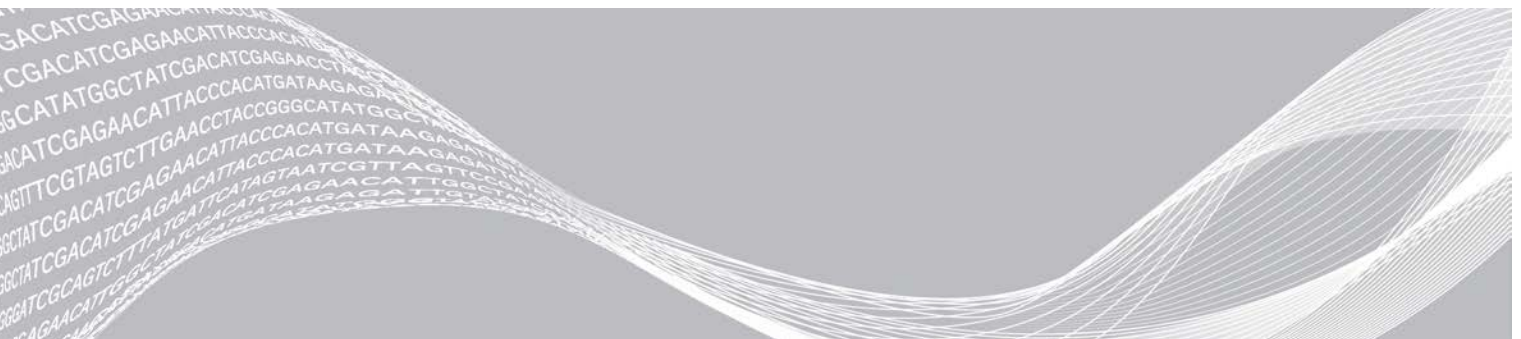


# Local Run Manager Somatic Variant Analysis Module

Οδηγός ροής εργασιών για το NextSeq 550Dx

ΓΙΑ ΧΡΗΣΗ ΣΤΗ ΔΙΑΓΝΩΣΗ IN VITRO

Επισκόπηση	3
Εισαγωγή πληροφοριών εκτέλεσης	4
Μέθοδοι ανάλυσης	6
Προβολή δεδομένων εκτέλεσης και δείγματος	8
Αναφορά ανάλυσης	8
Αρχεία εξόδου ανάλυσης	10
Αντιστοίχιση βάσης και ποικιλομορφία ευρετηρίων	17
Ιστορικό αναθεώρησης	18
Τεχνική βοήθεια	19



Το παρόν έγγραφο και τα περιεχόμενά του αποτελούν ιδιοκτησία της Illumina, Inc. και των συνδεδεμένων εταιρειών της («Illumina») και προορίζονται αποκλειστικά για τη συμβατική χρήση του πελάτη της σε συνδυασμό με τη χρήση του (-ων) προϊόντος(-ων) που περιγράφονται στο παρόν έγγραφο και για κανέναν άλλον σκοπό. Απαγορεύεται η χρήση ή η διανομή του παρόντος εγγράφου και των περιεχομένων του για οποιονδήποτε άλλον σκοπό ή/και άλλη κοινοποίηση, αποκάλυψη ή αναπαραγωγή τους με οποιονδήποτε τρόπο χωρίς την πρότερη έγγραφη συναίνεση της Illumina. Η Illumina δεν μεταβιβάζει διά του παρόντος εγγράφου καμία άδεια δυνάμει διπλώματος ευρεσιτεχνίας, εμπορικού σήματος, πνευματικού δικαιώματος ή δικαιωμάτων κοινού δικαίου της.

Οι οδηγίες στο παρόν έγγραφο πρέπει να τηρούνται αυστηρά και με ακρίβεια από ειδικευμένο και κατάλληλα εκπαιδευμένο προσωπικό, προκειμένου να διασφαλιστεί η ορθή και ασφαλής χρήση του(-ων) προϊόντος(-ων) που περιγράφονται στο παρόν. Όλα τα περιεχόμενα του παρόντος εγγράφου πρέπει να αναγνωσθούν και να γίνουν πλήρως κατανοητά πριν από τη χρήση του(-ων) εν λόγω προϊόντος(-ων).

ΣΕ ΠΕΡΙΠΤΩΣΗ ΜΗ ΠΛΗΡΟΥΣ ΑΝΑΓΝΩΣΗΣ ΚΑΙ ΤΗΡΗΣΗΣ ΜΕ ΑΚΡΙΒΕΙΑ ΟΛΩΝ ΤΩΝ ΟΔΗΓΙΩΝ ΠΟΥ ΠΕΡΙΕΧΟΝΤΑΙ ΣΤΟ ΠΑΡΟΝ, ΜΠΟΡΕΙ ΝΑ ΠΡΟΚΛΗΘΕΙ ΖΗΜΙΑ ΣΤΟ(-Α) ΠΡΟΪΟΝ(-ΤΑ), ΤΡΑΥΜΑΤΙΣΜΟΣ ΑΤΟΜΩΝ, ΣΥΜΠΕΡΙΛΑΜΒΑΝΟΜΕΝΩΝ ΤΩΝ ΧΡΗΣΤΩΝ Ή ΑΛΛΩΝ, ΚΑΘΩΣ ΚΑΙ ΑΛΛΗ ΥΛΙΚΗ ΖΗΜΙΑ, ΚΑΙ ΘΑ ΚΑΤΑΣΤΕΙ ΑΚΥΡΗ Η ΕΓΓΥΗΣΗ ΠΟΥ ΙΣΧΥΕΙ ΓΙΑ ΤΟ(-Α) ΠΡΟΪΟΝ(-ΤΑ).

Η ILLUMINA ΔΕΝ ΑΝΑΛΑΜΒΑΝΕΙ ΚΑΜΙΑ ΕΥΘΥΝΗ ΠΟΥ ΑΠΟΡΡΕΕΙ ΑΠΟ ΕΣΦΑΛΜΕΝΗ ΧΡΗΣΗ ΤΟΥ(-ΩΝ) ΠΡΟΪΟΝΤΟΣ(-ΩΝ) ΠΟΥ ΠΕΡΙΓΡΑΦΟΝΤΑΙ ΣΤΟ ΠΑΡΟΝ [ΣΥΜΠΕΡΙΛΑΜΒΑΝΟΜΕΝΩΝ ΤΩΝ ΕΞΑΡΤΗΜΑΤΩΝ ΤΟΥ(-ΟΥΣ) Ή ΤΟΥ ΛΟΓΙΣΜΙΚΟΥ].

© 2021 Illumina, Inc. Με την επιφύλαξη παντός δικαιώματος.

Όλα τα σήματα κατατεθέντα είναι ιδιοκτησία της Illumina, Inc. ή των αντίστοιχων κατόχων τους. Για συγκεκριμένες πληροφορίες σχετικά με τα σήματα κατατεθέντα, επισκεφτείτε την ηλεκτρονική διεύθυνση [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html).

## Επισκόπηση

Το Local Run Manager Somatic Variant Module προορίζεται για χρήση με τον προσδιορισμό Illumina TruSeq Custom Amplicon Kit Dx και με το NextSeq 550Dx. Όταν χρησιμοποιείται με το Somatic Variant Module, ο προσδιορισμός προορίζεται για την προετοιμασία βιβλιοθηκών που χρησιμοποιούνται για την αλληλούχιση DNA από δείγματα ιστού μονιμοποιημένα με φορμαλίνη και εγκλεισμένα σε παραφίνη (FFPE). Με τον προσδιορισμό ανιχνεύονται σωματικές μεταλλάξεις σε χαμηλές συχνότητες παραλλαγών.

Η μονάδα ανάλυσης αξιολογεί μικρές περιοχές ενισχυμένου DNA ή αμπλικόνια για παραλλαγές. Η εστιασμένη αλληλούχιση των αμπλικονίων προσφέρει υψηλή κάλυψη συγκεκριμένων περιοχών για μεγάλο αριθμό δειγμάτων. Η μονάδα ανάλυσης εκτελεί δευτερεύουσα ανάλυση και δημιουργεί αναφορές από εκτελέσεις αλληλούχισης χρησιμοποιώντας προσέγγιση διπλής αλυσίδας που περιλαμβάνει δεξαμενές ολιγονουκλεϊδίων ευθείας και αντίστροφης διαδικασίας. Ανατρέξτε στο ένθετο συσκευασίας *TruSeq Custom Amplicon Kit Dx* (αρ. εγγράφου 1000000029772).

Για τη μονάδα ανάλυσης Somatic Variant απαιτούνται αναλώσιμα αλληλούχισης 300 κύκλων. Για περισσότερες πληροφορίες, βλ. το ένθετο συσκευασίας *NextSeq 550Dx High Output Reagent Kit v2* ή *NextSeq 550Dx High Output Reagent Kit v2.5*.

## Σχετικά με αυτόν τον οδηγό

Αυτός ο οδηγός περιλαμβάνει οδηγίες για τον καθορισμό παραμέτρων εκτέλεσης για αλληλούχιση και ανάλυση με τη μονάδα ανάλυσης Somatic Variant. Για πληροφορίες σχετικά με τον πίνακα εργαλείων του Local Run Manager και τις ρυθμίσεις του συστήματος, βλ. Οδηγός αναφοράς οργάνου *NextSeq 550Dx* (αρ. εγγράφου 1000000009513).

## Προβολή Local Run Manager

Το περιβάλλον εργασίας χρήστη του Local Run Manager είναι προσβάσιμο μέσω του NextSeq 550Dx Operating Software (NOS) ή μέσω κάποιου προγράμματος περιήγησης στο διαδίκτυο.

Το υποστηριζόμενο πρόγραμμα περιήγησης στο διαδίκτυο είναι το Chromium.



### ΣΗΜΕΙΩΣΗ

Εάν χρησιμοποιείτε πρόγραμμα περιήγησης στο διαδίκτυο που δεν υποστηρίζεται, πραγματοποιήστε λήψη του υποστηριζόμενου προγράμματος περιήγησης όταν σας ζητηθεί στο μήνυμα Confirm Unsupported Browser (Επιβεβαίωση μη υποστηριζόμενου προγράμματος περιήγησης). Επιλέξτε «**εδώ**» για να πραγματοποιήσετε λήψη της υποστηριζόμενης έκδοσης του Chromium.

## Προβολή στην οθόνη του οργάνου

- 1 Για να δείτε το περιβάλλον εργασίας του Local Run Manager στην οθόνη του οργάνου, ορίστε μία από τις παρακάτω επιλογές:
  - ▶ Από την αρχική οθόνη του NOS, επιλέξτε **Local Run Manager**. Επιλέξτε το «X» στην επάνω δεξιά γωνία για να επιστρέψετε στο NOS.
  - ▶ Επιλέξτε το εικονίδιο ελαχιστοποίησης στο NOS, ανοίξτε το πρόγραμμα περιήγησης στο διαδίκτυο Chromium που είναι εγκατεστημένο στο όργανο και εισαγάγετε τη διεύθυνση **http://localhost** στη γραμμή διεύθυνσης. Μόνο οι διαχειριστές μπορούν να εκτελέσουν ελαχιστοποίηση του NOS.

## Προβολή σε υπολογιστή στο δίκτυο

- 1 Ανοίξτε το πρόγραμμα περιήγησης στο διαδίκτυο Chromium σε έναν υπολογιστή με πρόσβαση στο ίδιο δίκτυο με το όργανο και συνδεθείτε χρησιμοποιώντας τη διεύθυνση IP ή το όνομα του οργάνου. Για παράδειγμα, <http://myinstrument>.

## Εισαγωγή πληροφοριών εκτέλεσης

### Καθορισμός παραμέτρων

- 1 Συνδεθείτε στο Local Run Manager.
- 2 Επιλέξτε **Create Run** (Δημιουργία εκτέλεσης) και **Somatic Variant**.
- 3 Εισαγάγετε ένα όνομα εκτέλεσης που ταυτοποιεί την εκτέλεση από την αλληλούχηση έως την ανάλυση.  
Χρησιμοποιήστε αλφαριθμητικούς χαρακτήρες, κενά, κάτω παύλες ή παύλες.
- 4 **[Προαιρετικό]** Εισαγάγετε μια περιγραφή της εκτέλεσης για ευκολότερη ταυτοποίηση της εκτέλεσης.  
Χρησιμοποιήστε αλφαριθμητικούς χαρακτήρες, κενά, κάτω παύλες ή παύλες.
- 5 Επιλέξτε τον αριθμό των δειγμάτων και το σύνολο των ευρετηρίων από την αναπτυσσόμενη λίστα. Κατά την επιλογή σας, λάβετε υπόψη σας τις παρακάτω πληροφορίες.
  - ▶ Η αναπτυσσόμενη λίστα περιλαμβάνει αριθμούς δειγμάτων με ένα σύνολο ευρετηρίων. Για παράδειγμα, το "24-Set 1" υποδεικνύει ότι υπάρχουν 24 δείγματα προς εξέταση, με ευρετήρια από το σύνολο ευρετηρίων 1.
  - ▶ Οι αριθμοί των συνόλων των ευρετηρίων αναφέρονται σε διαφορετικά σύνολα ευρετηρίων i5. Το Σύνολο 1 και το Σύνολο 2 προσφέρουν και τα δύο ποικιλομορφία ευρετηρίων. Παρέχονται δύο σύνολα ευρετηρίων για την αποφυγή εξάντλησης ενός μόνο συνόλου.
  - ▶ Επιλέξτε τον αριθμό δειγμάτων που πλησιάζει περισσότερο στον αριθμό των δειγμάτων που θα υποβληθούν σε εξέταση. Εάν στη λίστα δεν υπάρχει ο ακριβής αριθμός των δειγμάτων, επιλέξτε τον πλησιέστερο αριθμό, αλλά φροντίστε να είναι μικρότερος από τον αριθμό των προς εξέταση δειγμάτων. Για παράδειγμα, αν πρόκειται να εξεταστούν 18 δείγματα, επιλέξτε 16 δείγματα.
  - ▶ Τα βοηθία δειγμάτων και οι συνδυασμοί ευρετηρίων που πληρούν τις απαιτήσεις ποικιλομορφίας ευρετηρίων επισημαίνονται με πράσινο χρώμα. Εάν επιλέξετε άλλα βοηθία και άλλους συνδυασμούς ευρετηρίων, κατά την αποθήκευση της εκτέλεσης θα λάβετε μια ειδοποίηση σε περίπτωση που δεν πληρούνται οι απαιτήσεις ποικιλομορφίας.

### Εισαγωγή αρχείων δελτίου δείγματος για την εκτέλεση

- 1 Βεβαιωθείτε ότι τα δελτία δείγματος που θέλετε να εισαγάγετε είναι διαθέσιμα σε μια προσβάσιμη θέση στο δίκτυο ή σε μια μονάδα USB.
- 2 Επιλέξτε **Import Manifests** (Εισαγωγή δελτίων δείγματος).
- 3 Μεταβείτε στο αρχείο του δελτίου δείγματος και επιλέξτε τα δελτία δείγματος που θέλετε να προσθέσετε.



## ΣΗΜΕΙΩΣΗ

Για να είναι διαθέσιμα τα αρχεία δελτίων δείγματος για όλες τις εκτελέσεις με τη μονάδα ανάλυσης Somatic Variant, προσθέστε τα δελτία δείγματος με τη λειτουργία Module Settings (Ρυθμίσεις μονάδας). Για τη λειτουργία αυτή απαιτούνται άδειες επιπέδου διαχειριστή. Για περισσότερες λεπτομέρειες, βλ. *Οδηγός αναφοράς οργάνου NextSeq 550Dx (αρ. εγγράφου 100000009513)*.

## Καθορισμός δειγμάτων για την εκτέλεση


Καθορίστε τα δείγματα για την εκτέλεση χρησιμοποιώντας μία από τις επιλογές και τις οδηγίες που ακολουθούν.

- ▶ **Μη αυτόματη εισαγωγή δειγμάτων**—Χρησιμοποιήστε τον κενό πίνακα στην οθόνη Create Run (Δημιουργία εκτέλεσης).
- ▶ **Εισαγωγή δειγμάτων**—Μεταβείτε σε ένα εξωτερικό αρχείο με μορφή τιμών χωρισμένων με κόμμα (\*.csv). Στην οθόνη Create Run (Δημιουργία εκτέλεσης) υπάρχει ένα πρότυπο προς λήψη.


Αφού συμπληρώσετε τον πίνακα των δειγμάτων, μπορείτε να εξαγάγετε τις πληροφορίες των δειγμάτων σε ένα εξωτερικό αρχείο. Χρησιμοποιήστε το αρχείο ως πρότυπο για την προετοιμασία βιβλιοθηκών ή εισαγάγετε το αρχείο για μια άλλη εκτέλεση.

## Μη αυτόματη εισαγωγή δειγμάτων

- 1 Εισαγάγετε ένα μοναδικό όνομα δείγματος στο πεδίο Sample Name (Όνομα δείγματος). Χρησιμοποιήστε αλφαριθμητικούς χαρακτήρες, παύλες ή κάτω παύλες. Το όνομα δείγματος συμπληρώνει αυτόματα το αντίστοιχο βοθρίο στην άλλη δεξαμενή.
- 2 **[Προαιρετικό]** Για δείγματα με θετικούς ή αρνητικούς μάρτυρες, κάντε δεξί κλικ και επιλέξτε τον τύπο του μάρτυρα. Ο μάρτυρας στο ένα βοθρίο δείγματος συμπληρώνει αυτόματα το αντίστοιχο βοθρίο στην άλλη δεξαμενή με τον ίδιο μάρτυρα.
- 3 **[Προαιρετικό]** Εισαγάγετε μια περιγραφή δείγματος στο πεδίο Sample Description (Περιγραφή δείγματος). Χρησιμοποιήστε αλφαριθμητικούς χαρακτήρες, παύλες ή κάτω παύλες. Η περιγραφή δείγματος συμπληρώνει αυτόματα το αντίστοιχο βοθρίο στην άλλη δεξαμενή. Οι περιγραφές των δειγμάτων συνδέονται με ένα αναγνωριστικό δείγματος. Οι περιγραφές των δειγμάτων αντικαθίστανται αν το ίδιο αναγνωριστικό δείγματος χρησιμοποιηθεί ξανά σε επόμενη εκτέλεση.
- 4 Επιλέξτε έναν προσαρμογέα Ευρετηρίου 1 από την αναπτυσσόμενη λίστα Index 1 (Ευρετήριο 1) (i7). Όταν χρησιμοποιείτε τα προτεινόμενα βοθρία δείγματος, το λογισμικό συμπληρώνει αυτόματα προσαρμογείς ευρετηρίου i7 και i5 που πληρούν τις απαιτήσεις ποικιλομορφίας ευρετηρίων. Εάν στη λίστα δεν υπάρχει ακριβώς ο αριθμός των δειγμάτων προς εξέταση, φροντίστε να επιλέξετε προσαρμογείς ευρετηρίου για επιπλέον βοθρία. Εάν πρέπει να επιλέξετε ευρετήρια για επιπλέον βοθρία ή αν δεν χρησιμοποιήσετε τους προτεινόμενους συνδυασμούς προσαρμογέων ευρετηρίου, προτού επιλέξετε ευρετήρια, διαβάστε την ενότητα *Αντιστοίχιση βάσης και ποικιλομορφία ευρετηρίων στη σελίδα 17*.
- 5 Επιλέξτε έναν προσαρμογέα Ευρετηρίου 2 από την αναπτυσσόμενη λίστα Index 2 (Ευρετήριο 2) (i5).

- 6 Επιλέξτε ένα αρχείο δελτίου δείγματος από την αναπτυσσόμενη λίστα Manifest (Δελτίο δείγματος). Για τα δείγματα στη Δεξαμενή A απαιτείται διαφορετικό δελτίο δείγματος από τα δείγματα στη Δεξαμενή B.
- 7 Ορίστε μια επιλογή για προβολή, εκτύπωση ή αποθήκευση της διάταξης πλάκας ως αναφορά για την προετοιμασία βιβλιοθηκών:
  - ▶ Επιλέξτε το εικονίδιο  **Print** (Εκτύπωση) για να εμφανίσετε τη διάταξη της πλάκας. Επιλέξτε **Print** (Εκτύπωση) για την εκτύπωση της διάταξης πλάκας.
  - ▶ Επιλέξτε **Export** (Εξαγωγή) για να εξαγάγετε τις πληροφορίες δείγματος σε ένα εξωτερικό αρχείο.Βεβαιωθείτε ότι οι πληροφορίες του δελτίου δείγματος και του δείγματος είναι σωστές. Τυχόν εσφαλμένες πληροφορίες μπορεί να επηρεάσουν τα αποτελέσματα.
- 8 Επιλέξτε **Save Run** (Αποθήκευση εκτέλεσης).

## Εισαγωγή δειγμάτων

- 1 Επιλέξτε **Import Samples** (Εισαγωγή δειγμάτων) και μεταβείτε στη θέση του αρχείου με τις πληροφορίες του δείγματος. Υπάρχουν δύο τύποι αρχείων που μπορείτε να εισαγάγετε.
  - ▶ Επιλέξτε **Template** (Πρότυπο) στην οθόνη Create Run (Δημιουργία εκτέλεσης) για να δημιουργήσετε μια νέα διάταξη πλάκας. Το αρχείο του προτύπου περιέχει τις σωστές επικεφαλίδες στηλών για την εισαγωγή. Εισαγάγετε τις πληροφορίες δείγματος σε κάθε στήλη για τα δείγματα της εκτέλεσης. Διαγράψτε τις πληροφορίες των παραδειγμάτων στα κελιά που δεν χρησιμοποιούνται και, στη συνέχεια, αποθηκεύστε το αρχείο.
  - ▶ Χρησιμοποιήστε ένα αρχείο πληροφοριών δείγματος που έχει εξαχθεί από τη μονάδα Somatic Variant χρησιμοποιώντας τη λειτουργία Export (Εξαγωγή).
- 2 Επιλέξτε το εικονίδιο  **Print** (Εκτύπωση) για να εμφανίσετε τη διάταξη της πλάκας.
- 3 Επιλέξτε **Print** (Εκτύπωση) για να εκτυπώσετε τη διάταξη της πλάκας ως αναφορά για την προετοιμασία βιβλιοθηκών.
- 4 **[Προαιρετικό]** Επιλέξτε **Export** (Εξαγωγή) για να εξαγάγετε πληροφορίες δείγματος σε εξωτερικό αρχείο. Βεβαιωθείτε ότι οι πληροφορίες του δελτίου δείγματος και του δείγματος είναι σωστές. Τυχόν εσφαλμένες πληροφορίες μπορεί να επηρεάσουν τα αποτελέσματα.
- 5 Επιλέξτε **Save Run** (Αποθήκευση εκτέλεσης).

## Επεξεργασία εκτέλεσης

Για οδηγίες σχετικά με την επεξεργασία των πληροφοριών στην εκτέλεση πριν από την αλληλούχιση, βλ. *Οδηγός αναφοράς οργάνου NextSeq 550Dx (αρ. εγγράφου 1000000009513)*.

## Μέθοδοι ανάλυσης

Η μονάδα ανάλυσης Somatic Variant εκτελεί τα παρακάτω βήματα ανάλυσης και, στη συνέχεια, εγγράφει τα αρχεία εξόδου ανάλυσης στον φάκελο Alignment.

- ▶ Εκτελεί αποπολύπλεξη στις αναγνώσεις ευρετηρίου
- ▶ Δημιουργεί αρχεία FASTQ
- ▶ Εκτελεί ευθυγράμμιση με ένα στοιχείο αναφοράς
- ▶ Εντοπίζει παραλλαγές

## Αποπολύπλεξη

Κατά την αποπολύπλεξη γίνεται μια σύγκριση κάθε αλληλουχίας ανάγνωσης ευρετηρίου με τις αλληλουχίες ευρετηρίου που έχουν οριστεί για την εκτέλεση. Σε αυτό το βήμα δεν λαμβάνονται υπόψη τιμές ποιότητας.

Οι αλληλουχίες ευρετηρίου ταυτοποιούνται με τα παρακάτω βήματα:

- ▶ Τα δείγματα αριθμούνται από το 1 με βάση τη σειρά με την οποία αναγράφονται για την εκτέλεση.
- ▶ Ο αριθμός δείγματος 0 προορίζεται για συστάδες που δεν αντιστοιχίστηκαν σε δείγμα.
- ▶ Οι συστάδες αντιστοιχίζονται σε δείγμα όταν υπάρχει απόλυτη αντιστοιχία της αλληλουχίας ευρετηρίου ή όταν υπάρχει έως και μία αναντιστοιχία ανά ανάγνωση ευρετηρίου.

## Δημιουργία αρχείου FASTQ

Μετά την αποπολύπλεξη, το λογισμικό δημιουργεί ενδιάμεσα αρχεία ανάλυσης σε μορφή FASTQ, η οποία είναι μια μορφή κειμένου που χρησιμοποιείται για την αναπαράσταση αλληλουχιών. Τα αρχεία FASTQ περιέχουν αναγνώσεις για κάθε δείγμα και τις σχετικές βαθμολογίες ποιότητας. Οι συστάδες που δεν πέρασαν το φίλτρο αποκλείονται.

Κάθε αρχείο FASTQ περιέχει αναγνώσεις για ένα μόνο δείγμα και το όνομα του δείγματος περιλαμβάνεται στο όνομα του αρχείου FASTQ. Τα αρχεία FASTQ είναι τα βασικά δεδομένα εισόδου για την ευθυγράμμιση. Δημιουργούνται οκτώ αρχεία FASTQ ανά δείγμα ανά δεξαμενή ολιγονουκλεοτιδίων, τέσσερα από την Ανάγνωση 1 και τέσσερα από την Ανάγνωση 2, τα οποία δημιουργούν ένα σύνολο από 16 αρχεία FASTQ ανά δείγμα.

## Ευθυγράμμιση

Στο βήμα της ευθυγράμμισης, ο αλγόριθμος Smith-Waterman με ζώνες ευθυγραμμίζει συστάδες από κάθε δείγμα βάσει αλληλουχιών αμπλικονίων που καθορίζονται στο αρχείο δελτίου δείγματος.

Ο αλγόριθμος Smith-Waterman με ζώνες εκτελεί ημικαθολικές ευθυγραμμίσεις αλληλουχιών για τον καθορισμό παρεμφερών περιοχών μεταξύ δύο αλληλουχιών. Αντί της σύγκρισης του συνόλου της αλληλουχίας, ο αλγόριθμος Smith-Waterman συγκρίνει τμήματα όλων των πιθανών τιμών μήκους.

Κάθε ανάγνωση συζευγμένων άκρων αξιολογείται ως προς την ευθυγράμμισή της με τις αλληλουχίες της αντίστοιχης κεφαλής για τη συγκεκριμένη ανάγνωση.

- ▶ Η ανάγνωση 1 αξιολογείται ως προς το αντίστροφο συμπλήρωμα του Downstream Locus-Specific Oligos (DLSO).
- ▶ Η ανάγνωση 2 αξιολογείται ως προς το Upstream Locus-Specific Oligos (ULSO).
- ▶ Εάν η αρχή μιας ανάγνωσης ταιριάζει με μια αλληλουχία κεφαλής με όχι περισσότερες από τρεις διαφορές (αναντιστοιχίες ή μεταβολές λόγω αρχικών indel), το πλήρες μήκος της ανάγνωσης ευθυγραμμίζεται ως προς τον στόχο του αμπλικονίου αυτής της αλληλουχίας.
- ▶ Τα indel με DLSO και ULSO δεν παρατηρούνται, δεδομένης της χημείας του προσδιορισμού.

Οι ευθυγραμμίσεις φιλτράρονται από τα αποτελέσματα της ευθυγράμμισης βάσει των ποσοστών αναντιστοιχίας, είτε σε σχέση με την περιοχή ενδιαφέροντος είτε με ολόκληρο το αμπλικόνιο, ανάλογα με το μήκος του αμπλικονίου. Οι φιλτραρισμένες ευθυγραμμίσεις εγγράφονται ως αρχεία ευθυγράμμισης ως μη ευθυγραμμισμένα στοιχεία και δεν χρησιμοποιούνται στην αντιστοίχιση παραλλαγής.

## Αντιστοίχιση παραλλαγής



Το PISCES Variant Caller, το οποίο έχει αναπτυχθεί από την Illumina, αναγνωρίζει παραλλαγές που υπάρχουν στο δείγμα DNA.

Ο αντιστοιχιστής παραλλαγής PISCES αναγνωρίζει SNV, MNV και μικρά indel σε τρία βήματα:

- ▶ Εξετάζει ξεχωριστά κάθε θέση στο γονιδίωμα αναφοράς
- ▶ Μετρά τις βάσεις στη δεδομένη θέση για ευθυγραμμισμένες αναγνώσεις που υπερκαλύπτουν τη θέση
- ▶ Υπολογίζει μια βαθμολογία ποιότητας που υπολογίζει την ποιότητα της αντιστοίχισης με το μοντέλο Poisson. Οι παραλλαγές με βαθμολογία ποιότητας κάτω από Q30 αποκλείονται.

Οι παραλλαγές αντιστοιχίζονται πρώτα για κάθε δεξαμενή χωριστά. Στη συνέχεια, οι παραλλαγές από κάθε δεξαμενή συγκρίνονται και συνδυάζονται σε ένα ενιαίο αρχείο εξόδου. Εάν μια παραλλαγή υπάρχει και στις δύο δεξαμενές και διέλθει από όλα τα φίλτρα που αναγράφονται στην ενότητα *Επισημειώσεις αρχείου VCF στη σελίδα 13*, η παραλλαγή αυτή επισημαίνεται ως PASS στο αρχείο αντιστοίχισης παραλλαγής (VCF).

## Προβολή δεδομένων εκτέλεσης και δείγματος

- 1 Από τον πίνακα εργαλείων Local Run Manager, κάντε κλικ στο όνομα της εκτέλεσης.
- 2 Από την καρτέλα Run Overview (Επισκόπηση εκτέλεσης), δείτε τις μετρήσεις της εκτέλεσης αλληλούχισης.
- 3 **[Προαιρετικό]** Κάντε κλικ στο εικονίδιο **Copy to Clipboard**  (Αντιγραφή στο πρόχειρο) για να αντιγράψετε τη διαδρομή στον φάκελο εξόδου εκτέλεσης.
- 4 Κάντε κλικ στην καρτέλα Sequencing Information (Πληροφορίες αλληλούχισης) για να δείτε τις παραμέτρους της εκτέλεσης και τις πληροφορίες για τα αναλώσιμα.
- 5 Κάντε κλικ στην καρτέλα Samples and Results (Δείγματα και αποτελέσματα) για να δείτε τη θέση της αναφοράς της ανάλυσης.
  - ▶ Για επανειλημμένη ανάλυση, αναπτύξτε την αναπτυσσόμενη λίστα Select Analysis (Επιλογή ανάλυσης) και επιλέξτε την κατάλληλη ανάλυση.
- 6 Κάντε κλικ στο εικονίδιο **Copy to Clipboard**  (Αντιγραφή στο πρόχειρο) για να αντιγράψετε τη διαδρομή του φακέλου Analysis.

Για περισσότερες πληροφορίες σχετικά με τις καρτέλες Run Overview (Επισκόπηση εκτέλεσης) και Sequencing Information (Πληροφορίες αλληλούχισης), αλλά και για τον τρόπο επανατοποθέτησης μιας ανάλυσης σε ουρά, βλ. *Οδηγός αναφοράς οργάνου NextSeq 550Dx (αρ. εγγράφου 1000000009513)*.

## Αναφορά ανάλυσης

Τα αποτελέσματα της ανάλυσης συνοψίζονται στην καρτέλα Samples and Results (Δείγματα και αποτελέσματα), αλλά και με τη μορφή μιας συγκεντρωτικής αναφοράς στον φάκελο Alignment. Διατίθεται επίσης μια αναφορά για κάθε δείγμα με τη μορφή αρχείου PDF.



## Πληροφορίες καρτέλας Samples and Results (Δείγματα και αποτελέσματα)

1 Κάντε κλικ σε ένα δείγμα της λίστας για να δείτε την αναφορά δείγματος.

Πίνακας 1 Πληροφορίες εκτέλεσης και δείγματος

Επικεφαλίδα στήλης	Περιγραφή
Κατάσταση εκτέλεσης	Αναφέρει το κατά πόσον η εκτέλεση αλληλούχισης ήταν επιτυχής ή απέτυχε.
Total Yield (GB) [Συνολική απόδοση (GB)]	Αριθμός των βάσεων που αντιστοιχίστηκαν στην εκτέλεση αλληλούχισης. Εμφανίζει το όριο επιτυχίας και την κατάσταση επιτυχίας ή αποτυχίας.
% ≥ Q30	Το ποσοστό των αναγνώσεων στην εκτέλεση αλληλούχισης με βαθμολογία ποιότητας 30 (Q30) και άνω. Εμφανίζει το όριο επιτυχίας και την κατάσταση επιτυχίας ή αποτυχίας.
Sample Name (Όνομα δείγματος)	Το όνομα του δείγματος που εκχωρήθηκε κατά τη δημιουργία της εκτέλεσης.
Total PF Reads (Σύνολο αναγνώσεων που διέρχονται από το φίλτρο)	Ο συνολικός αριθμός των αναγνώσεων που διέρχονται από το φίλτρο.
Read 1% ≥ Q30 (Ανάγνωση 1% ≥ Q30)	Το ποσοστό των αναγνώσεων στην Ανάγνωση 1 με βαθμολογία ποιότητας 30 (Q30) και άνω για το δείγμα.
Read 2% ≥ Q30 (Ανάγνωση 2% ≥ Q30)	Το ποσοστό των αναγνώσεων στην Ανάγνωση 2 με βαθμολογία ποιότητας 30 (Q30) και άνω για το δείγμα.
Autosome Call Rate (Ποσοστό αντιστοίχισης αυτοσωμάτων)	Ο αριθμός των θέσεων γονιδιώματος στο σύνολο των αυτοσωμάτων (χρωμοσώματα 1 έως 22) που πληρούν ένα προκαθορισμένο όριο τιμής αξιοπιστίας, διαιρούμενος με τον συνολικό αριθμό των αυτοσωμικών θέσεων γονιδιώματος που ερευνήθηκαν. Το ποσοστό αντιστοίχισης περιγράφεται σε μια βάση ανά δείγμα και αναφέρεται ως ποσοστό που υπολογίζεται ως μείον 1 (αριθμός αυτοσωμικών θέσεων με μη ολοκληρωμένες αντιστοιχίσεις διαιρούμενος με τον συνολικό αριθμό των αυτοσωμικών θέσεων που υπεβλήθησαν σε αλληλούχιση).

Πίνακας 2 Πληροφορίες αναφοράς δείγματος

Επικεφαλίδα στήλης	Περιγραφή
Δείγμα	Το όνομα του δείγματος που εκχωρήθηκε κατά τη δημιουργία της εκτέλεσης.
Report Date (Ημερομηνία αναφοράς)	Η ημερομηνία που δημιουργήθηκε η αναφορά.
Sample Information (Πληροφορίες δείγματος)	Το αναγνωριστικό του δείγματος που εκχωρήθηκε κατά τη δημιουργία της εκτέλεσης, το σύνολο των αναγνώσεων που διήλθαν από το φίλτρο στο δείγμα, το ποσοστό των αναγνώσεων με βαθμολογία ποιότητας 30 (Q30) και άνω, καθώς και το ποσοστό αντιστοίχισης αυτοσωμάτων.
Amplicon Summary (Σύνοψη αμπλικονίων)	Ο συνολικός αριθμός περιοχών αμπλικονίου που υπεβλήθησαν σε αλληλούχιση, καθώς και το συνολικό μήκος σε ζεύγη βάσεων αμπλικονίων που υπεβλήθησαν σε αλληλούχιση στις περιοχές στόχο, για το δείγμα στη Δεξαμενή Α και στη Δεξαμενή Β, και το αρχείο του δελτίου δείγματος που χρησιμοποιείται για κάθε δεξαμενή. Στο αρχείο δελτίου δείγματος καθορίζεται το γονιδίωμα αναφοράς και οι στοχευόμενες περιοχές αναφοράς που χρησιμοποιούνται στο βήμα ευθυγράμμισης.
Read Level Statistics (Στατιστικά επιπέδου ανάγνωσης)	Ο αριθμός και το ποσοστό αναγνώσεων για το δείγμα που καλύπτουν κάθε θέση στην αναφορά, για την Ανάγνωση 1 και την Ανάγνωση 2 στη Δεξαμενή Α και στη Δεξαμενή Β.

Επικεφαλίδα στήλης	Περιγραφή
Variants Summary (Σύνοψη παραλλαγών)	Αριθμός SNV, ενθέσεις και διαγραφές που ανιχνεύτηκαν για το δείγμα που πέρασε τις προτεινόμενες τιμές, για να καθοριστεί το κατά πόσον τα αποτελέσματα της ποιότητας βρίσκονται εντός ενός αποδεκτού εύρους.
Coverage Summary (Σύνοψη κάλυψης)	Ο συνολικός αριθμός των ευθυγραμμισμένων βάσεων διαιρούμενος με το μέγεθος της στοχευόμενης περιοχής, και το ποσοστό των περιοχών αμπλικονίων με τιμές κάλυψης μεγαλύτερες από το κατώτερο όριο κάλυψης της μέσης κάλυψης αμπλικονίων 0,2 *, για το δείγμα στη Δεξαμενή A και στη Δεξαμενή B.
Coverage Plots (Γραφήματα κάλυψης)	Το γράφημα κάλυψης ανά περιοχή αμπλικονίων δείχνει την κάλυψη στις περιοχές αμπλικονίων για το δείγμα. Οι περιοχές με τιμές κάλυψης χαμηλότερες από το όριο κάλυψης επισημαίνονται με κόκκινο χρώμα. Ο μέσος όρος όλων των τιμών υποδεικνύεται από μια γραμμή πορτοκαλί χρώματος. Παρέχεται ένα γράφημα για την κάλυψη της Δεξαμενής A και της Δεξαμενής B.
Εκδόσεις λογισμικού	Οι εκδόσεις λογισμικού κατά την αλληλούχιση του δείγματος. Περιλαμβάνει το NextSeq 550Dx Operating Software (NOS), το λογισμικό του Local Run Manager, το λογισμικό RTA, και την έκδοση του Somatic Variant Module.

## Αρχεία εξόδου ανάλυσης

Τα παρακάτω αρχεία εξόδου ανάλυσης δημιουργούνται για τη μονάδα ανάλυσης Somatic Variant και παρέχουν αποτελέσματα ανάλυσης για ευθυγράμμιση και αντιστοίχιση παραλλαγής. Τα αρχεία εξόδου ανάλυσης βρίσκονται στον φάκελο Alignment.

Όνομα αρχείου	Περιγραφή
Αποπολύπλεξη (*.txt)	Ενδιάμεσα αρχεία που περιέχουν συνοπτικά αποτελέσματα της αποπολύπλεξης.
FASTQ (*.fastq.gz)	Ενδιάμεσα αρχεία που περιέχουν αντιστοιχίσεις βάσης με βαθμολογία ποιότητας. Τα αρχεία FASTQ είναι τα βασικά δεδομένα εισόδου για το βήμα της ευθυγράμμισης.
Αρχεία ευθυγράμμισης σε μορφή BAM (*.bam)	Περιέχει ευθυγραμμισμένες αναγνώσεις για ένα συγκεκριμένο δείγμα.
Αρχεία αντιστοίχισης παραλλαγής ανά δεξαμενή σε μορφή VCF (*.vcf)	Περιέχει παραλλαγές που αντιστοιχίστηκαν σε κάθε θέση είτε από τη δεξαμενή ευθείας είτε αντίστροφης αντιστοίχισης.
Αρχεία αντιστοίχισης παραλλαγής σε μορφή VCF γονιδιώματος (*.genome.vcf.gz)	Περιέχει τον γονότυπο για κάθε θέση, είτε με αντιστοίχιση παραλλαγής είτε με αντιστοίχιση αναφοράς.
Αρχεία αντιστοίχισης παραλλαγής ομοφωνίας σε μορφή VCF (*.vcf.gz)	Περιέχει παραλλαγές που αντιστοιχίστηκαν σε κάθε θέση και από τις δύο δεξαμενές.
AmpliconCoverage_M1.tsv	Περιέχει πληροφορίες για την κάλυψη ανά αμπλικόνιο ανά δείγμα για κάθε δελτίο δείγματος που παρέχεται. Το M# δηλώνει τον αριθμό του δελτίου δείγματος.

## Μορφή αρχείου αποπολύπλεξης

Η διαδικασία αποπολύπλεξης διαβάζει την αλληλουχία ευρετηρίου που συνδέεται με κάθε συστάδα, για να καθοριστεί από ποιο δείγμα προέρχεται η συστάδα. Η αντιστοίχιση μεταξύ των συστάδων και του αριθμού του δείγματος εγγράφεται σε ένα αρχείο αποπολύπλεξης (\*.demux) για κάθε πλακίδιο στην κυψελίδα ροής.

Η μορφή του ονόματος του αρχείου αποπολύπλεξης είναι **s\_1\_X.demux**, όπου X είναι ο αριθμός του πλακιδίου.

Τα αρχεία αποπολύπλεξης ξεκινούν με μια επικεφαλίδα:

- ▶ Έκδοση (ακέραιος 4 byte), αυτήν τη στιγμή 1
- ▶ Αριθμός συστάδων (ακέραιος 4 byte)

Το υπόλοιπο αρχείο αποτελείται από τους αριθμούς των δειγμάτων για κάθε συστάδα του πλακιδίου.

Όταν ολοκληρωθεί το βήμα της αποπολύπλεξης, το λογισμικό δημιουργεί ένα αρχείο αποπολύπλεξης με όνομα **DemultiplexSummaryF1L1.txt**.

- ▶ Στο όνομα του αρχείου, το **F1** αναπαριστά τον αριθμό της κυψελίδας ροής.
- ▶ Στο όνομα του αρχείου, το **L1** αναπαριστά τον αριθμό της λωρίδας.
- ▶ Η αποπολύπλεξη έχει ως αποτέλεσμα τη δημιουργία ενός πίνακα με μία γραμμή ανά πλακίδιο και μία στήλη ανά δείγμα, συμπεριλαμβανομένου του δείγματος 0.
- ▶ Οι πιο συχνές αλληλουχίες σε αναγνώσεις ευρετηρίου.

## Μορφή αρχείου FASTQ

Η μορφή FASTQ είναι μια μορφή αρχείων βάσει κειμένου που περιέχουν αντιστοιχίσεις βάσης και τιμές ποιότητας ανά ανάγνωση. Κάθε εγγραφή περιέχει τέσσερις γραμμές:

- ▶ Το αναγνωριστικό
- ▶ Την αλληλουχία
- ▶ Το σύμβολο συν (+)
- ▶ Τις βαθμολογίες ποιότητας Phred σε κωδικοποιημένη μορφή ASCII + 33

Η μορφή του αναγνωριστικού έχει ως εξής:

**@Όργανο:Αναγνωριστικό εκτέλεσης:Αναγνωριστικό κυψέλης ροής:Λωρίδα:Πλακίδιο:X:Y Αρ. ανάγνωσης:Επισημάνση φίλτρου:0:Αριθμός δείγματος**

Παράδειγμα:

```
@SIM:1:FCX:1:15:6329:1045 1:N:0:2
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
+
<>;##=><9=AAAAAAAAA9#:<#<;<<<?????#=#
```

## Μορφή αρχείου BAM

Τα αρχεία BAM (\*.bam) είναι η συμπιεσμένη δυαδική μορφή των αρχείων SAM που χρησιμοποιούνται για την αναπαράσταση ευθυγραμμισμένων αλληλουχιών έως και 128 Mb. Οι μορφές SAM και BAM περιγράφονται λεπτομερώς στο [samtools.github.io/hts-specs/SAMv1.pdf](https://samtools.github.io/hts-specs/SAMv1.pdf).

Η μορφή του ονόματος των αρχείων BAM είναι **SampleName\_S#.bam**, όπου το # είναι ο αριθμός του δείγματος που ορίζεται από τη σειρά με την οποία αναφέρονται τα δείγματα για την εκτέλεση.

Τα αρχεία BAM έχουν μια ενότητα επικεφαλίδας και μια ενότητα ευθυγράμμισης:

- ▶ **Header** (Επικεφαλίδα)—Περιέχει πληροφορίες για ολόκληρο το αρχείο, όπως το όνομα και το μήκος του δείγματος, αλλά και τη μέθοδο ευθυγράμμισης. Οι ευθυγραμμίσεις στην ενότητα των ευθυγραμμίσεων σχετίζονται με συγκεκριμένες πληροφορίες στην ενότητα επικεφαλίδας.
- ▶ **Alignments** (Ευθυγραμμίσεις)—Περιέχει το όνομα, την αλληλουχία και την ποιότητα της ανάγνωσης, τις πληροφορίες της ευθυγράμμισης και προσαρμοσμένες ετικέτες. Το όνομα της ανάγνωσης περιλαμβάνει το χρωμόσωμα, τις συντεταγμένες έναρξης, την ποιότητα της ευθυγράμμισης και τη συμβολοσειρά του περιγραφέα αντιστοιχίας.

Στην ενότητα ευθυγράμμισης περιλαμβάνονται οι παρακάτω πληροφορίες για κάθε ανάγνωση ή ζεύγος ανάγνωσης:

- ▶ **AS:** Ποιότητα ευθυγράμμισης συζευγμένων άκρων
- ▶ **BC:** Ετικέτα γραμμωτού κωδικού, η οποία υποδεικνύει το αναγνωριστικό του αποπολυπλεγμένου δείγματος που σχετίζεται με την ανάγνωση.
- ▶ **SM:** Ποιότητα ανάγνωσης μονού άκρου.
- ▶ **XC:** Συμβολοσειρά περιγραφέα αντιστοιχίας
- ▶ **XN:** Ετικέτα ονόματος αμπλικονίου, η οποία καταγράφει το αναγνωριστικό του αμπλικονίου που σχετίζεται με την ανάγνωση

Τα αρχεία ευρετηρίου BAM (\*.bam.bai) παρέχουν ένα ευρετήριο του αντίστοιχου αρχείου BAM.

## Μορφή αρχείου VCF

Η μορφή VCF (Variant Call Format) είναι μια συνηθισμένη μορφή αρχείου που έχει αναπτυχθεί από την επιστημονική κοινότητα που ασχολείται με τα γονιδιώματα. Περιλαμβάνει πληροφορίες για τις παραλλαγές που βρέθηκαν σε συγκεκριμένες θέσεις σε ένα γονιδίωμα αναφοράς. Τα αρχεία VCF λήγουν με το επίθημα ".vcf"

Η επικεφαλίδα του αρχείου VCF περιλαμβάνει την έκδοση της μορφής του αρχείου VCF και την έκδοση του αντιστοιχιστή παραλλαγής και αναφέρει τις επισημειώσεις που χρησιμοποιούνται στο υπόλοιπο αρχείο. Η επικεφαλίδα VCF περιλαμβάνει και το αρχείο του γονιδιώματος αναφοράς και το αρχείο BAM. Η τελευταία γραμμή στην επικεφαλίδα περιέχει τις επικεφαλίδες των στηλών για τις γραμμές δεδομένων. Κάθε μία από τις γραμμές δεδομένων του αρχείου VCF περιέχει πληροφορίες για μία παραλλαγή.

## Επικεφαλίδες αρχείου VCF

Επικεφαλίδα	Περιγραφή
<b>CHROM</b>	Το χρωμόσωμα του γονιδιώματος αναφοράς. Τα χρωμοσώματα εμφανίζονται με την ίδια σειρά όπως στο αρχείο αναφοράς FASTQ.
<b>POS</b>	Η θέση μονής βάσης της παραλλαγής στο χρωμόσωμα αναφοράς. Για SNP, αυτή η θέση είναι η βάση αναφοράς με την παραλλαγή. Για indel ή διαγραφές, αυτή η θέση είναι βάση αναφοράς που βρίσκεται αμέσως πριν από την παραλλαγή.
<b>ID</b>	Ο αριθμός rs για την παραλλαγή, ο οποίος ελήφθη από το dbSNP.txt, κατά περίπτωση. Εάν υπάρχουν περισσότεροι αριθμοί rs σε αυτήν τη θέση, οι τιμές στη λίστα χωρίζονται με ερωτηματικό. Εάν δεν υπάρχει εισαγωγή dbSNP σε αυτήν τη θέση, χρησιμοποιείται ένας δείκτης απουσίας τιμής ('.').
<b>REF</b>	Ο γονότυπος αναφοράς. Για παράδειγμα, μια διαγραφή ενός T αναπαριστάται ως TT αναφοράς και ως εναλλακτικό T. Μια παραλλαγή μονού νουκλεοτιδίου A σε T αναπαριστάται ως A αναφοράς και ως εναλλακτικό T.
<b>ALT</b>	Τα αλληλόμορφα που διαφέρουν από την ανάγνωση αναφοράς. Για παράδειγμα, η ένθεση ενός T αναπαριστάται ως A αναφοράς και ως εναλλακτικό AT. Μια παραλλαγή μονού νουκλεοτιδίου A σε T αναπαριστάται ως A αναφοράς και εναλλακτικό T.
<b>QUAL</b>	Μια βαθμολογία ποιότητας με κλίμακα Phred που εκχωρείται από τον αντιστοιχιστή παραλλαγής. Οι υψηλότερες βαθμολογίες δηλώνουν μεγαλύτερη αξιοπιστία στην παραλλαγή και χαμηλότερη πιθανότητα σφαλμάτων. Για βαθμολογία ποιότητας Q, η εκτιμώμενη πιθανότητα σφάλματος είναι $10^{-(Q/10)}$ . Για παράδειγμα, το σύνολο αντιστοιχίσεων Q30 έχει ποσοστό σφάλματος 0,1%. Πολλοί αντιστοιχιστές παραλλαγής εκχωρούν βαθμολογίες ποιότητας με βάση τα στατιστικά τους μοντέλα, τα οποία είναι υψηλά σε σχέση με το ποσοστό σφάλματος που παρατηρείται.

## Επισημειώσεις αρχείου VCF

Επικεφαλίδα	Περιγραφή
<b>FILTER (ΦΙΛΤΡΟ)</b>	<p>Εάν το στοιχείο διέλθει από όλα τα φίλτρα, στη στήλη των φίλτρων προστίθεται η επισήμανση <b>PASS</b>.</p> <ul style="list-style-type: none"> <li>• <b>LowDP</b>—Εφαρμόζεται σε θέσεις με βάθος κάλυψης κάτω από 450x σε οποιαδήποτε από τις δύο δεξαμενές. Για θέσεις αμπλικονίων που καλύπτονται τόσο από την ευθεία όσο και από την αντίστροφη ανάγνωση, αυτό ισοδυναμεί με κάλυψη μεμονωμένης ανάγνωσης 900x.</li> <li>• <b>LowGQ</b>—Η ποιότητα γονότυπου (GQ) είναι χαμηλότερη από μια τιμή αποκοπής.</li> <li>• <b>q30</b>—Βαθμολογία ποιότητας &lt; 30.</li> <li>• <b>LowVariantFreq</b>—Η συχνότητα παραλλαγής είναι χαμηλότερη από το ορισμένο κατώφλι.</li> <li>• <b>PB</b>—Μεροληψία κεφαλής δεξαμενής. Η παραλλαγή δεν βρέθηκε ή βρέθηκε με χαμηλή συχνότητα σε μία ή δύο δεξαμενές κεφαλής.</li> <li>• <b>R3x6</b>—Ο αριθμός των παρακείμενων επαναλήψεων (με μήκος από 1 έως 3 bp) στις αντιστοιχίσεις παραλλαγής ≥ 6.</li> <li>• <b>SB</b>—Η μεροληψία αλυσίδας είναι μεγαλύτερη από το ορισμένο κατώφλι.</li> </ul>
<b>INFO (ΠΛΗΡΟΦΟΡΙΕΣ)</b>	<p>Στις πιθανές καταχωρήσεις στη στήλη INFO (ΠΛΗΡΟΦΟΡΙΕΣ) περιλαμβάνονται τα εξής:</p> <ul style="list-style-type: none"> <li>• <b>AC</b>—Αριθμός αλληλόμορφων στους γονότυπους για κάθε αλληλόμορφο ALT, με τη σειρά που αναγράφονται.</li> <li>• <b>AF</b>—Συχνότητα αλληλόμορφων για κάθε αλληλόμορφο ALT, με τη σειρά που αναγράφονται.</li> <li>• <b>AN</b>—Ο συνολικός αριθμός αλληλόμορφων στους γονότυπους της αντιστοίχισης.</li> <li>• <b>CD</b>—Μια επισήμανση που υποδεικνύει ότι το SNP εμφανίζεται εντός της κωδικοποιητικής αλληλουχίας τουλάχιστον μίας εισαγωγής RefGene.</li> <li>• <b>DP</b>—Το βάθος (αριθμός αντιστοιχίσεων βάσης που ευθυγραμμίζονται σε μια θέση και χρησιμοποιούνται στην αντιστοίχιση παραλλαγής).</li> <li>• <b>Exon</b>—Μια λίστα περιοχών εξονίου, διαχωρισμένων με κόμμα, οι οποίες προέρχονται από την ανάγνωση του RefGene.</li> <li>• <b>FC</b>—Λειτουργική συνέπεια.</li> <li>• <b>GI</b>—Μια λίστα των αναγνωριστικών των γονιδίων, διαχωρισμένων με κόμμα, τα οποία προέρχονται από την ανάγνωση του RefGene.</li> <li>• <b>QD</b>—Αξιοπιστία παραλλαγής/Ποιότητα ανά βάθος.</li> <li>• <b>TI</b>—Μια λίστα των αναγνωριστικών των μεταγραφών, διαχωρισμένων με κόμμα, τα οποία προέρχονται από την ανάγνωση του RefGene.</li> </ul>
<b>FORMAT (ΜΟΡΦΗ)</b>	<p>Η στήλη της μορφής παραθέτει πεδία που διαχωρίζονται με άνω-κάτω τελεία. Για παράδειγμα, GT:GQ. Η λίστα των πεδίων εξαρτάται από τον αντιστοιχιστή παραλλαγής που χρησιμοποιείται. Στα διαθέσιμα πεδία περιλαμβάνονται τα εξής:</p> <ul style="list-style-type: none"> <li>• <b>AD</b>—Εισαγωγή μορφής X,Y, όπου X είναι ο αριθμός των αντιστοιχίσεων γονιδιώματος αναφοράς, και Y ο αριθμός των αντιστοιχίσεων εναλλακτικού γονιδιώματος.</li> <li>• <b>DP</b>—Κατά προσέγγιση βάθος ανάγνωσης. Οι αναγνώσεις με MQ=255 ή με μη ικανοποιητικά ζεύγη φιλτράρονται.</li> <li>• <b>GQ</b>—Ποιότητα γονότυπου.</li> <li>• <b>GQX</b>—Ποιότητα γονότυπου. Το GQX είναι το ελάχιστο της τιμής GQ και της στήλης QUAL. Σε γενικές γραμμές, αυτές οι τιμές είναι παρεμφερείς. Σε περίπτωση επιλογής της ελάχιστης τιμής, το GQX γίνεται η πιο συντηρητική μέτρηση της ποιότητας του γονότυπου.</li> <li>• <b>GT</b>—Γονότυπος. Το 0 αντιστοιχεί στη βάση αναφοράς, το 1 αντιστοιχεί στην πρώτη εισαγωγή στη στήλη ALT, και ούτω καθεξής. Η κάθετος (/) υποδεικνύει ότι δεν διατίθενται πληροφορίες phasing (καθυστέρησης βάσης).</li> <li>• <b>NL</b>—Επίπεδο θορύβου. Μια εκτίμηση του θορύβου στην αντιστοίχιση βάσης σε αυτήν τη θέση.</li> <li>• <b>PB</b>—Μεροληψία κεφαλής δεξαμενής. Οι τιμές που είναι πιο κοντά στο 0 δηλώνουν μεγαλύτερη μεροληψία προς μία δεξαμενή κεφαλής και χαμηλότερη αξιοπιστία σε μια αντιστοίχιση παραλλαγής.</li> <li>• <b>SB</b>—Μεροληψία αλυσίδας σε αυτήν τη θέση. Οι μεγαλύτερες αρνητικές τιμές δηλώνουν μικρότερη μεροληψία, οι τιμές κοντά στο 0 δηλώνουν μεγαλύτερη μεροληψία.</li> <li>• <b>VF</b>—Συχνότητα παραλλαγής. Το ποσοστό των αναγνώσεων που υποστηρίζουν το εναλλακτικό αλληλόμορφο.</li> </ul>
<b>SAMPLE (ΔΕΙΓΜΑ)</b>	<p>Η στήλη δείγματος δίνει τις τιμές που προσδιορίζονται στη στήλη FORMAT (ΜΟΡΦΗ).</p>

## Αρχεία γονιδιώματος VCF

Τα αρχεία γονιδιώματος VCF (gVCF) είναι αρχεία VCF v4.1 που ακολουθούν ένα σύνολο συμβάσεων για την αναπαράσταση όλων των θέσεων του γονιδιώματος σε μια εύλογα μικρή μορφή. Τα αρχεία gVCF (\*.genome.vcf.gz) περιλαμβάνουν όλες τις θέσεις εντός της περιοχής ενδιαφέροντος σε ένα αρχείο για κάθε δείγμα.

Το αρχείο gVCF εμφανίζει μη αντιστοιχίσεις σε θέσεις που δεν διέρχονται από όλα τα φίλτρα. Μια ετικέτα γονότυπου (GT) *./.* υποδεικνύει μη αντιστοίχιση.

Για περισσότερες πληροφορίες, βλ. [sites.google.com/site/gvcftools/home/about-gvcf](https://sites.google.com/site/gvcftools/home/about-gvcf).

## Αρχεία ανά δεξαμενή και αρχεία ομοφωνίας VCF

Η ροή εργασιών του Somatic Variant δημιουργεί δύο σετ αρχείων αντιστοίχισης παραλλαγής.

- ▶ **Αρχεία VCF ανά δεξαμενή**—Περιέχει παραλλαγές που αντιστοιχίστηκαν στη δεξαμενή ευθείας ή αντίστροφης διαδικασίας. Τα αρχεία ανά δεξαμενή εγγράφονται στον φάκελο VariantCallingLogs.
- ▶ **Αρχεία ομοφωνίας VCF**—Περιέχουν παραλλαγές που αντιστοιχίστηκαν και από τις δύο δεξαμενές. Τα αρχεία ομοφωνίας εγγράφονται στον φάκελο Alignment.

Τα αρχεία ανά δεξαμενή και τα αρχεία ομοφωνίας VCF περιλαμβάνουν τόσο αρχεία VCF (\*.vcf) όσο και αρχεία gVCF (\*.genome.vcf), και χρησιμοποιούν την ακόλουθη σύμβαση ονοματοθεσίας, όπου το S# δηλώνει τη σειρά με την οποία το δείγμα αναγράφεται για την εκτέλεση:

- ▶ **Αναφορές για όλες τις θέσεις**—SampleName\_S#.genome.vcf
- ▶ **Αναφορές μόνο για παραλλαγές**—SampleName\_S#.vcf

Το λογισμικό συγκρίνει τα αρχεία VCF ανά δεξαμενή και συνδυάζει τα δεδομένα κάθε θέσης για τη δημιουργία ενός αρχείου ομοφωνίας VCF για το δείγμα.

Οι αντιστοιχίσεις παραλλαγής από κάθε δεξαμενή συγχωνεύονται σε αρχεία ομοφωνίας VCF με τα παρακάτω κριτήρια.

Κριτήρια	Αποτέλεσμα
Μια αντιστοίχιση αναφοράς σε κάθε δεξαμενή	Αντιστοίχιση αναφοράς
Μια αντιστοίχιση αναφοράς σε μία δεξαμενή και μια αντιστοίχιση παραλλαγής στην άλλη δεξαμενή	Αντιστοίχιση παραλλαγής με φίλτρο
Ταιριαστές αντιστοιχίσεις παραλλαγής με παρεμφερείς συχνότητες σε κάθε δεξαμενή	Αντιστοίχιση παραλλαγής
Ταιριαστές αντιστοιχίσεις παραλλαγής με ελαφρώς διαφορετικές συχνότητες σε κάθε δεξαμενή	Αντιστοίχιση παραλλαγής με φίλτρο
Αταιρίαστες αντιστοιχίσεις παραλλαγής σε κάθε δεξαμενή	Αντιστοίχιση παραλλαγής με φίλτρο

Οι μετρήσεις από κάθε δεξαμενή συγχωνεύονται με τη χρήση των παρακάτω τιμών.

Μέτρηση	Τιμή
Βάθος	Άθροισμα τιμών βάθους και από τις δύο δεξαμενές
Συχνότητα παραλλαγής	Το σύνολο των παραλλαγών διαιρούμενο με το συνολικό βάθος κάλυψης
Βαθμολογία ποιότητας	Η ελάχιστη τιμή και των δύο δεξαμενών

## Αρχείο κάλυψης αμπλικονίου

Για κάθε αρχείο δελτίου δείγματος δημιουργείται ένα αρχείο κάλυψης αμπλικονίου. Το M# στο όνομα του αρχείου δηλώνει τον αριθμό του δελτίου δείγματος.

Κάθε αρχείο περιλαμβάνει μια γραμμή επικεφαλίδας που περιέχει τα αναγνωριστικά του δείγματος, τα οποία σχετίζονται με το δελτίο δείγματος. Το αρχείο περιλαμβάνει τις παρακάτω πληροφορίες.

- ▶ Το αναγνωριστικό στόχου όπως αναγράφεται στο δελτίο δείγματος.
- ▶ Το βάθος κάλυψης των αναγνώσεων που διέρχονται από το φίλτρο.

## Συμπληρωματικά αρχεία εξόδου

Τα παρακάτω αρχεία εξόδου παρέχουν συμπληρωματικές πληροφορίες ή συνοψίζουν τα αποτελέσματα της εκτέλεσης και τα σφάλματα της ανάλυσης. Παρόλο που τα αρχεία αυτά δεν είναι απαραίτητα για την αξιολόγηση των αποτελεσμάτων της ανάλυσης, μπορούν να χρησιμοποιηθούν για διαδικασίες αντιμετώπισης προβλημάτων. Όλα τα αρχεία βρίσκονται στον φάκελο Alignment εκτός και αν ορίζεται διαφορετικά.

Όνομα αρχείου	Περιγραφή
AnalysisLog.txt	Αρχείο καταγραφής επεξεργασίας που περιγράφει κάθε βήμα που έλαβε χώρα κατά τη διάρκεια της ανάλυσης για τον τρέχοντα φάκελο της εκτέλεσης. Αυτό το αρχείο δεν περιέχει μηνύματα σφάλματος. Βρίσκεται στον φάκελο Alignment.
AnalysisError.txt	Αρχείο καταγραφής επεξεργασίας που περιέχει τυχόν σφάλματα που έλαβαν χώρα κατά τη διάρκεια της ανάλυσης. Αυτό το αρχείο θα είναι κενό αν δεν έλαβαν χώρα σφάλματα. Βρίσκεται στον φάκελο Alignment.
DemultiplexSummaryF1L1#.txt	Αναφέρει τα αποτελέσματα της αποπολύπλεξης σε έναν πίνακα με μία γραμμή ανά πλακίδιο και μία στήλη ανά δείγμα. Το # αναπαριστά τη λωρίδα 1, 2, 3, ή 4 της κυψελίδας ροής. Βρίσκεται στον φάκελο Alignment.
AmpliconRunStatistics.xml	Περιέχει συνοπτικά στατιστικά στοιχεία για τη συγκεκριμένη εκτέλεση. Βρίσκεται στον φάκελο Alignment.

## Φάκελος ανάλυσης

Ο φάκελος ανάλυσης περιέχει τα αρχεία που δημιουργούνται από το λογισμικό του Local Run Manager.


Η σχέση μεταξύ του φακέλου εξόδου και του φακέλου ανάλυσης συνοψίζεται ως εξής:


- ▶ Κατά τη διάρκεια της αλληλούχησης, το Real-Time Analysis (RTA) τοποθετεί στον φάκελο εξόδου αρχεία που δημιουργήθηκαν κατά την ανάλυση εικόνων, την αντιστοίχιση βάσης και τη βαθμολόγηση ποιότητας.
- ▶ Το RTA αντιγράφει τα αρχεία στον φάκελο ανάλυσης σε πραγματικό χρόνο. Αφού το RTA εκχωρήσει μια βαθμολογία ποιότητας σε κάθε βάση για κάθε κύκλο, το λογισμικό εγγράφει το αρχείο RTAComplete.txt και στους δύο φακέλους.
- ▶ Όταν υπάρχει το αρχείο RTAComplete.txt, ξεκινά η ανάλυση.
- ▶ Καθώς η ανάλυση συνεχίζεται, το Local Run Manager εγγράφει τα αρχεία εξόδου στον φάκελο ανάλυσης και, στη συνέχεια, τα αντιγράφει και πάλι στον φάκελο εξόδου.

## Φάκελοι ευθυγράμμισης


Κάθε φορά που μια ανάλυση επανατοποθετείται στην ουρά, το Local Run Manager δημιουργεί έναν φάκελο ευθυγράμμισης με όνομα **Alignment\_N**, όπου N είναι ένας αύξων αριθμός.

### Δομή φακέλων

 **Alignment**—Περιέχει αρχεία \*.bam, \*.vcf, FASTQ, καθώς και αρχεία που αφορούν τη συγκεκριμένη μονάδα ανάλυσης.

 **Date and Time Stamp**—Η ημερομηνία και η χρονική σήμανση της ανάλυσης με τη μορφή ΕΕΕΕΜΜΗΗ\_ΩΩΛΛΔΔ

 AnalysisError.txt

 AnalysisLog.txt

 aggregate.report.html

 aggregate.report.pdf

 aggregate.summary.csv

 AmpliconCoverage\_M#.tsv

 AmpliconRunStatistics.xml

 Sample1.genome.vcf.gz

 Sample1.coverage.csv

 Sample1.report.pdf

 Sample1.summary.csv

 Sample1.vcf.gz

 Sample1.bam

 **FASTQ**

 **Sample1**

 Sample1\_L001\_R1\_001\_fastq.gz

 **Stats**

 DemuxSummaryF1L1.txt

 FastqSummaryF1L1.txt

 **Data**


 **Intensities**

 **BaseCalls**


 **L001**—Περιέχει αρχεία \*.bcl.

 **L001**—Περιέχει αρχεία \*.locs.

 **RTA Logs**—Περιέχει αρχεία καταγραφής από ανάλυση του λογισμικού RTA.


 **InterOp**—Περιέχει δυαδικά αρχεία που χρησιμοποιούνται για την αναφορά μετρήσεων εκτέλεσης αλληλούχισης.



 **Logs**—Περιέχει αρχεία καταγραφής στα οποία περιγράφονται τα βήματα που εκτελέστηκαν κατά τη διάρκεια της αλληλούχησης.

 RTAComplete.txt

 RunInfo.xml

 RunParameters.xml

## Αντιστοίχιση βάσης και ποικιλομορφία ευρετηρίων

Όταν εκτελείται αλληλούχηση δειγμάτων στο όργανο NextSeq 550Dx, η αντιστοιχία βάσης ορίζει μια βάση (A, C, G, ή T) για κάθε συστάδα ενός συγκεκριμένου πλακιδίου ή μια περιοχή απεικόνισης στην κυψελίδα ροής, σε έναν συγκεκριμένο κύκλο. Το όργανο NextSeq 550Dx χρησιμοποιεί αλληλούχηση δύο καναλιών, η οποία απαιτεί μόνο δύο εικόνες για την κωδικοποίηση των δεδομένων για τις τέσσερις βάσεις του DNA, μία από το κόκκινο κανάλι και μία από το πράσινο κανάλι.

Η διαδικασία των αναγνώσεων ευρετηρίου αντιστοίχισης βάσης διαφέρει από την αντιστοίχιση βάσης κατά τη διάρκεια άλλων αναγνώσεων.

Οι αναγνώσεις ευρετηρίου πρέπει να ξεκινούν με τουλάχιστον μία βάση άλλη από τη G σε κάποιον από τους δύο πρώτους κύκλους. Εάν μια ανάγνωση ευρετηρίου ξεκινά με δύο αντιστοιχίσεις βάσης G, δεν δημιουργείται ένταση σήματος. Το σήμα πρέπει να υπάρχει σε έναν από τους δύο πρώτους κύκλους για τη διασφάλιση της απόδοσης αποπολύπλεξης.

Στην επιλογή ευρετηρίων κατά τη διάρκεια της δημιουργίας μιας εκτέλεσης, θα εμφανιστεί μια προειδοποίηση χαμηλής ποικιλομορφίας, αν τα ευρετήρια δεν πληρούν τις απαιτήσεις ποικιλομορφίας. Για την αποφυγή της εμφάνισης της προειδοποίησης χαμηλής ποικιλομορφίας, πρέπει να επιλέγετε αλληλουχίες ευρετηρίου με σήμα και στα δύο κανάλια για κάθε κύκλο.

- ▶ Κόκκινο κανάλι—A ή C
- ▶ Πράσινο κανάλι—A ή T

Αυτή η διαδικασία αντιστοίχισης βάσης διασφαλίζει ακρίβεια κατά την ανάλυση μικρότερου αριθμού δειγμάτων. Για περισσότερες πληροφορίες σχετικά με τις αλληλουχίες των ευρετηρίων σας, δείτε το ένθετο συσκευασίας *TruSeq Custom Amplicon Kit Dx* (αρ. εγγράφου 1000000029772).

Κατά τη δημιουργία εκτελέσεων στο Local Run Manager θα επιλέξετε τον αριθμό των δειγμάτων προς εξέταση. Το λογισμικό συμπληρώνει αυτόματα τους προτεινόμενους συνδυασμούς ευρετηρίων που πληρούν τις απαιτήσεις ποικιλομορφίας ευρετηρίων. Εάν και δεν είναι υποχρεωτικό να χρησιμοποιήσετε τους προτεινόμενους συνδυασμούς ευρετηρίων, συνιστάται να το πράξετε.

## Ιστορικό αναθεώρησης

Έγγραφο	Ημερομηνία	Περιγραφή αλλαγής
Αρ. εγγράφου 1000000030330 v04	Αύγουστος 2021	Ενημέρωση της διεύθυνσης εξουσιοδοτημένου εκπροσώπου για την ΕΕ.
Αρ. εγγράφου 1000000030330 v03	Δεκέμβριος 2019	Ενημέρωση της διεύθυνσης εξουσιοδοτημένου εκπροσώπου για την ΕΕ. Ενημέρωση της διεύθυνσης του χορηγού στην Αυστραλία.
Αρ. εγγράφου 1000000030330 v02	Ιανουάριος 2019	Προσθήκη πληροφοριών για τα κιτ αντιδραστηρίων v2.5.
Αρ. εγγράφου 1000000030330 v01	Αύγουστος 2018	Ενημέρωση κανονιστικών επισημάνσεων.
Αρ. εγγράφου 1000000030330 v00	Νοέμβριος 2017	Αρχική δημοσίευση.

## Τεχνική βοήθεια

Για τεχνική βοήθεια, επικοινωνήστε με το τμήμα τεχνικής υποστήριξης της Illumina.

Ιστότοπος: [www.illumina.com](http://www.illumina.com)  
 Email: [techsupport@illumina.com](mailto:techsupport@illumina.com)

### Τηλεφωνικοί αριθμοί εξυπηρέτησης πελατών της Illumina

Περιοχή	Χωρίς χρέωση	Τοπικό
Βόρεια Αμερική	+1.800.809.4566	
Αυστραλία	+1.800.775.688	
Αυστρία	+43 800006249	+43 19286540
Βέλγιο	+32 80077160	+32 34002973
Γαλλία	+33 805102193	+33 170770446
Γερμανία	+49 8001014940	+49 8938035677
Δανία	+45 80820183	+45 89871156
Ελβετία	+41 565800000	+41 800200442
Ηνωμένο Βασίλειο	+44 8000126019	+44 2073057197
Ιαπωνία	0800.111.5011	
Ιρλανδία	+353 1800936608	+353 016950506
Ισπανία	+34 911899417	+34 800300143
Ιταλία	+39 800985513	+39 236003759
Κάτω Χώρες	+31 8000222493	+31 207132960
Κίνα	400.066.5835	
Νέα Ζηλανδία	0800.451.650	
Νορβηγία	+47 800 16836	+47 21939693
Σιγκαπούρη	+1.800.579.2745	
Σουηδία	+46 850619671	+46 200883979
Ταϊβάν	00806651752	
Φινλανδία	+358 800918363	+358 974790110
Χονγκ Κονγκ	800960230	
Άλλες χώρες	+44.1799.534000	

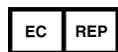
**Φύλλα δεδομένων ασφάλειας (SDS)**—Διαθέσιμα στον ιστότοπο της Illumina στη διεύθυνση [support.illumina.com/sds.html](http://support.illumina.com/sds.html).

**Τεκμηρίωση προϊόντος**—Διαθέσιμη για λήψη σε μορφή PDF από τον ιστότοπο της Illumina. Μεταβείτε στη διεύθυνση [support.illumina.com](http://support.illumina.com), επιλέξτε ένα προϊόν και κατόπιν επιλέξτε την καρτέλα **Documentation & Literature** (Τεκμηρίωση και βιβλιογραφία).



Illumina  
5200 Illumina Way  
San Diego, California 92122  
Η.Π.Α.  
+1.800.809.ILMN (4566)  
+1.858.202.4566  
(εκτός Βορείου Αμερικής)  
techsupport@illumina.com  
www.illumina.com

CE



Illumina Netherlands B. V.  
Steenoven 19  
5626 DK Eindhoven  
Κάτω Χώρες

**Χορηγός στην Αυστραλία**  
Illumina Australia Pty Ltd  
Nursing Association Building  
Level 3, 535 Elizabeth Street  
Melbourne, VIC 3000  
Αυστραλία

ΓΙΑ ΧΡΗΣΗ ΣΤΗ ΔΙΑΓΝΩΣΗ IN VITRO

© 2021 Illumina, Inc. Με την επιφύλαξη παντός δικαιώματος.

**illumina**<sup>®</sup>