan	00806651752
Irland	1.800.812949	Andere Länder	+44.1799.534000
Italien	800.874909		

### Materialsicherheitsdatenblätter

Sicherheitsdatenblätter (SDS, Safety Data Sheets) stehen auf der Illumina-Website unter [support.illumina.com/sds.html](http://support.illumina.com/sds.html) zur Verfügung.

### Produktdokumentation

Die Produktdokumentation steht auf der Illumina-Website im PDF-Format zum Herunterladen zur Verfügung. Gehen Sie zu [support.illumina.com](http://support.illumina.com), wählen Sie ein Produkt aus und klicken Sie anschließend auf (**Dokumentation und Literatur**).

