

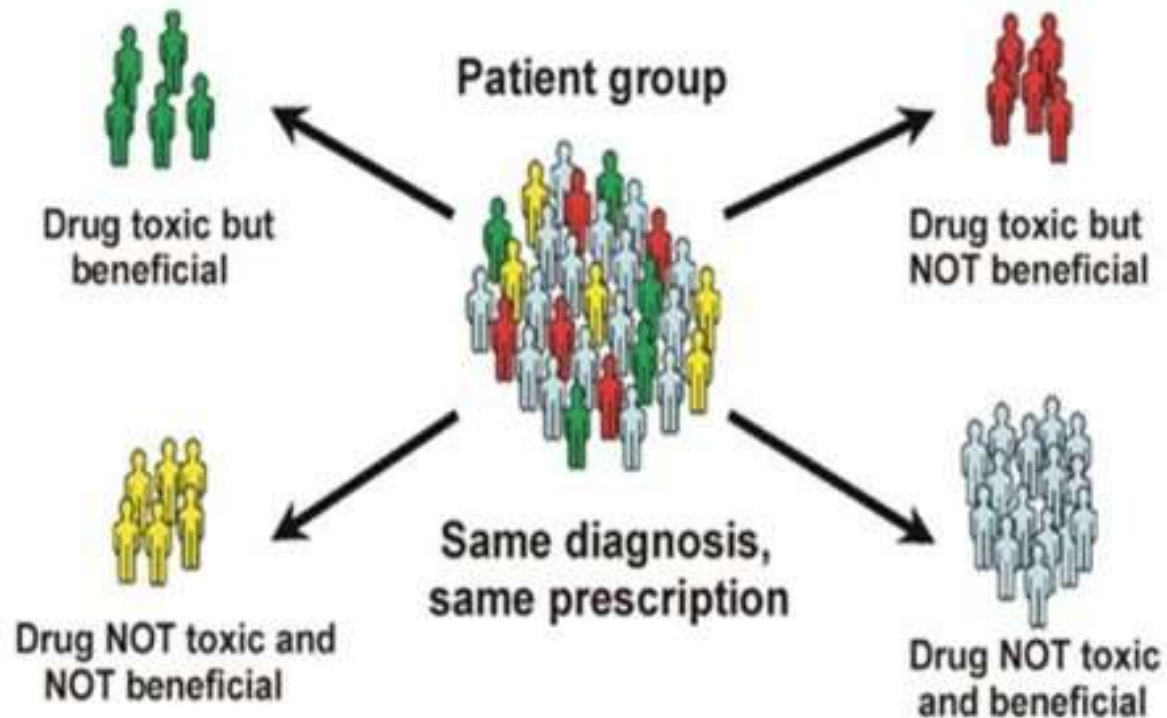


Pharmacogenomic technologies

Marianna Lucafò
mlucafo@units.it

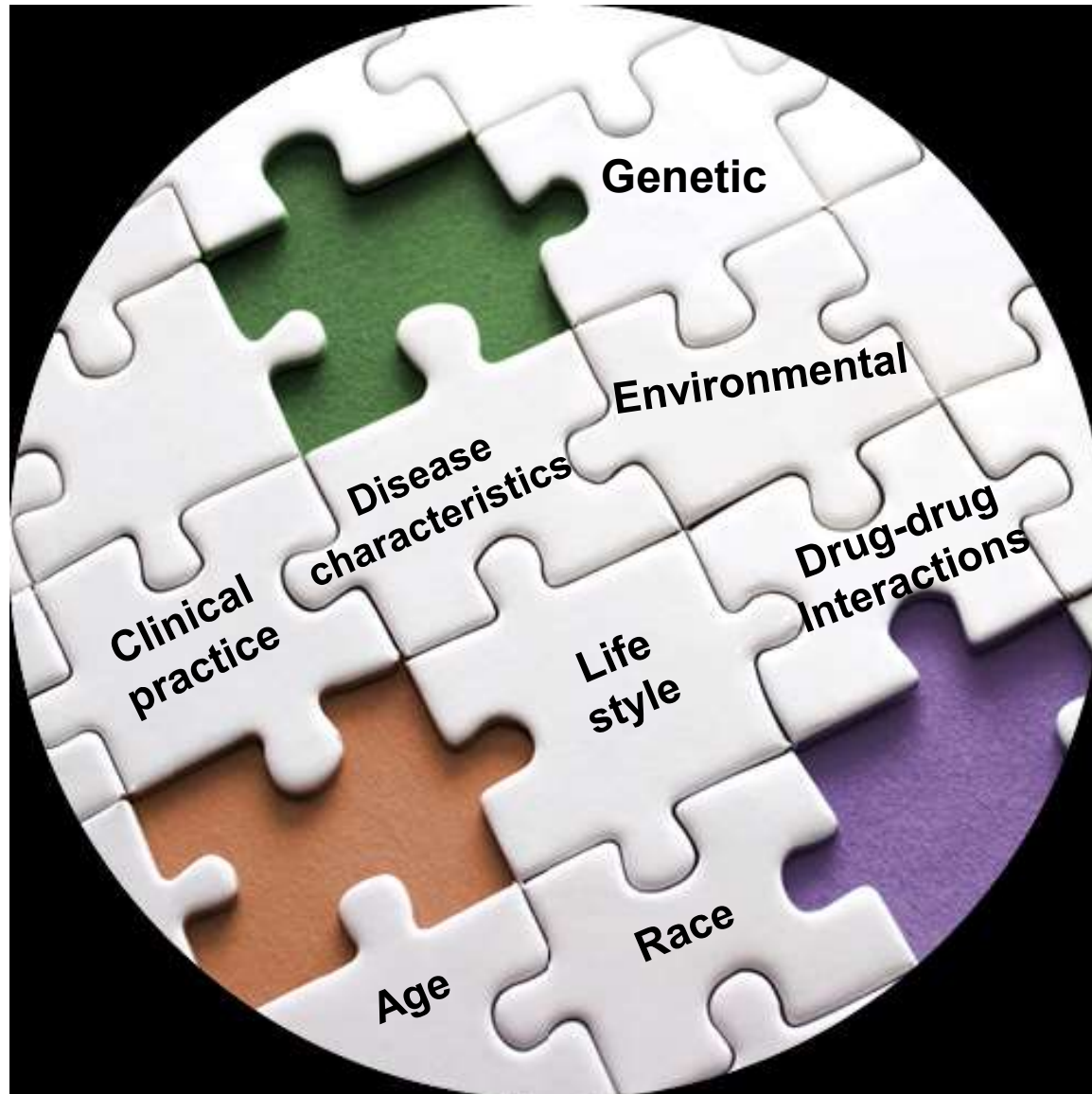
Pharmacogenomics

contribution of genetic factors to the interindividual variability in drug efficacy and safety

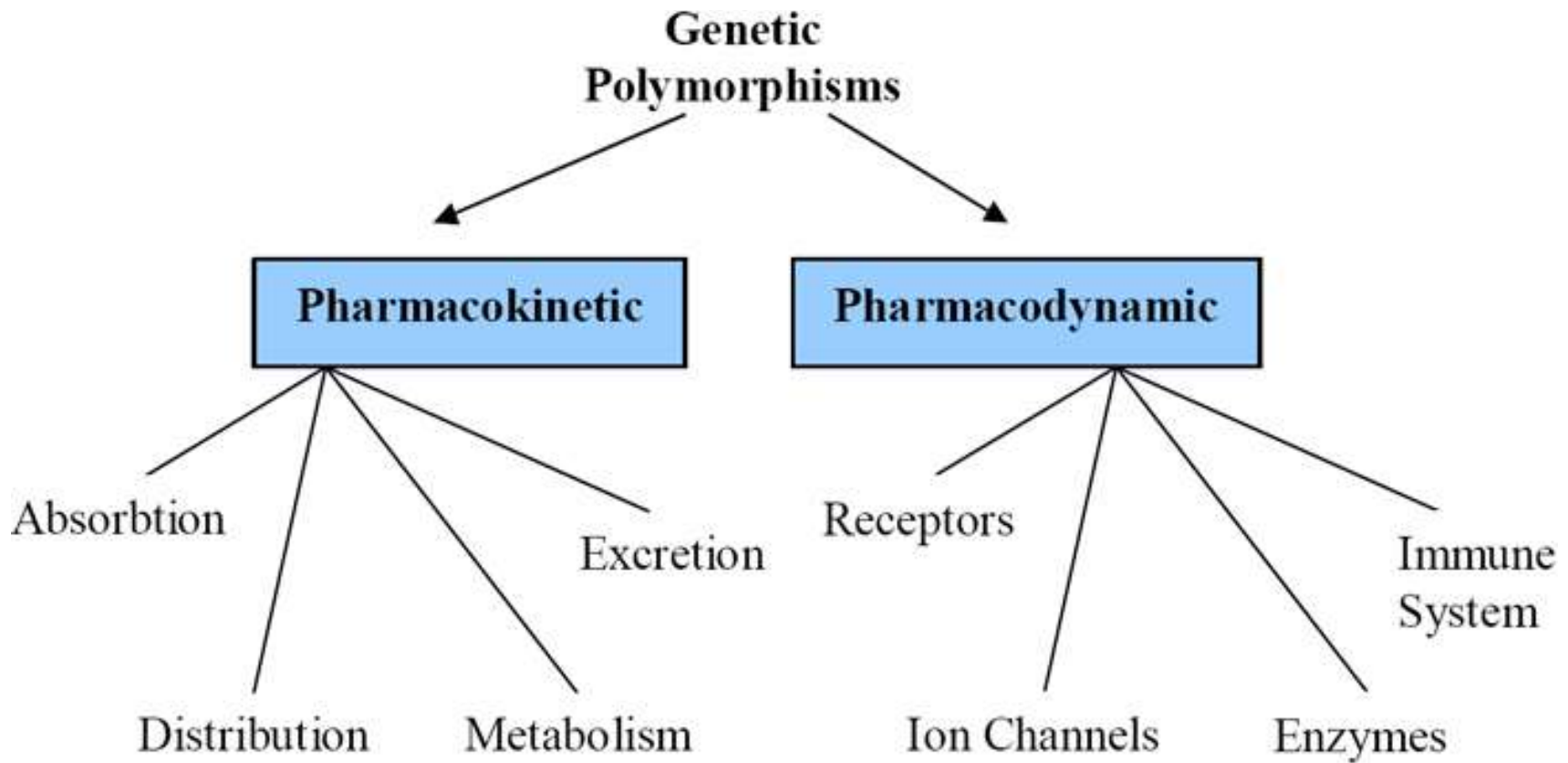


Pharmacogenomics

contribution of genetic factors to the interindividual variability in drug efficacy and safety



Pharmacogenetics



Pharmacogenomics and Personalized Medicine

By: Jill U. Adams, Ph.D. (Freelance science writer in Albany, NY) © 2008 Nature Education

Types of DNA sequence variation

- SNP: Single Nucleotide Polymorphism
- RFLP: Restriction Fragment Length Polymorphism
- VNRT: Variable Number of Tandem Repeats
- CNV: Copy Number Variations
- SSR: Simple Sequence Repeat
- Insertions or Deletions
- Rearrangement

PharmGKB: The Pharmacogenomics Knowledge Base

← → ↻ 🏠 🔒 https://www.pharmgkb.org 50% PHARMGKB →

PHARMGKB Publications News Downloads Contact Help

Search PharmGKB 🔍

Search for a molecule, gene, variant, or combination

Therapeutic Resource for COVID-19

PharmGKB data are under a Creative Commons license. More details are in our [Data Usage Policy](#). Please [cite PharmGKB](#) if you use our information or images.

Annotated Drugs 708

Curated Pathways 150

Clinical Guideline Annotations 161

Drug Label Annotations 780

WHAT IS PHARMACOGENOMICS?
The study of the relationship between genetic variations and how our body responds to medications.
[Pretty cool right? Tell me more.](#)

PHARMACOGENOMICS. KNOWLEDGE. IMPLEMENTATION.
PharmGKB is a comprehensive resource that curates knowledge about the impact of genetic variation on drug response for clinicians and researchers.
[Learn more about PharmGKB](#)

Tecniche di genotipizzazione

- **A basso rendimento (low throughput)**
 - **PCR-RFLP**
 - **PCR allele specifica**
 - **Ibridazione oligo (ASO)**

Polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP)

Amplificazione DNA (PCR)



Digestione dell'amplificato con enzima di restrizione (<http://rebase.neb.com/rebase/rebase.html>)

creazione /abolizione sito di taglio
frammentazione anomala

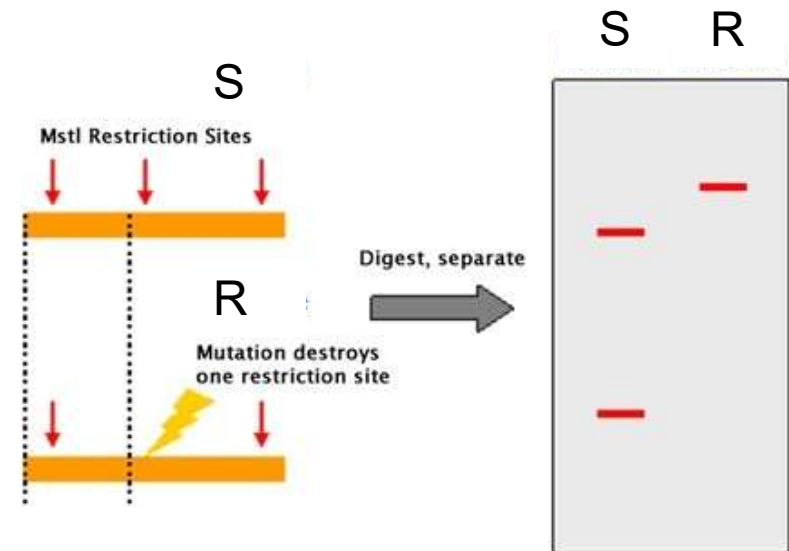


Elettroforesi su gel di agarosio
colorazione con GEL RED o
etidio bromuro



Pattern di frammenti caratteristici
creazione sito di taglio: + 1 banda
abolizione sito di taglio: - 1 banda

Gli enzimi di restrizione sono endonucleasi di origine batterica in grado di riconoscere specifiche sequenze di 4-8 nucleotidi e di tagliare in quelle posizioni il DNA.

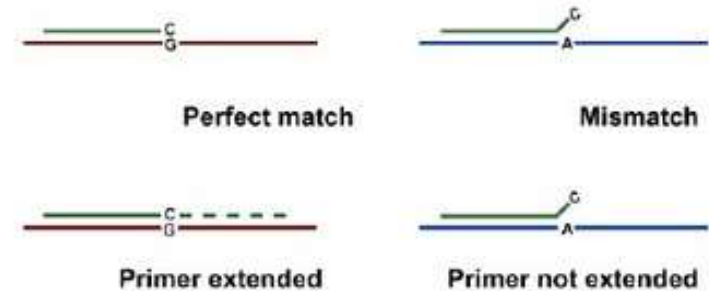


Allele-specific PCR

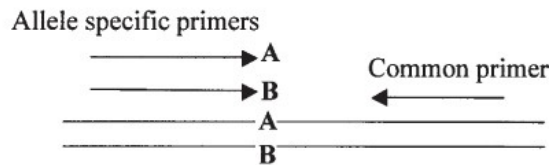
Reazione di PCR nella quale vengono utilizzati separatamente primers allele-specifici che differiscono per il nucleotide all'estremità 3' del filamento.

La sintesi dell'amplificato nella reazione di PCR dipende dal corretto appaiamento dell'estremità 3'.

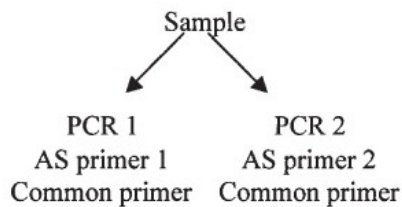
Per ogni campione vengono allestite due reazioni di PCR utilizzando in ogni reazione il primer comune in combinazione con quello corrispondente alla sequenza wt e a quella mutata.



A Assay design



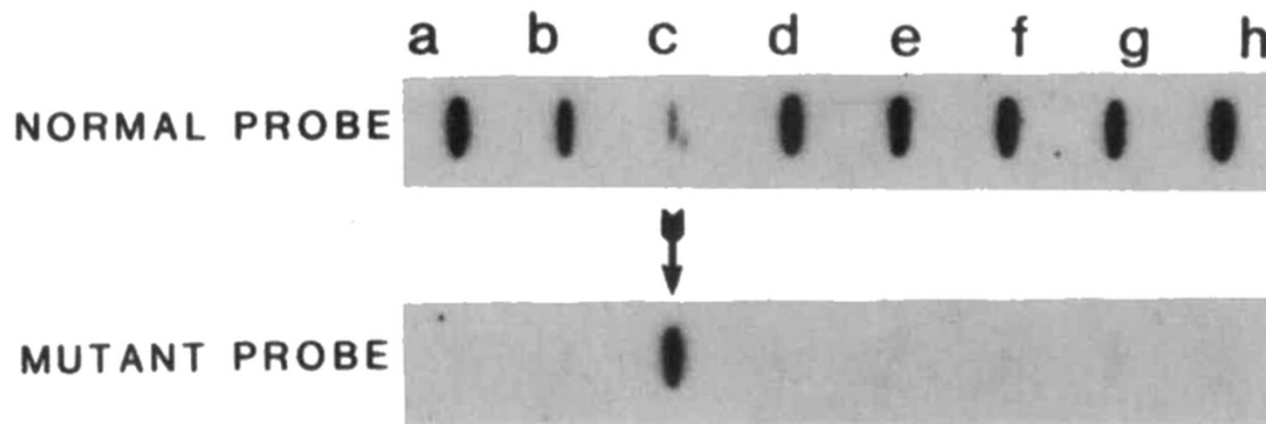
B Amplification



La presenza degli amplificati viene evidenziata mediante elettroforesi su gel di agarosio e colorazione con gel red.

Allele-specific oligonucleotides PCR (ASO)

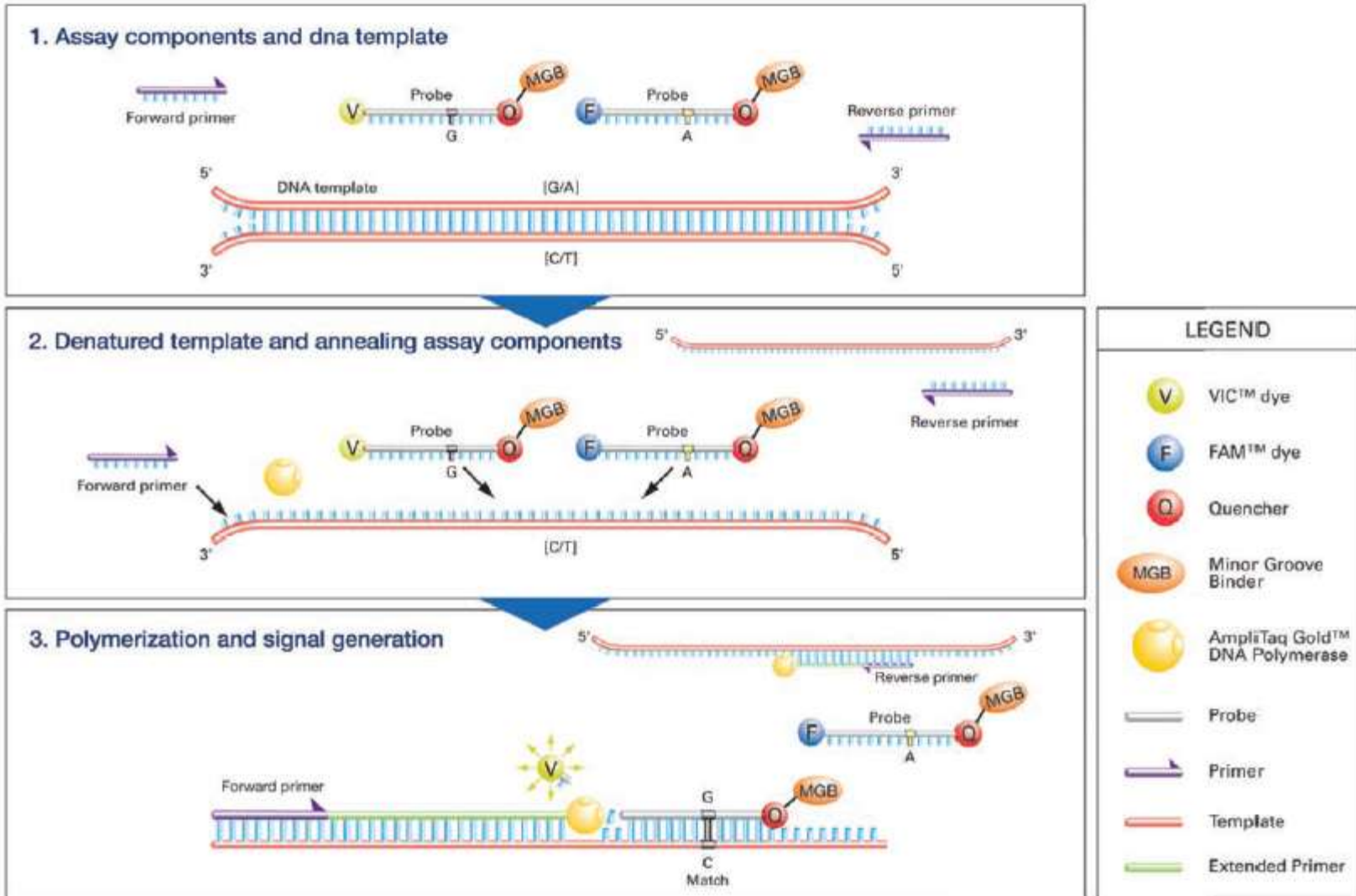
- La presenza di una mutazione viene riconosciuta facendo reagire il DNA amplificato con sonde oligonucleotidiche complementari alla sequenza mutata o normale (separatamente).
- Solo in presenza del 100% di omologia tra le sequenze si avrà l'ibridazione della sonda con il DNA.



Tecniche di genotipizzazione

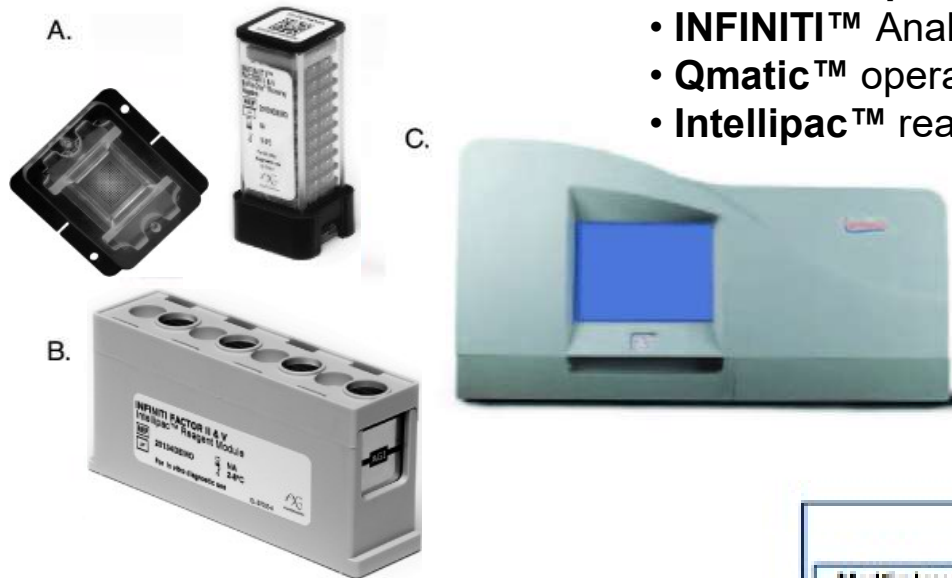
- **A medio rendimento (medium throughput)**
 - **Taqman® (Applied Biosystem)**
 - **Infiniti™ (Autogenomics)**
 - **Pyrosequencing® (QIAGEN)**
 - **Invader® Assay (Hologic)**

TaqMan® SNP Genotyping Assays



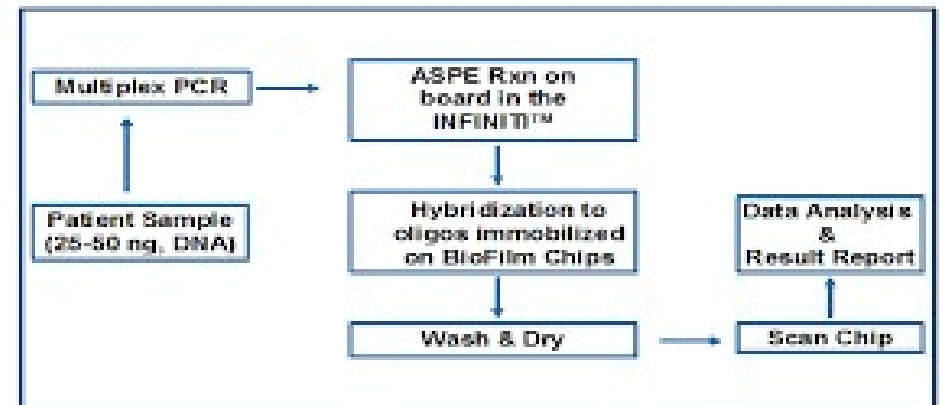
Infiniti™ (Autogenomics)

- BioFilmChip™ microarray (A)
- INFINITI™ Analyzer with sample to result automation (C)
- Qmatic™ operating software with applications interface
- Intellipac™ reagent management module (B)

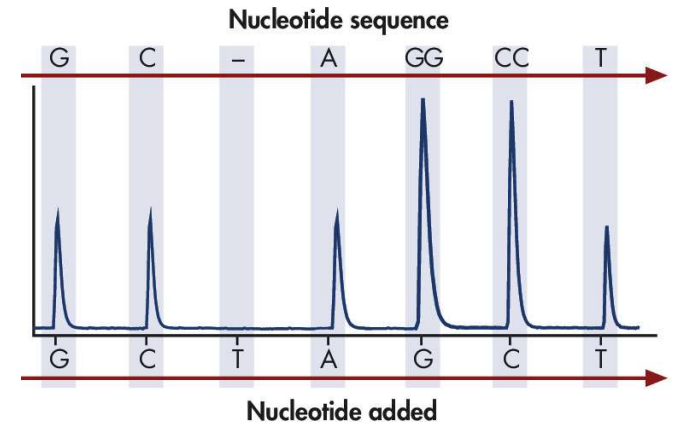
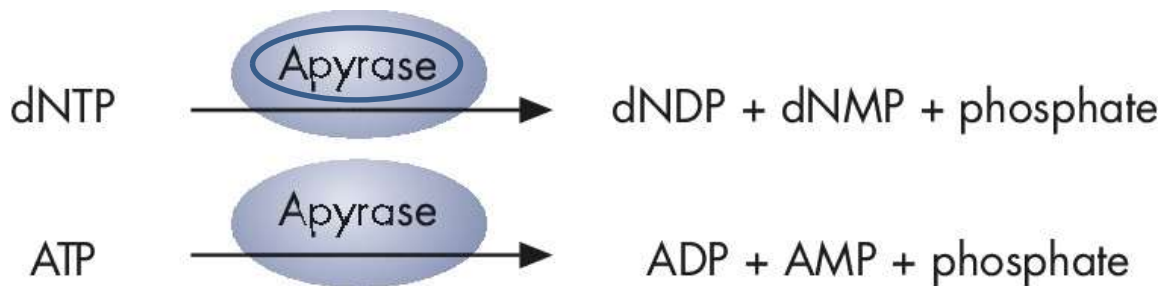
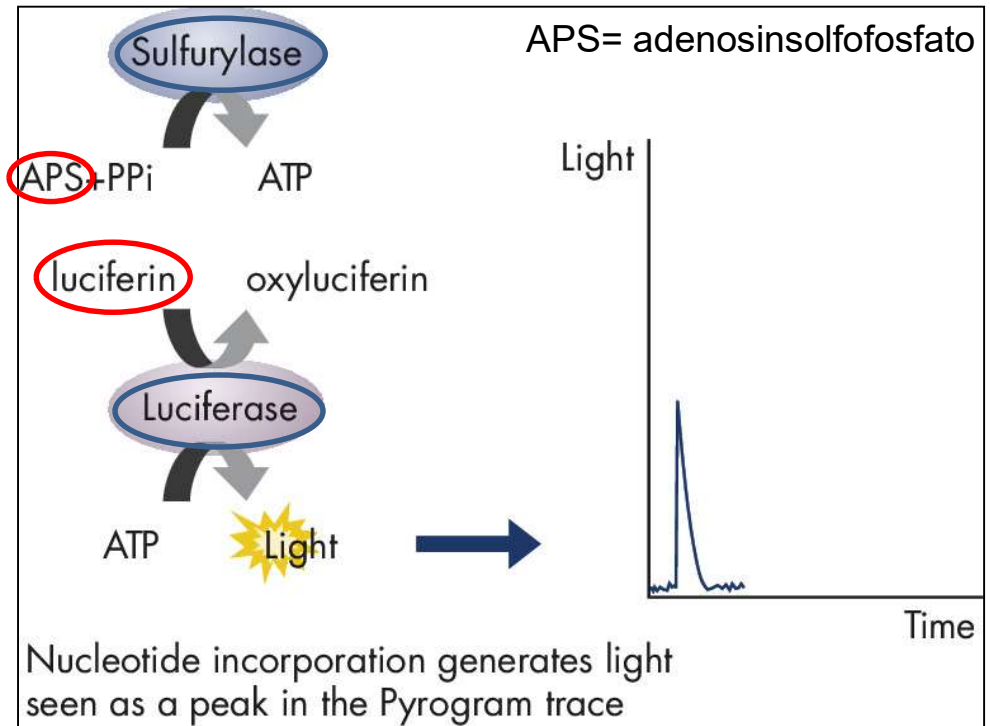
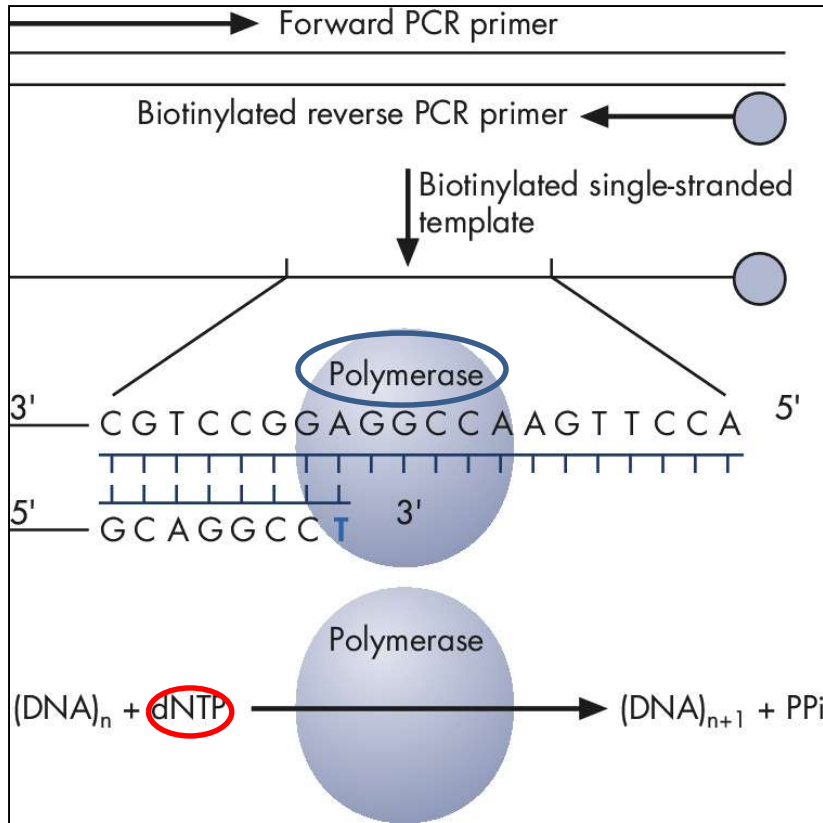


Protocollo del test:

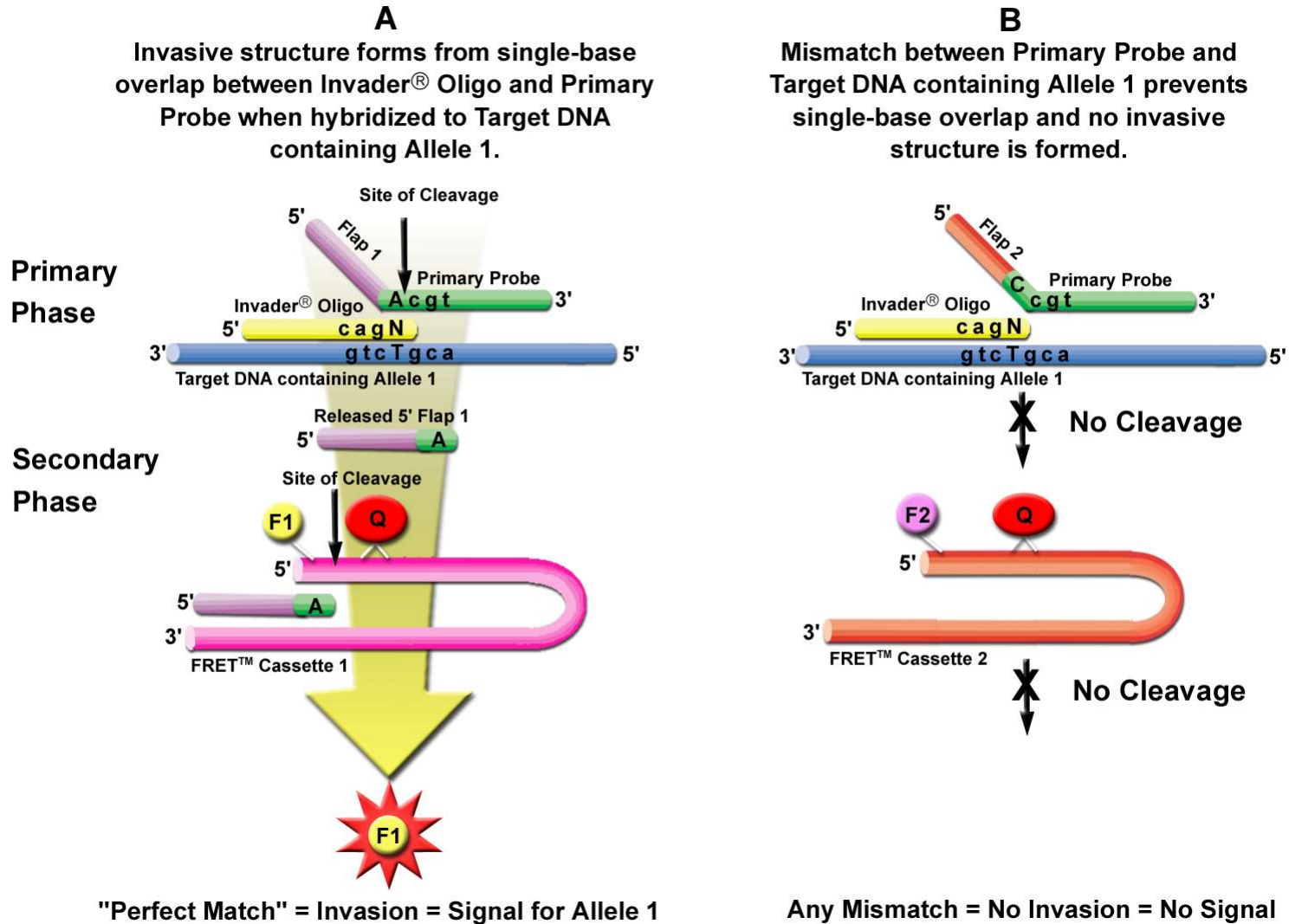
- Estrazione del DNA da campioni di sangue umano
- Amplificazione mediante PCR
- Estensione del primer allele specifico con incorporazione del marcato (ASPE)
- Ibridazione dei primer esteso e fluorescente su un microarray e successivo lavaggio
- Scansione del microarray
- Rilevazione del segnale e analisi (determinazione del genotipo)



Pyrosequencing® (QIAGEN)



Invader[®] Assay (Hologic)



Tecniche di genotipizzazione

- Ad altissimo rendimento (super-high throughput)
 - **Genome-wide Human SNP Array 6.0 (Affymetrix)**

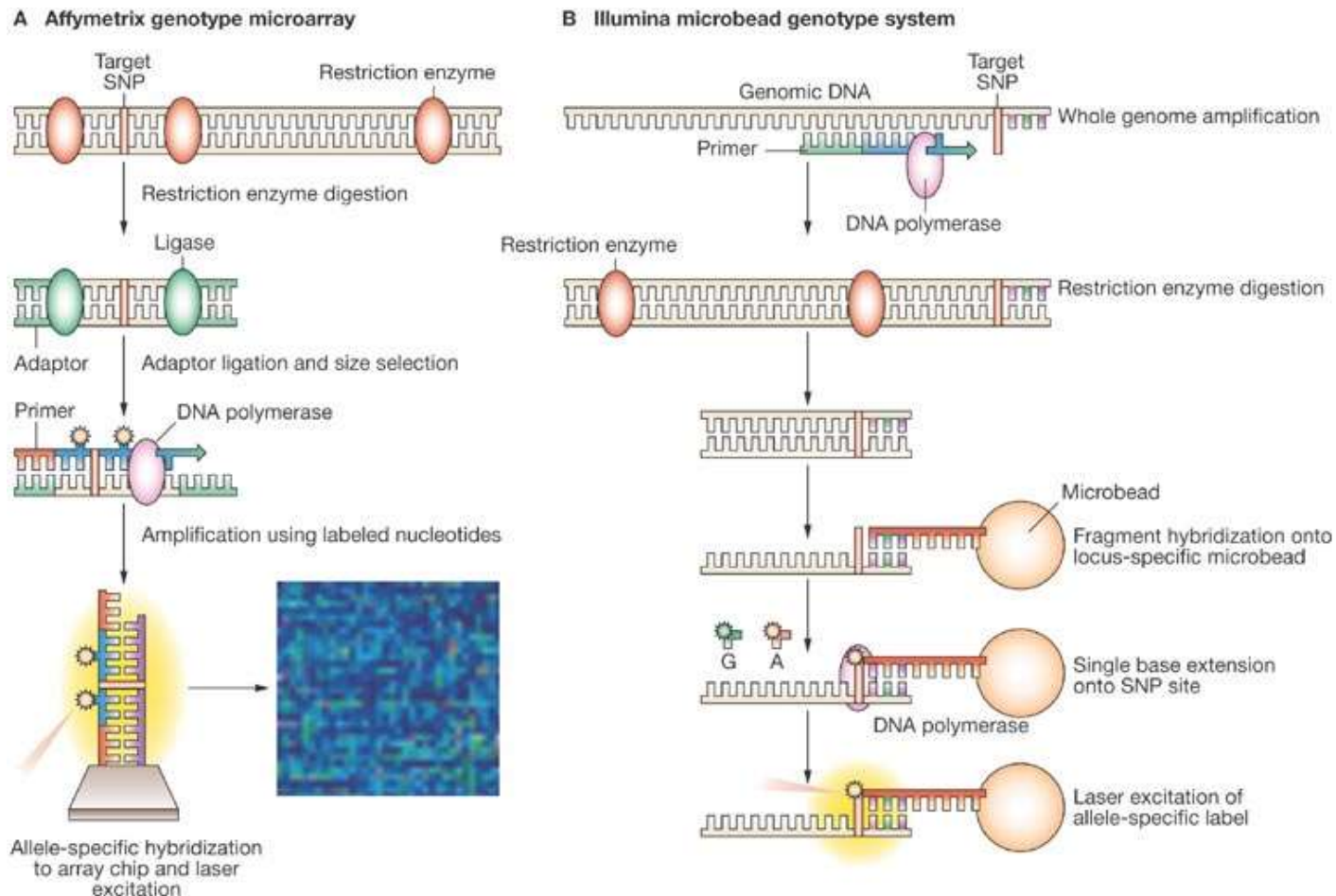
1.8 million markers, including 946,000 probes for the detection of copy number variants and 906,600 SNPs

- **BeadArray™ System (Illumina)**

1.2 million markers



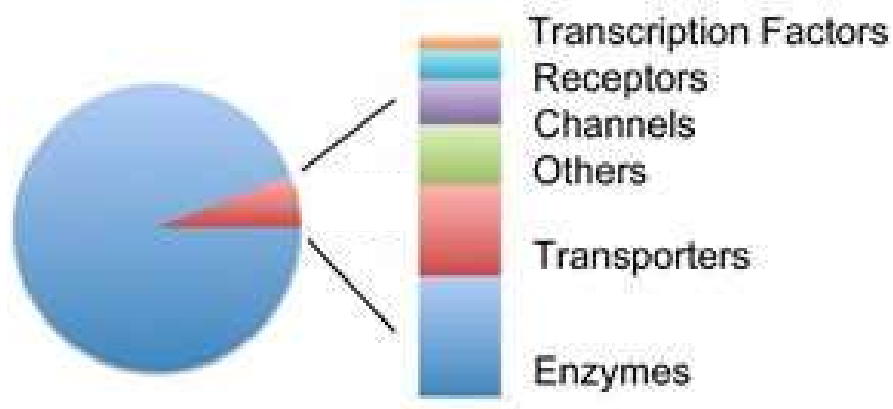
Two ultra-high-throughput single nucleotide polymorphism genotyping platforms for use in genome-wide association analyses.



Walker EJ and Siminovitch KA (2007) Primer: genomic and proteomic tools for the molecular dissection of disease *Nat Clin Pract Rheumatol* 3: 580-589 doi:10.1038/ncprheum0595

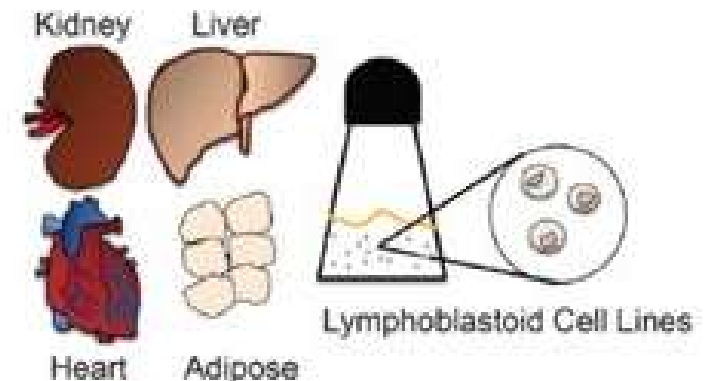
Transcriptomic variation of pharmacogenes

Pharmacogene candidates

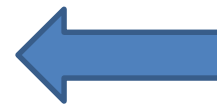


Realtime PCR
SYBR green
TaqMan Technology

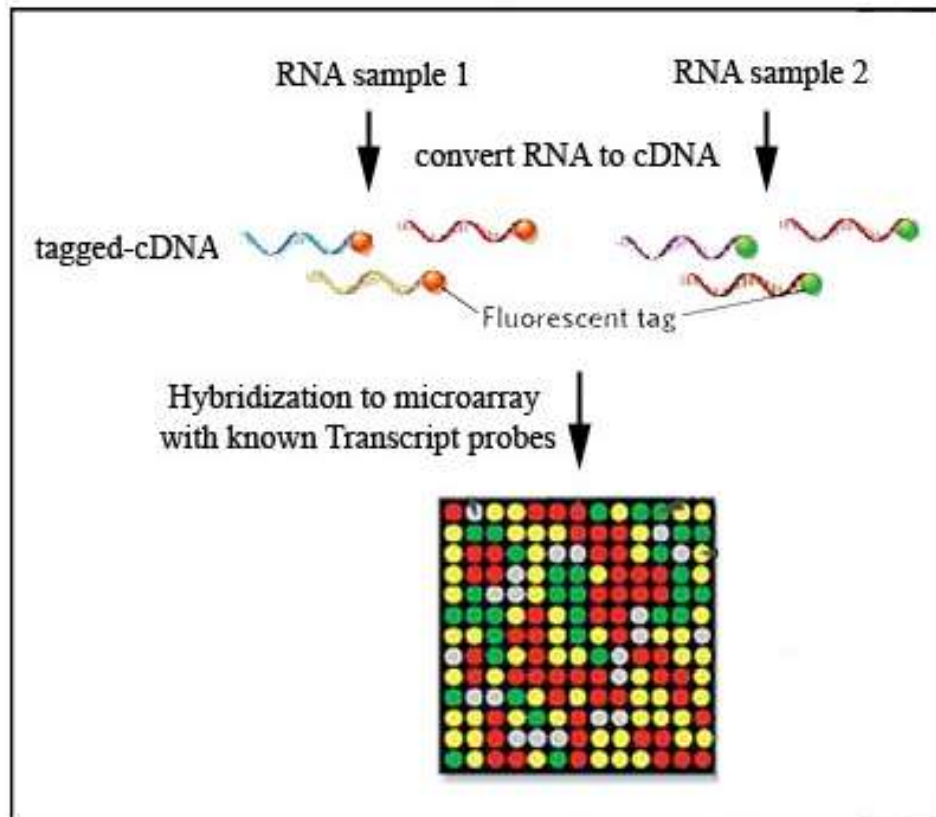
High-throughput screening



MICROARRAY
Next Generation Sequencing
Roche 454
Illumina Hi-Seq
Ion Torrent
Nanopore Technologies



Microarray



relative intensity
=
expression levels

Low sensitivity
Low dynamic range
known transcript only
No alternative splicing information
lower cost

Perché RNA-sequencing?

MICROARRAY:

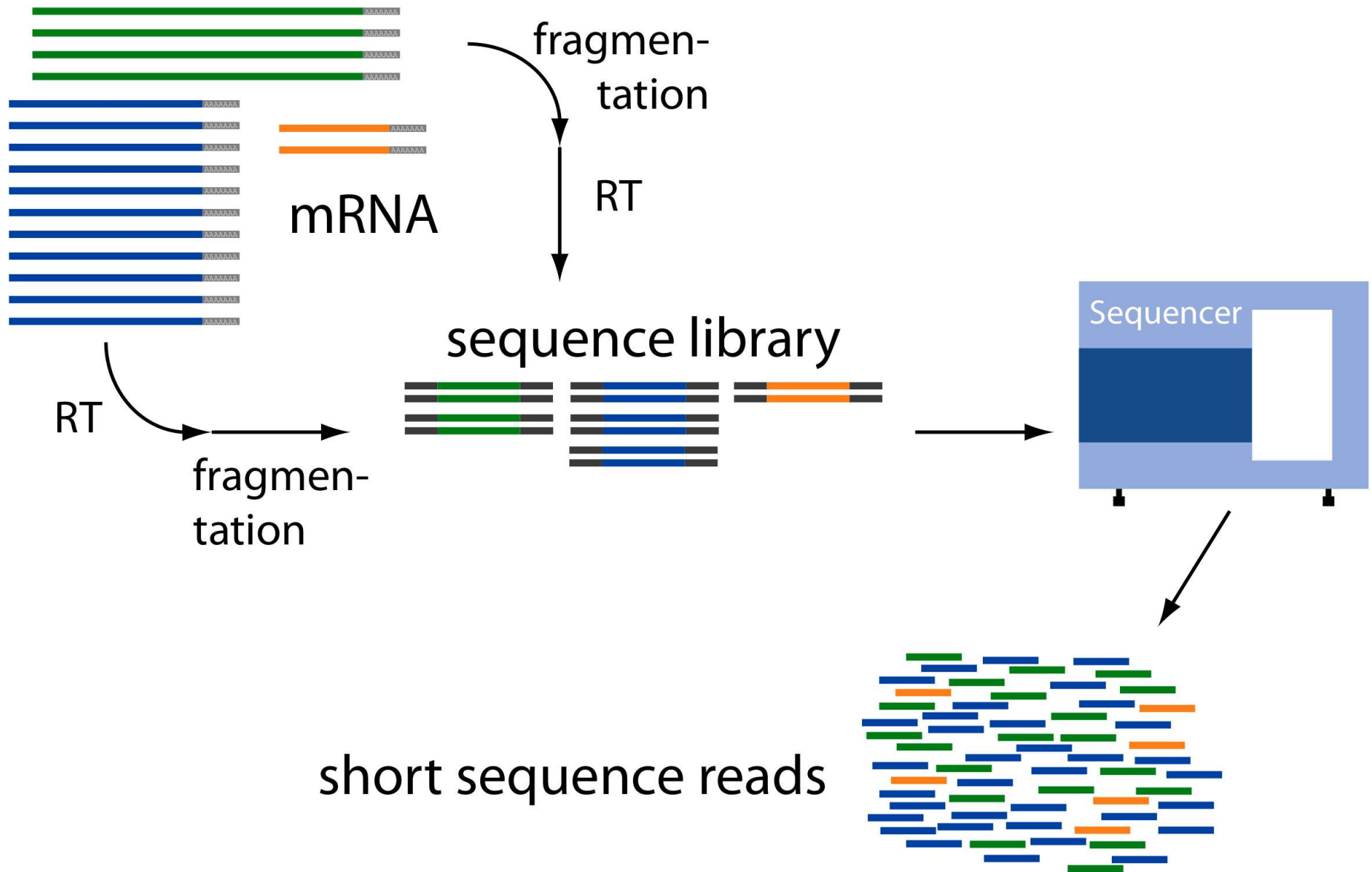
1. Limitato alle seq. spottate sul chip (non ho il trascrittoma completo);
2. Si basa su segnali di ibridazione competitiva strettamente dipendenti da come sono preparate le librerie (variabilità di fondo = scarsa affidabilità statistica);
3. Si basa sull'ibridazione di sonde corte che legano solo una specifica porzione del trascritto quindi non ho un'informazione riguardo alle possibili varianti (es. splicing alternativi).

RNA-seq:

1. Non ho limitazioni, posso analizzare tutte le seq. espresse;
2. Posso discriminare un campione dall'altro, ottengo dei valori di espressione assoluti e direttamente comparabili (direttamente proporzionali al numero di conte = statisticamente affidabile);
3. La mappatura delle reads avviene sul genoma quindi è possibile rilevare le possibili varianti di un singolo trascritto;
4. Maggiori applicazioni-potenzialità-informazioni

I COSTI SONO COMPARABILI

RNA-sequencing: la tecnica



RNA-seq: preparazione campioni

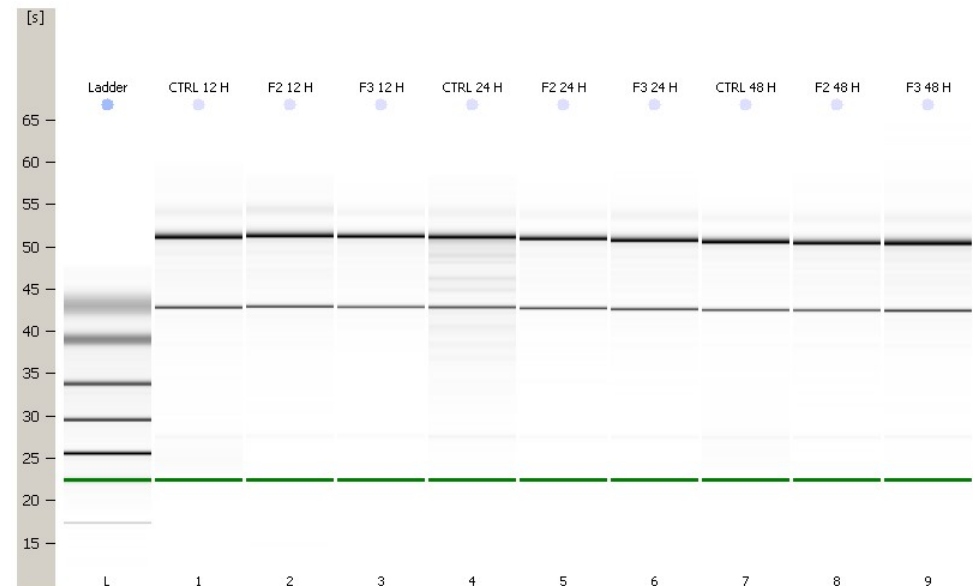
**VALUTAZIONE DELLA QUALITA' DELL'RNA MEDIANTE ELETTROFORESI CAPILLARE SU CHIP
(BIOANALYZER dell'AGILENT)**

Il campione di RNA:

- non deve essere degradato
- deve avere un rapporto A260/A280 compreso tra 1.8 e 2
- deve avere un rapporto A260/A230 compreso tra 2 e 2,2
- 1-2 µg di RNA

RNA INTEGRITY NUMBER ≥ 7!!!

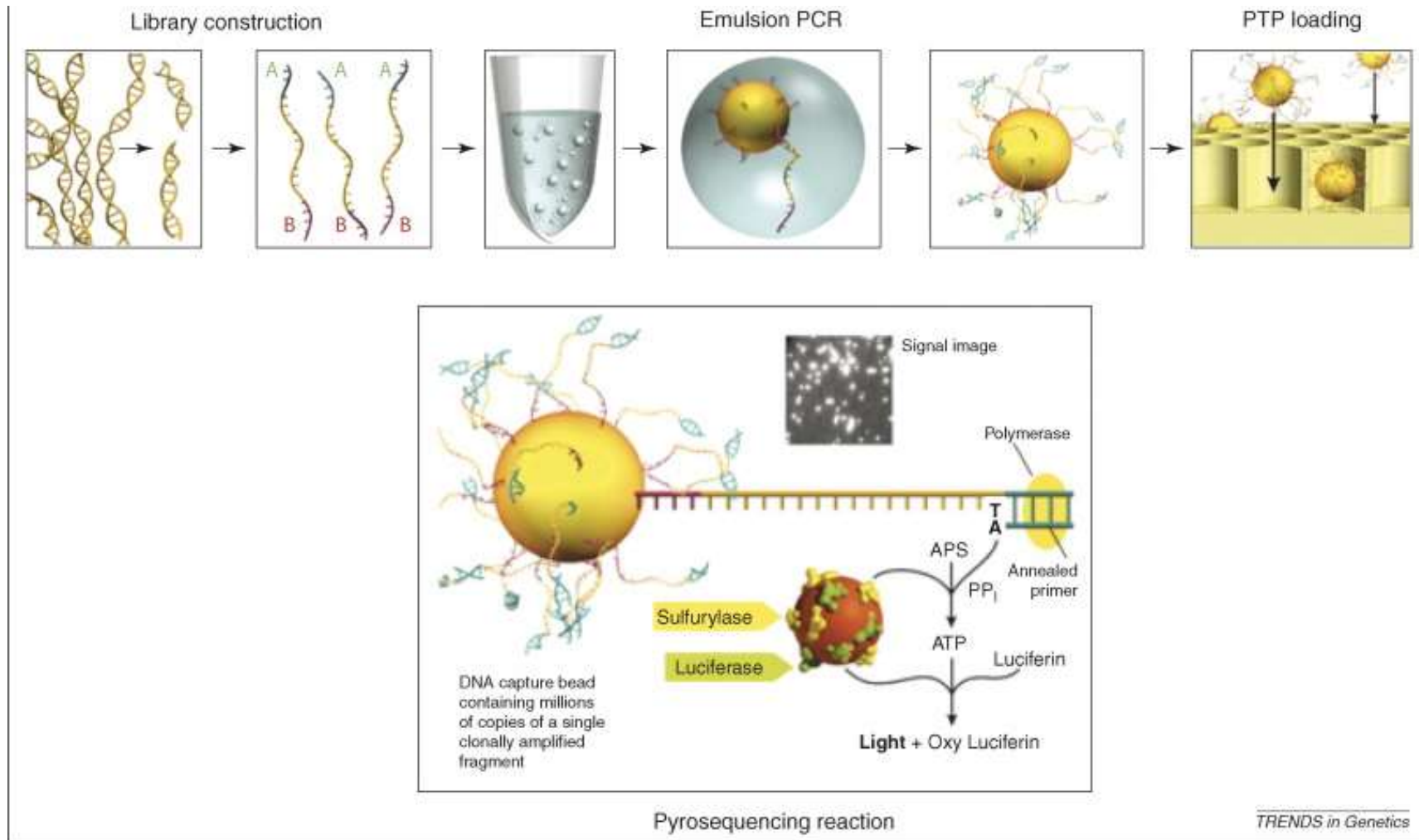
**DATO DAL RAPPORTO DELL'INTENSITA' DELLE
2 BANDE DI rRNA RISPETTO AL RUMORE DI
FONDO**



NGS platform: Roche 454

Sequencing method: Pyrosequencing

Library amplification method: Emulsion PCR



NGS platform: Illumina platform

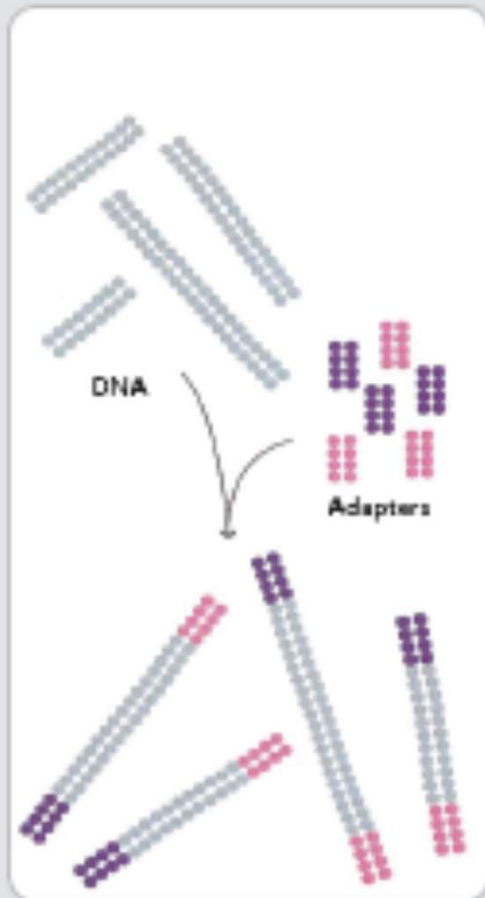
Illumina (Solexa) HiSeq 2000 Flow Cell



- 1 flow cell
- 8 canali
- Ogni canale può correre fino a 12 differenti librerie (Multiplexed Sequencing).
- Input richiesto: 0.1–1.0 μg (single- and paired-end reads), 10 μg (Mate Pair reads).
- 1.4-mm larghezza del canale.
- Si basa sul legame casuale di frammenti di oligo DNA fissati sul vetrino, una superficie ottica trasparente (flow cell).
- 96-120 milioni di reads (clusters) per lane

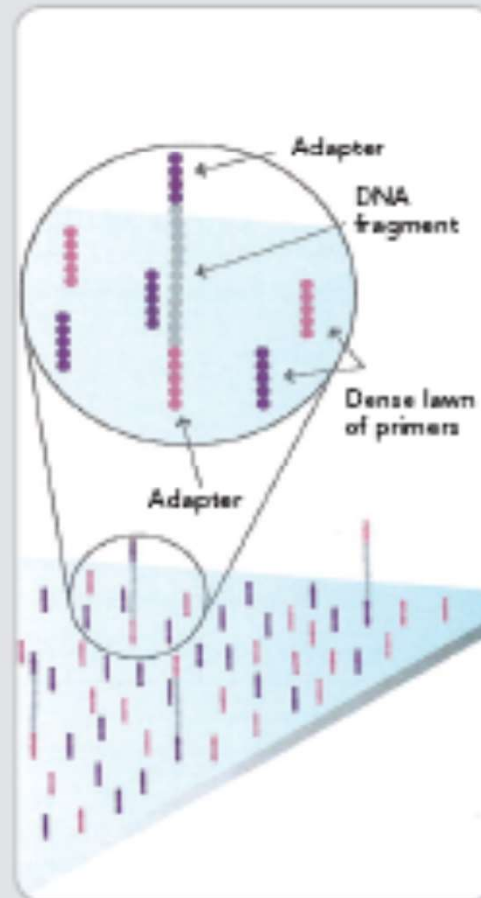
NGS platform: Illumina platform

1. PREPARE GENOMIC DNA SAMPLE



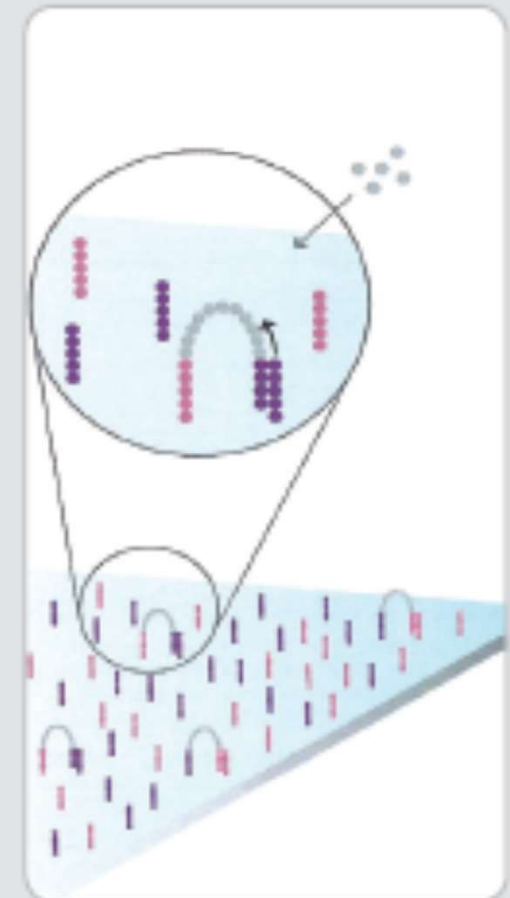
Randomly fragment genomic DNA and ligate adapters to both ends of the fragments.

2. ATTACH DNA TO SURFACE



Bind single-stranded fragments randomly to the inside surface of the flow cell channels.

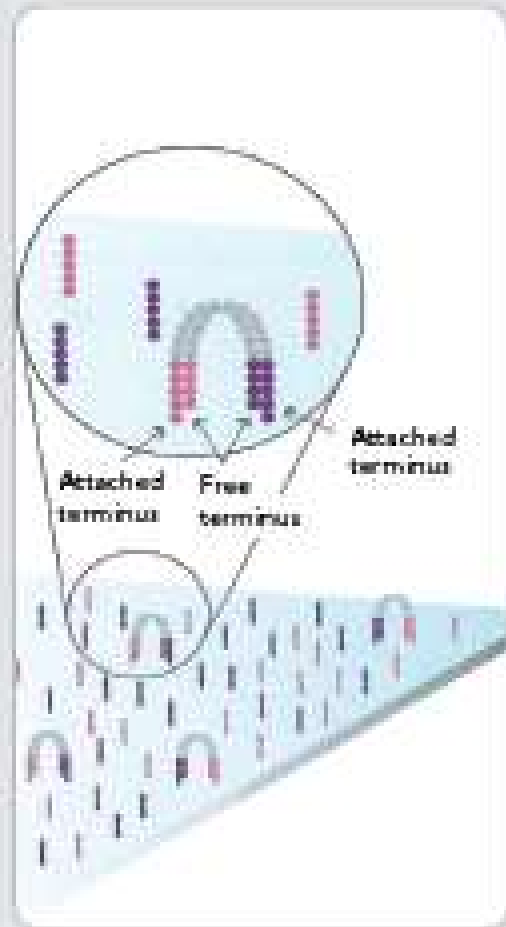
3. BRIDGE AMPLIFICATION



Add unlabeled nucleotides and enzyme to initiate solid-phase bridge amplification.

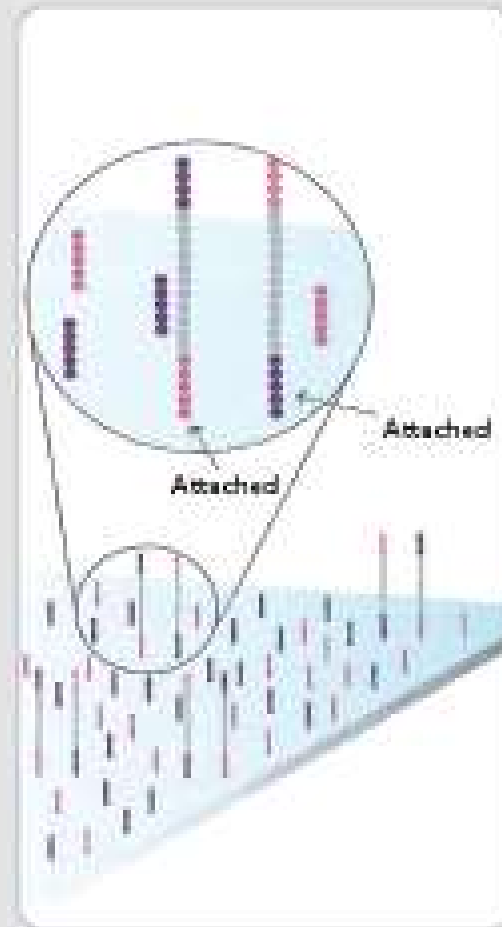
NGS platform: Illumina platform

4. FRAGMENTS BECOME DOUBLE-STRANDED



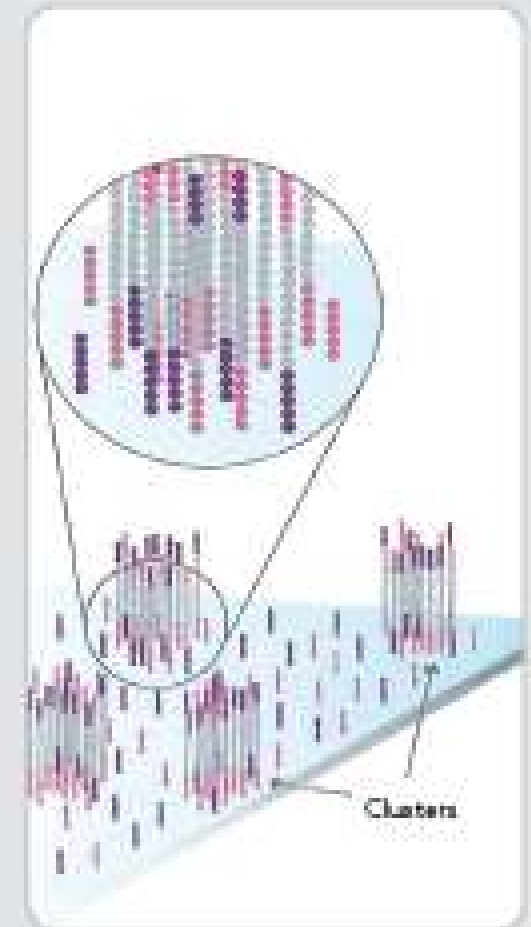
The enzyme incorporates nucleotides to build double-stranded bridges on the solid-phase substrate.

5. DENATURE THE DOUBLE-STRANDED MOLECULES



Denaturation leaves single-stranded templates anchored to the substrate.

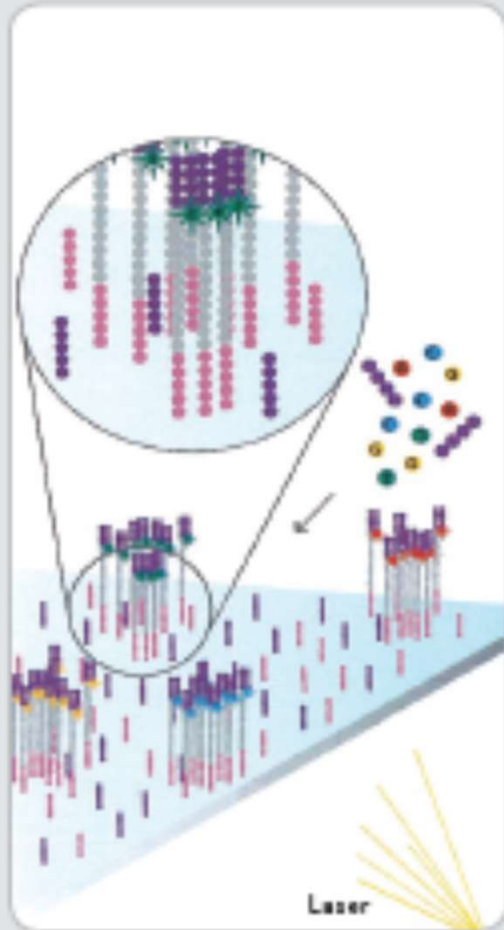
6. COMPLETE AMPLIFICATION



Several million dense clusters of double-stranded DNA are generated in each channel of the flow cell.

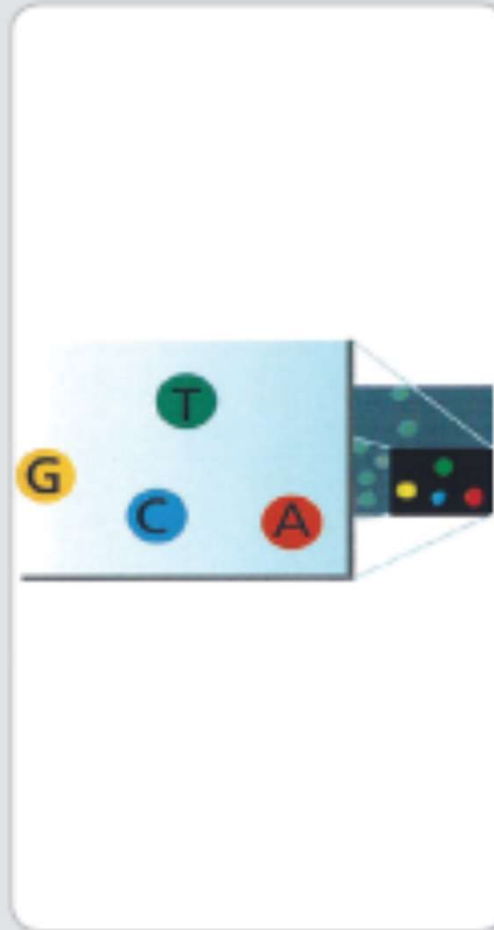
NGS platform: Illumina platform

7. DETERMINE FIRST BASE



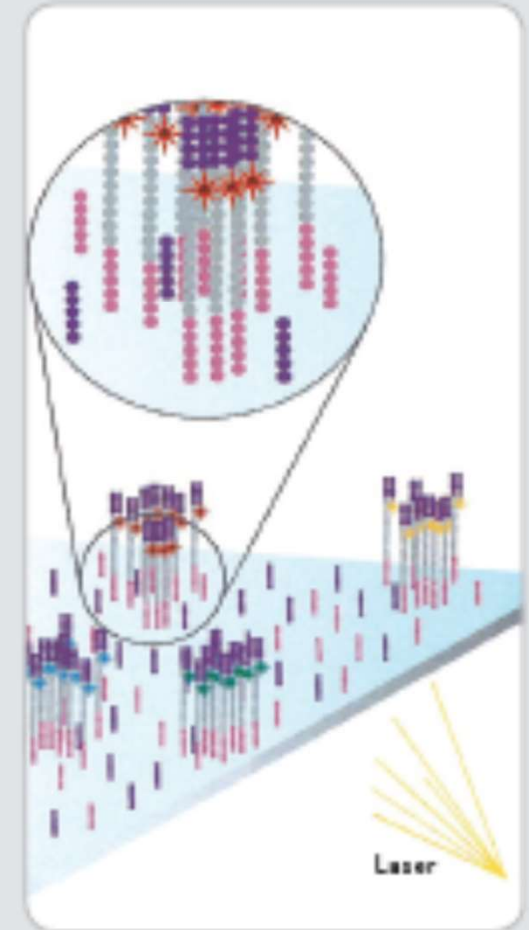
The first sequencing cycle begins by adding four labeled reversible terminators, primers, and DNA polymerase.

8. IMAGE FIRST BASE



After laser excitation, the emitted fluorescence from each cluster is captured and the first base is identified.

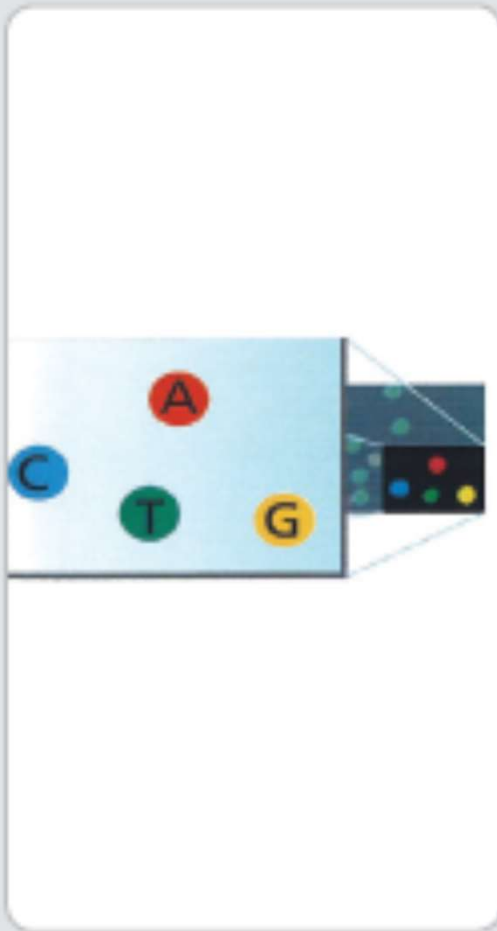
9. DETERMINE SECOND BASE



The next cycle repeats the incorporation of four labeled reversible terminators, primers, and DNA polymerase.

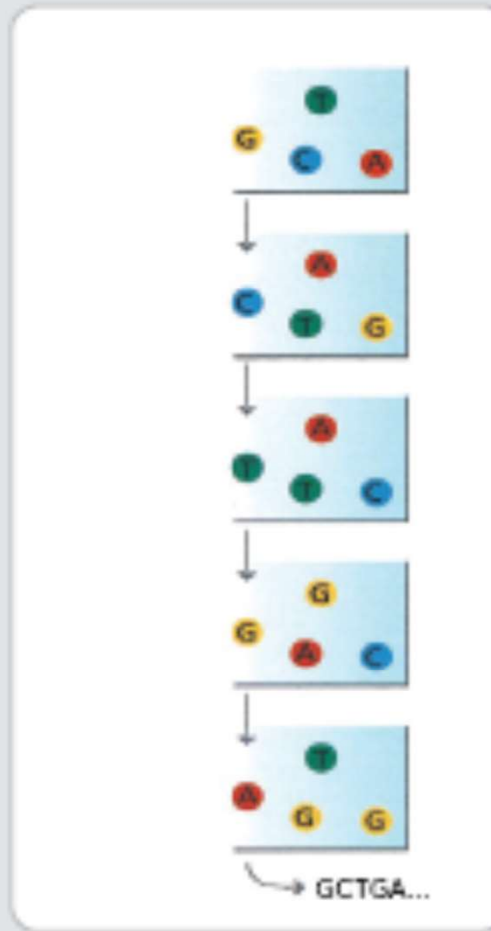
Sequencing Technology Overview

10. IMAGE SECOND CHEMISTRY CYCLE



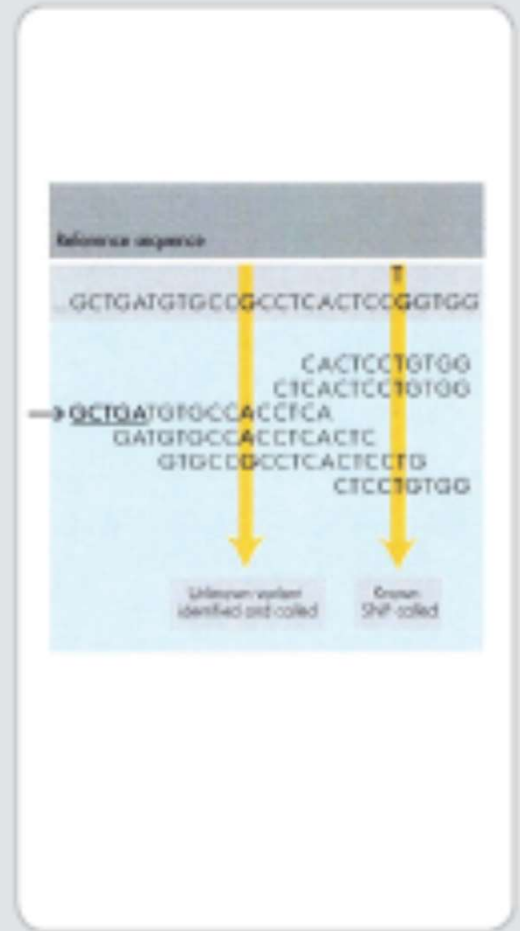
After laser excitation, the image is captured as before, and the identity of the second base is recorded.

11. SEQUENCING OVER MULTIPLE CHEMISTRY CYCLES



The sequencing cycles are repeated to determine the sequence of bases in a fragment, one base at a time.

12. ALIGN DATA

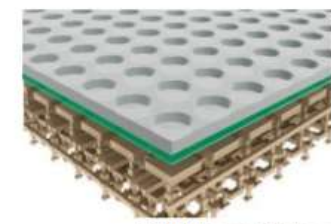
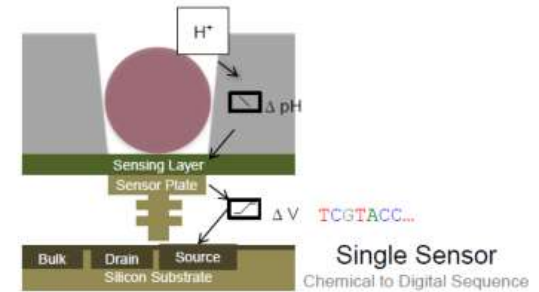
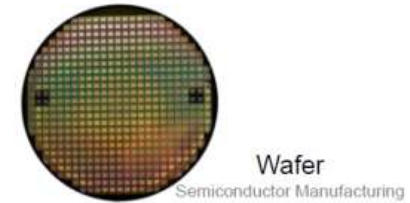
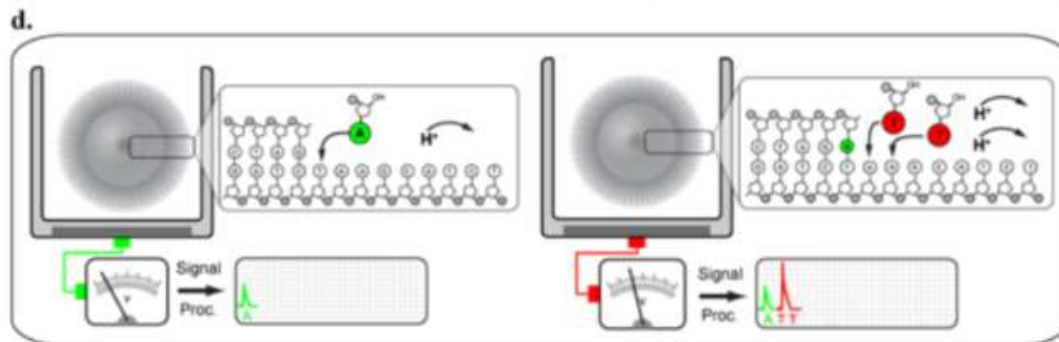
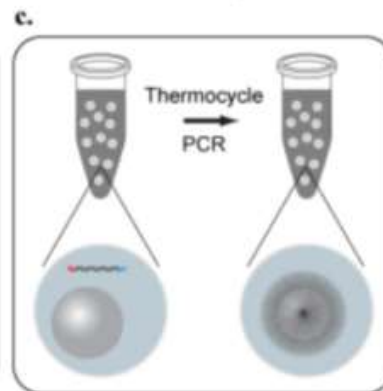
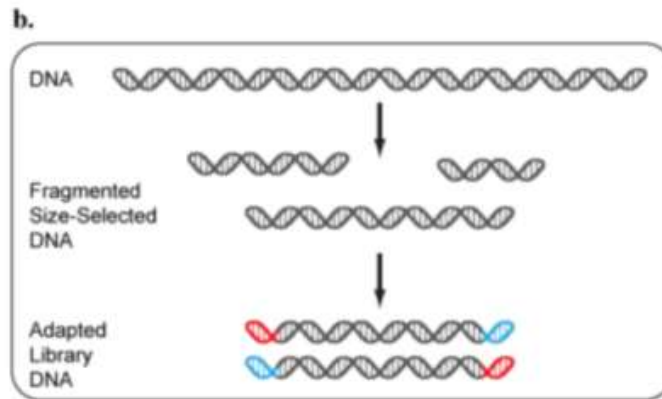
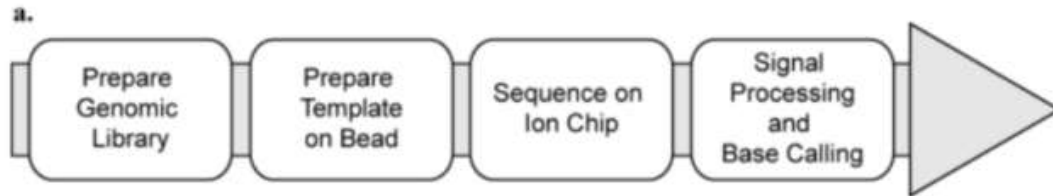


The data are aligned and compared to a reference, and sequencing differences are identified.

NGS platform: Ion Torrent platform

Sequencing method: Ion semiconductor

Library amplification method: Emulsion PCR

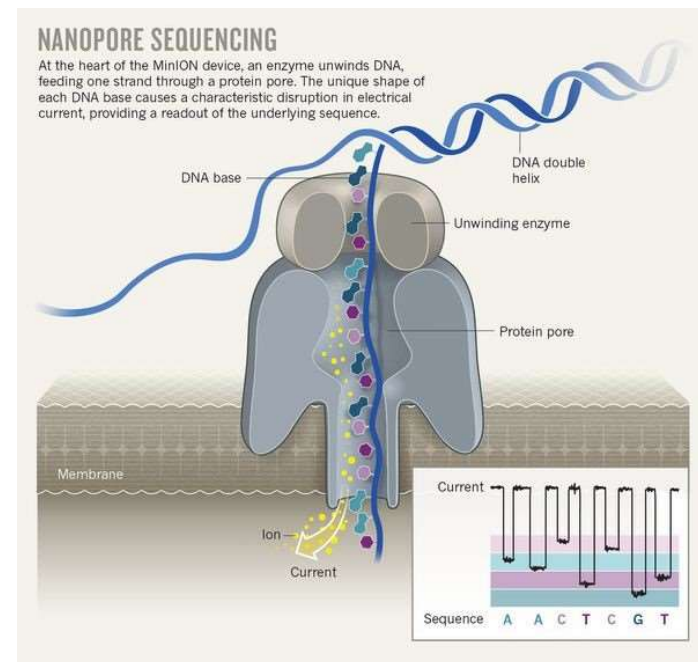
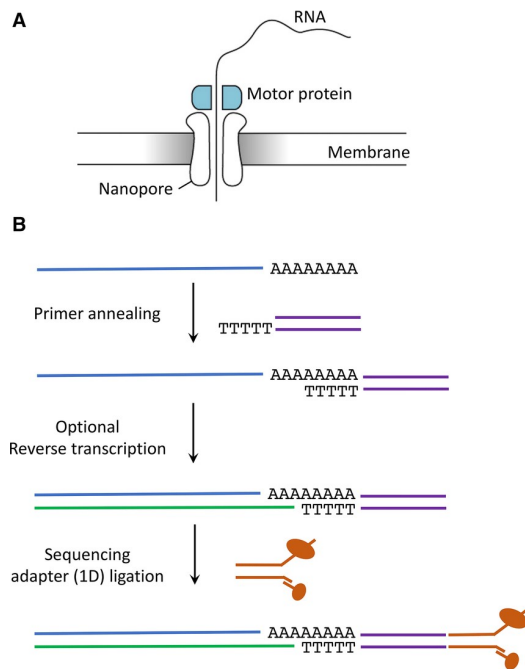


NGS platform: Nanopore Technologies

The technology produces full-length transcripts!

Direct RNA sequencing

A protein nanopore is set in an electrically resistant polymer membrane. An ionic current is passed through the nanopore by setting a voltage across this membrane. If an analyte passes through the pore or near its aperture, this event creates a characteristic disruption in current. Measurement of that current makes it possible to identify the molecule in question.



A strand of RNA is passed through a nanopore. The current is changed as the bases G, A, T and C pass through the pore in different combinations!

NGS platform: Nanopore Technologies



MinION

- Pocket-sized, portable device for biological analysis
- Up to 512 nanopore channels
- Simple 10-min sample prep available
- Real-time analysis for rapid, efficient workflows
- Adaptable to direct DNA or RNA sequencing
- [MinIT](#) available to support IT/software needs

Choose MinION if you:

- would like access to sequencing for \$1,000
- want to sequence immediately, not wait
- want to sequence outside a lab
- need 10–20Gb per 48 hrs
- want to avoid CapEx investments.

Coming soon



SmidgION

- Designed to be our smallest sequencing device so far
- Same nanopore sensing technology as MinION and PromethION
- Designed for use with a smartphone in any location

RNA-seq: analisi dei dati

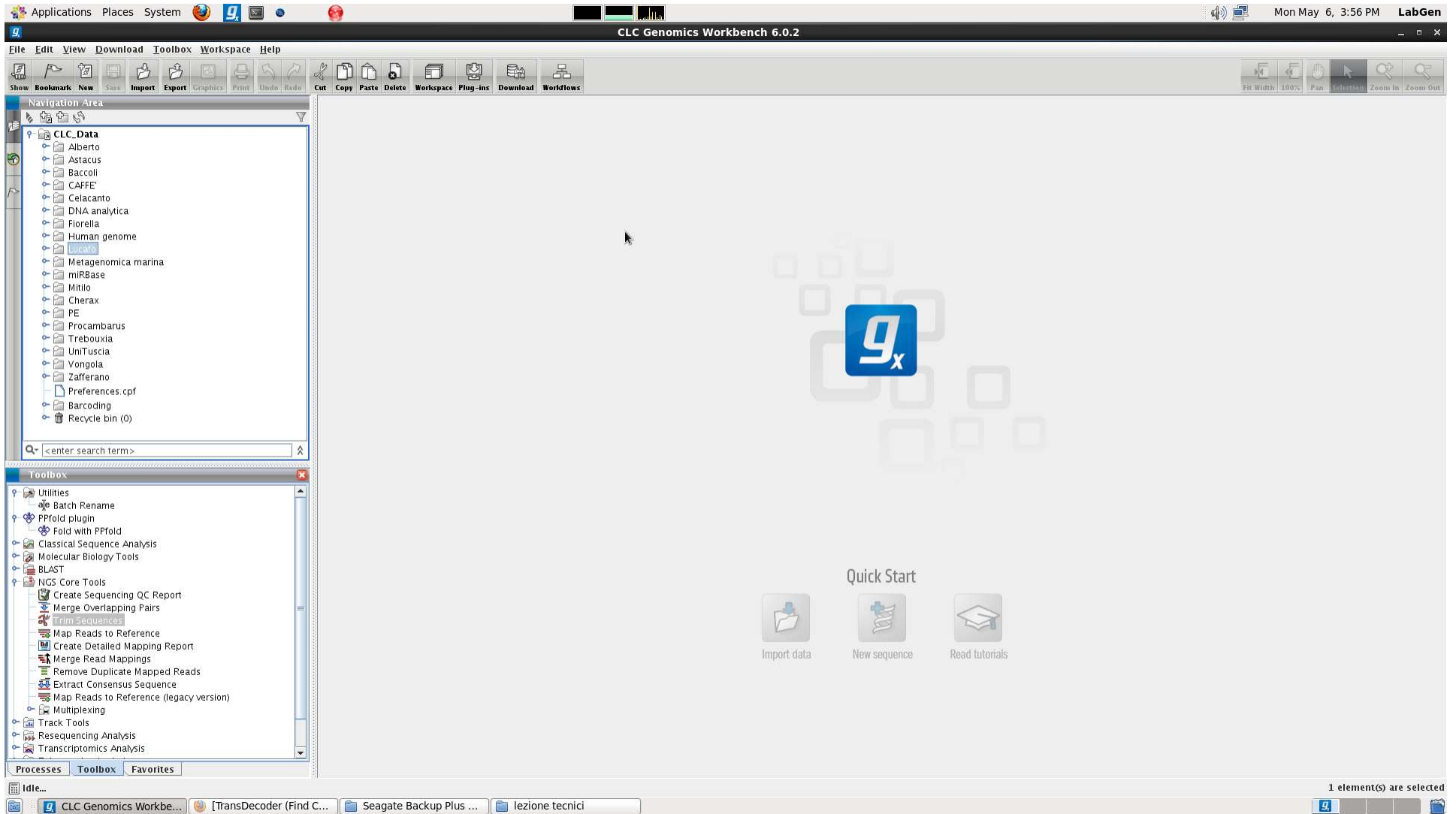
- Sequenza fornita in output dal sequenziatore = READ
- Formato dei file di output: FASTQ;
- Software: CLC Genomics Workbench (A comprehensive and user-friendly analysis package for analyzing, comparing, and visualizing next generation sequencing data)

TAPPE PRINCIPALI:

- 1) TRIMMING
- 2) ALLINEAMENTO (MAPPAGGIO)
- 3) CONTA DELLE READS
- 4) ANALISI STATISTICA
- 5) ANALISI SECONDARIE

CLC Genomics Workbench

TRIMMING



TRIMMING

The screenshot displays the CLC Genomics Workbench 6.0.2 interface. The main window shows a project tree on the left with 'CLC_Data' expanded to 'Lucafo'. A 'Trim Sequences' dialog box is open, titled '1. Select sequencing data'. The dialog has a 'Navigation Area' on the left and a 'Selected Elements (0)' area on the right. The 'Navigation Area' shows a tree view where 'Lucafo' is selected, and a list of sequencing data elements is displayed below it. The 'Selected Elements (0)' area is currently empty. At the bottom of the dialog, there are buttons for '?', 'Previous', 'Next', 'Finish', and 'Cancel'. Below these buttons are the options 'Import data', 'New sequence', and 'Read tutorials'. The status bar at the bottom right of the dialog indicates '1 element(s) are selected'. The main window's status bar shows 'Mon May 6, 3:57 PM' and 'LabGen'.

CLC Genomics Workbench 6.0.2

1. Select sequencing data

Select sequencing data

Navigation Area

Selected Elements (0)

CLC_Data

- Alberto
- Astacus
- Baccoli
- CAFFE'
- Celacanto
- DNA analytica
- Fiorella
- Human genome
- Lucafo
- Metagenomica marina
- miRBase
- Mitilo
- Cherax
- PE
- Procambarus
- Trebouxia
- UniTuscia
- Vongola
- Zafferano
- Preferences.cpf
- Barcoding
- Recycle bin (0)

Lucafo

- ctrl-12h-103_GCCAAT_L002_R1_001
- ctrl-24h-175_TAGCTT_L002_R1_001
- ctrl-48h-297_AGTCAA_L002_R1_001
- f2-12h-85_CAGATC_L002_R1_001
- f2-24h-220_GGTAC_L002_R1_001
- f2-48h-228_CGTACG_L002_R1_001
- f3-12h-99_GATCAG_L002_R1_001
- f3-24h-198_CTTGTA_L002_R1_001
- f3-48h-155_GAGTGG_L002_R1_001

sequence trimmate

- RNA-seq+ mitochondrio
- reads mapping mitochondriale
- RNaseq_lincRNA
- Metagenomica marina

Batch Select sequencing data

Previous Next Finish Cancel

Import data New sequence Read tutorials

1 element(s) are selected

TRIMMING

Applications Places System Mon May 6, 3:58 PM LabGen

CLC Genomics Workbench 6.0.2

File Edit View Download Toolbox Workspace Help

Show Bookmark New Save Import Export Graphics Print Undo Redo Cut Copy Paste Delete Workspace Plug-ins Download Workflows

Fit Width 100% Pan Selection Zoom In Zoom Out

Navigation Area

- CLC_Data
 - Alberto
 - Astacus
 - Baccoli
 - CAFFE'
 - Celacanto
 - DNA analytica
 - Fiorella
 - Human genome
 - Lucafo
 - Metagenomica marina
 - miRBase
 - Mitilo
 - Cherax
 - PE
 - Procambarus
 - Trebouxia
 - UniTuscia
 - Vongola
 - Zafferano
 - Preferences.cpf
 - Barcoding
 - Recycle bin (0)

Search: <enter search term>

Toolbox

- Utilities
 - Batch Rename
- PFfold plugin
 - Fold with PFfold
- Classical Sequence Analysis
- Molecular Biology Tools
- BLAST
- NCS Core Tools
 - Create Sequencing QC Report
 - Merge Overlapping Pairs
 - Trim Sequences
 - Map Reads to Reference
 - Create Detailed Mapping Report
 - Merge Read Mappings
 - Remove Duplicate Mapped Reads
 - Extract Consensus Sequence
 - Map Reads to Reference (legacy version)
- Multiplexing
- Track Tools
- Resequencing Analysis
- Transcriptomics Analysis

Processes Toolbox Favorites

Idle... 9 element(s) are selected

CLC Genomics Workbe... [TransDecoder (Find C... Seagate Backup Plus ...

Trim Sequences

1. Select sequencing data

Select sequencing data

Navigation Area

- CLC_Data
 - Alberto
 - Astacus
 - Baccoli
 - CAFFE'
 - Celacanto
 - DNA analytica
 - Fiorella
 - Human genome
 - Lucafo
 - sequenze grezze
 - ctrl-12h-103_GCCAAT_L002_R1_001
 - ctrl-24h-175_TAGCTT_L002_R1_001
 - ctrl-48h-297_AGTCAA_L002_R1_001
 - f2-12h-85_CAGATC_L002_R1_001
 - f2-24h-220_CGCTAC_L002_R1_001
 - f2-48h-228_CGTACG_L002_R1_001
 - f3-12h-99_GATCAG_L002_R1_001
 - f3-24h-198_CTTGTA_L002_R1_001
 - f3-48h-155_GAGTGG_L002_R1_001
 - sequenze trimmate
 - RNA-seq+ mitocondrio
 - reads mapping mitocondriale
 - RNaseq_lincRNA
 - Metagenomica marina

Search: <enter search term>

Batch

? ?

Previous Next Finish Cancel

Import data New sequence Read tutorials

Selected Elements (9)

- ctrl-12h-103_GCCAAT_L002_R1_001
- ctrl-24h-175_TAGCTT_L002_R1_001
- ctrl-48h-297_AGTCAA_L002_R1_001
- f2-12h-85_CAGATC_L002_R1_001
- f2-24h-220_CGCTAC_L002_R1_001
- f2-48h-228_CGTACG_L002_R1_001
- f3-12h-99_GATCAG_L002_R1_001
- f3-24h-198_CTTGTA_L002_R1_001
- f3-48h-155_GAGTGG_L002_R1_001

TRIMMING

The screenshot displays the CLC Genomics Workbench 6.0.2 interface. The main window shows a navigation area on the left with a tree view of data folders, including 'CLC_Data' and various sample folders like 'Alberto', 'Astacus', and 'Baccoli'. Below the navigation area is a 'Toolbox' with various analysis tools, and at the bottom, a 'Processes' and 'Favorites' section.

The 'Trim Sequences' dialog box is open, showing the 'Set parameters' tab. It includes a list of steps: '1. Select sequencing data', '2. Quality trimming', and '3. Adapter trimming'. The 'Adapter trimming' section is active, featuring a 'Trim adapter list' text box, a 'Use color space' checkbox, and a 'Search on both strands' checkbox. Below this is a 'Preview' section with a table showing summary statistics:

Name	Matches found	Reads discarded	Nucleotides removed	Avg. length

At the bottom of the dialog, there are buttons for '?', a magnifying glass, 'Previous', 'Next', 'Finish', and 'Cancel'. Below these buttons are links for 'Import data', 'New sequence', and 'Read tutorials'. The system tray at the bottom right indicates '9 element(s) are selected'.

TRIMMING

The screenshot displays the CLC Genomics Workbench 6.0.2 interface. The main window shows a navigation area on the left with a tree view of data folders, including 'CLC_Data' and various sample folders like 'Alberto', 'Astacus', and 'Lucato'. Below the navigation area is a 'Toolbox' with various analysis tools, including 'Trim Sequences'. The 'Trim Sequences' dialog box is open, showing a list of steps: 1. Select sequencing data, 2. Quality trimming, 3. Adapter trimming, and 4. Sequence filtering. The 'Set parameters' tab is active, showing options for 'Trim bases' and 'Filter on length'. The 'Trim bases' section has two checkboxes: 'Remove 5' terminal nucleotides' and 'Remove 3' terminal nucleotides', each with a spin box set to 1. The 'Filter on length' section has two checkboxes: 'Discard reads below length...' and 'Discard reads above length', with spin boxes set to 75 and 1,000 respectively. At the bottom of the dialog, there are buttons for '?', a hand icon, 'Previous', 'Next', 'Finish', and 'Cancel'. Below the dialog, there are links for 'Import data', 'New sequence', and 'Read tutorials'. The system tray at the bottom right shows '9 element(s) are selected'.

Applications Places System Mon May 6, 3:58 PM LabGen

CLC Genomics Workbench 6.0.2

File Edit View Download Toolbox Workspace Help

Show Bookmark New Save Import Export Graphics Print Undo Redo Cut Copy Paste Delete Workspace Plug-ins Download Workflows

Fix Width 100% Pan Selection Zoom In Zoom Out

Navigation Area

- CLC_Data
 - Alberto
 - Astacus
 - Baccoli
 - CAFFE
 - Celacanto
 - DNA analytica
 - Fiorella
 - Human genome
 - Lucato
 - Metagenomica marina
 - miRBase
 - Mitilo
 - Cherax
 - PE
 - Procambarus
 - Trebouxia
 - UniTuscia
 - Vongola
 - Zafferano
 - Preferences.cpf
 - Barcoding
 - Recycle bin (0)

Q- <enter search term>

Toolbox

- Utilities
 - Batch Rename
- PPfold plugin
 - Fold with PPfold
- Classical Sequence Analysis
- Molecular Biology Tools
- BLAST
- NCS Core Tools
 - Create Sequencing QC Report
 - Merge Overlapping Pairs
 - Trim Sequences
 - Map Reads to Reference
 - Create Detailed Mapping Report
 - Merge Read Mappings
 - Remove Duplicate Mapped Reads
 - Extract Consensus Sequence
 - Map Reads to Reference (legacy version)
- Multiplexing
- Track Tools
- Resequencing Analysis
- Transcriptomics Analysis

Processes Toolbox Favorites

Idle...

CLC Genomics Workbe... [TransDecoder (Find C... Seagate Backup Plus ...

9 element(s) are selected

Trim Sequences

- Select sequencing data
- Quality trimming
- Adapter trimming
- Sequence filtering

Set parameters

Trim bases

- Remove 5' terminal nucleotides
- Remove 3' terminal nucleotides

Filter on length

- Discard reads below leng...
- Discard reads above length

? Previous Next Finish Cancel

Import data New sequence Read tutorials

TRIMMING

The screenshot displays the CLC Genomics Workbench 6.0.2 interface. The main window shows a navigation area on the left with a tree view of data folders, including 'CLC_Data' and various project folders. Below the navigation area is a 'Toolbox' panel containing various analysis tools. The 'Trim Sequences' dialog box is open in the center, showing a list of steps: 1. Select sequencing data, 2. Quality trimming, 3. Adapter trimming, 4. Sequence filtering, and 5. Result handling. The 'Result handling' step is selected, and the dialog shows options for output, result, and log handling. The 'Output options' section includes checkboxes for 'Save discarded sequences', 'Save broken pairs', and 'Create report'. The 'Result handling' section has radio buttons for 'Open' and 'Save'. The 'Log handling' section has a checkbox for 'Make log'. At the bottom of the dialog, there are buttons for 'Previous', 'Next', 'Finish', and 'Cancel'. The status bar at the bottom right indicates '9 element(s) are selected'.

Applications Places System Mon May 6, 3:59 PM LabGen

CLC Genomics Workbench 6.0.2

File Edit View Download Toolbox Workspace Help

Show Bookmark New Save Import Export Graphics Print Undo Redo Cut Copy Paste Delete Workspace Plug-ins Download Workflows

Fit Width 100% Pan Selection Zoom In Zoom Out

Navigation Area

- CLC_Data
 - Alberto
 - Astacus
 - Baccoli
 - CAFFE
 - Celacanto
 - DNA analytica
 - Fiorella
 - Human genome
 - Lucato
 - Metagenomica marina
 - miRBase
 - Mitilo
 - Cherax
 - PE
 - Procambarus
 - Trebouxia
 - UniTuscia
 - Vongola
 - Zafferano
 - Preferences.cpf
 - Barcoding
 - Recycle bin (0)

Toolbox

- Utilities
 - Batch Rename
- PPfold plugin
 - Fold with PPfold
- Classical Sequence Analysis
- Molecular Biology Tools
- BLAST
- NCS Core Tools
 - Create Sequencing QC Report
 - Merge Overlapping Pairs
 - Trim Sequences
 - Map Reads to Reference
 - Create Detailed Mapping Report
 - Merge Read Mappings
 - Remove Duplicate Mapped Reads
 - Extract Consensus Sequence
 - Map Reads to Reference (legacy version)
- Multiplexing
- Track Tools
- Resequencing Analysis
- Transcriptomics Analysis

Processes Toolbox Favorites

Idle...

CLC Genomics Workbe... [TransDecoder (Find C... Seagate Backup Plus ...

9 element(s) are selected

Trim Sequences

- Select sequencing data
- Quality trimming
- Adapter trimming
- Sequence filtering
- Result handling

Result handling

Output options

- Save discarded sequences
- Save broken pairs
- Create report

Result handling

- Open
- Save

Log handling

- Make log

? Previous Next Finish Cancel

Import data New sequence Read tutorials

TRIMMING

The screenshot displays the CLC Genomics Workbench 6.0.2 interface. The main window shows a project named 'grezze' with a 'trimmate' sub-project. The 'Navigation Area' on the left lists various files, including trimmed reads (e.g., 'ctrl-12h_trimmed', 'f2-12h_trimmed') and their corresponding reports. The 'Toolbox' on the left contains various analysis tools, with 'Trim Sequences' highlighted. The main workspace features a large 'g_x' logo and a 'Quick Start' section with buttons for 'Import data', 'New sequence', and 'Read tutorials'. The status bar at the bottom indicates '1 element(s) are selected'.

Applications Places System Mon May 6, 4:02 PM LabGen

CLC Genomics Workbench 6.0.2

File Edit View Download Toolbox Workspace Help

Show Bookmark New Save Import Export Graphics Print Undo Redo Cut Copy Paste Delete Workspace Plug-ins Download Workflows

Fit Width 100% Pan Selection Zoom In Zoom Out

Navigation Area

- grezze
 - sequenze grezze
 - sequenze trimmate
 - ctrl-12h_trimmed
 - ctrl-24h_trimmed
 - ctrl-48h_trimmed
 - f2-12h_trimmed
 - f2-24h_trimmed
 - f2-48h_trimmed
 - f3-12h_trimmed
 - f3-24h_trimmed
 - f3-48h_trimmed
 - ctrl-12h_trimmed_report
 - ctrl-24h_trimmed_report
 - ctrl-48h_trimmed_report
 - f2-12h_trimmed_report
 - f2-24h_trimmed_report
 - f2-48h_trimmed_report
 - f3-12h_trimmed_report
 - f3-24h_trimmed_report
 - f3-48h_trimmed_report
 - RNA-seq+mitochondrio
 - reads mapping mitochondriale
 - RNAseq_lincRNA
 - Metagenomics_marina

Toolbox

- Utilities
 - Batch Rename
- PPfold plugin
 - Fold with PPfold
- Classical Sequence Analysis
- Molecular Biology Tools
- BLAST
- NCS Core Tools
 - Create Sequencing QC Report
 - Merge Overlapping Pairs
 - Trim Sequences
 - Map Reads to Reference
 - Create Detailed Mapping Report
 - Merge Read Mappings
 - Remove Duplicate Mapped Reads
 - Extract Consensus Sequence
 - Map Reads to Reference (legacy version)
- Multiplexing
- Track Tools
- Resequencing Analysis
- Transcriptomics Analysis

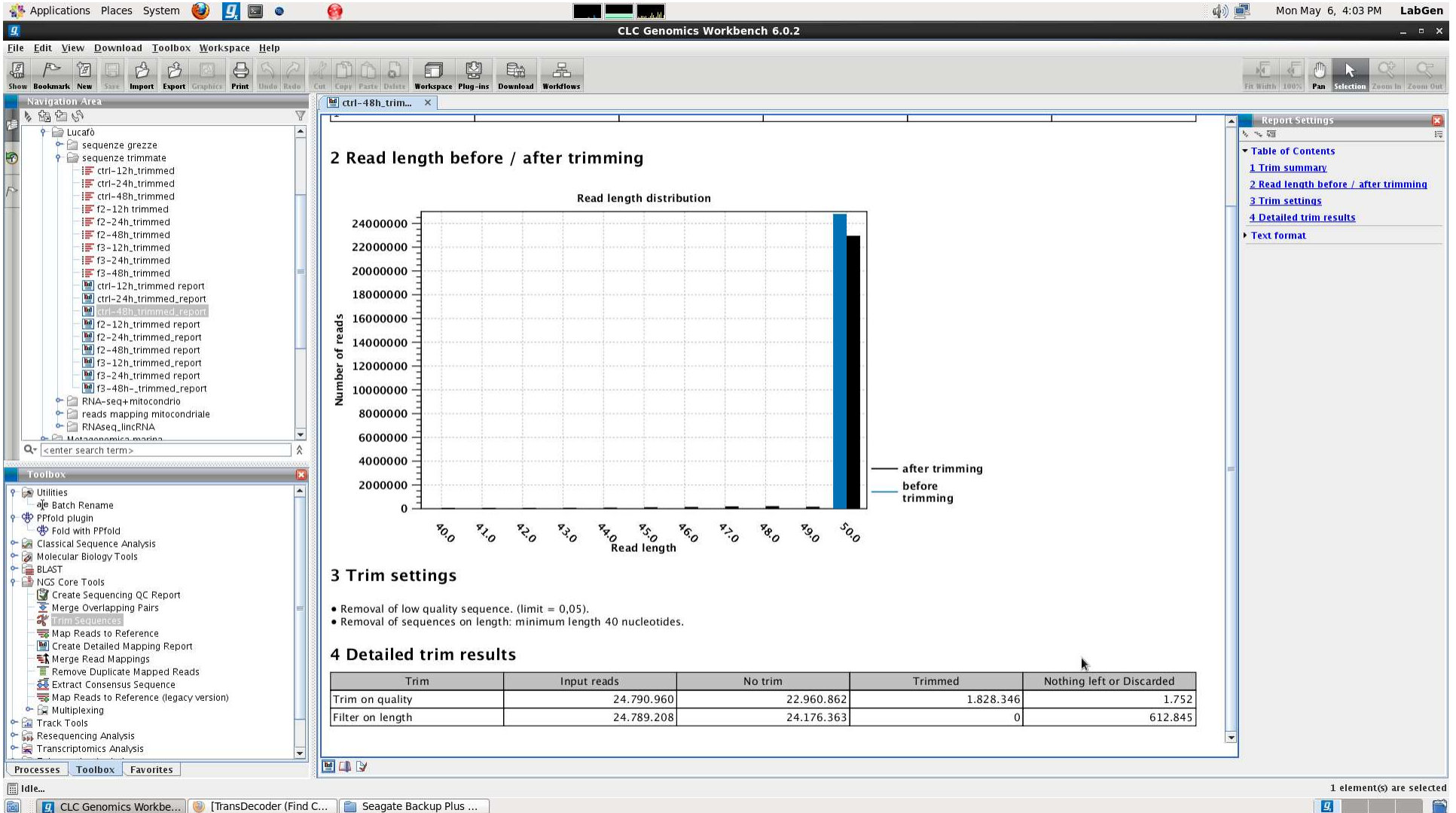
Processes Toolbox Favorites

Idle...

CLC Genomics Workbe... [TransDecoder (Find C... Seagate Backup Plus ...

1 element(s) are selected

TRIMMING: report finale



2) ALLINEAMENTO

Applications Places System Mon May 6, 4:04 PM LabGen

Ensembl genome browser 71: Homo sapiens - Description - Mozilla Firefox

File Edit View History Bookmarks Tools Help

TransDecoder (Find Coding ... TransDecoder (Find Coding ... Ensembl genome browser 7...

www.ensembl.org/Homo_sapiens/info/index

minibarcoding

Human (GRCh37) Login/Register

Human

Homo sapiens

Search Human...

e.g. BRCA2 or 6:133017695-133161157 or osteoarthritis

Genome assembly: GRCh37 (GCA_000001405.11)

- More information and statistics
- Download DNA sequence (FASTA)
- Convert your data to GRCh37 coordinates
- Display your data in Ensembl

Other assemblies

- NCBI36 (Ensembl release 54)

What's New in Human release 71

- Update to Ensembl-Havana GENCODE gene set (release 16)
- Added Kidney RNASeq models and intron supporting features
- Human: updated cDNA alignments

More news...

Gene annotation

What can I find? Protein-coding and non-coding genes, splice variants, cDNA and protein sequences, non-coding RNAs.

- More about this genebuild
- Download genes, cDNAs, ncRNA, proteins (FASTA)
- Update your old Ensembl IDs

Vega Additional manual annotation can be found in Vega

Example gene: Pax6, INS, FOXP2, BRCA2, DMU, ssh

Example transcript

Comparative genomics

What can I find? Homologues, gene trees, and whole genome alignments across multiple species.

- More about comparative analysis
- Download alignments (EMF)

Example gene tree

Regulation

What can I find? DNA methylation, transcription factor binding sites, histone modifications, and regulatory features such as enhancers and repressors, and microarray annotations.

- More about the Ensembl regulatory build and microarray annotation
- Download all regulatory features (GFF)

Example regulatory feature

ENCODE data in Ensembl

Variation

What can I find? Short sequence variants and longer structural variants; disease and other phenotypes

- More about variation in Ensembl
- Download all variants (GVF)
- Variant Effect Predictor *vep*

Example variant: ATCGAGCT, ATCCAGCT, ATCGAGAT

Example phenotype:

Example structural variant:

CLC Genomics Workbe... Ensembl genome bro... Seagate Backup Plus ...

2) ALLINEAMENTO

The screenshot displays the CLC Genomics Workbench 6.0.3 interface. The main window shows a bar chart titled "3.2 Match specificity (single)" with the y-axis labeled "Number of reads" (0 to 1,000,000) and the x-axis labeled "Number of match positions" (0.0 to 10.0). The highest bar is at 1.0 match position, exceeding 1,000,000 reads. Other bars are significantly lower, with the next highest at 0.0 match positions (around 250,000 reads).

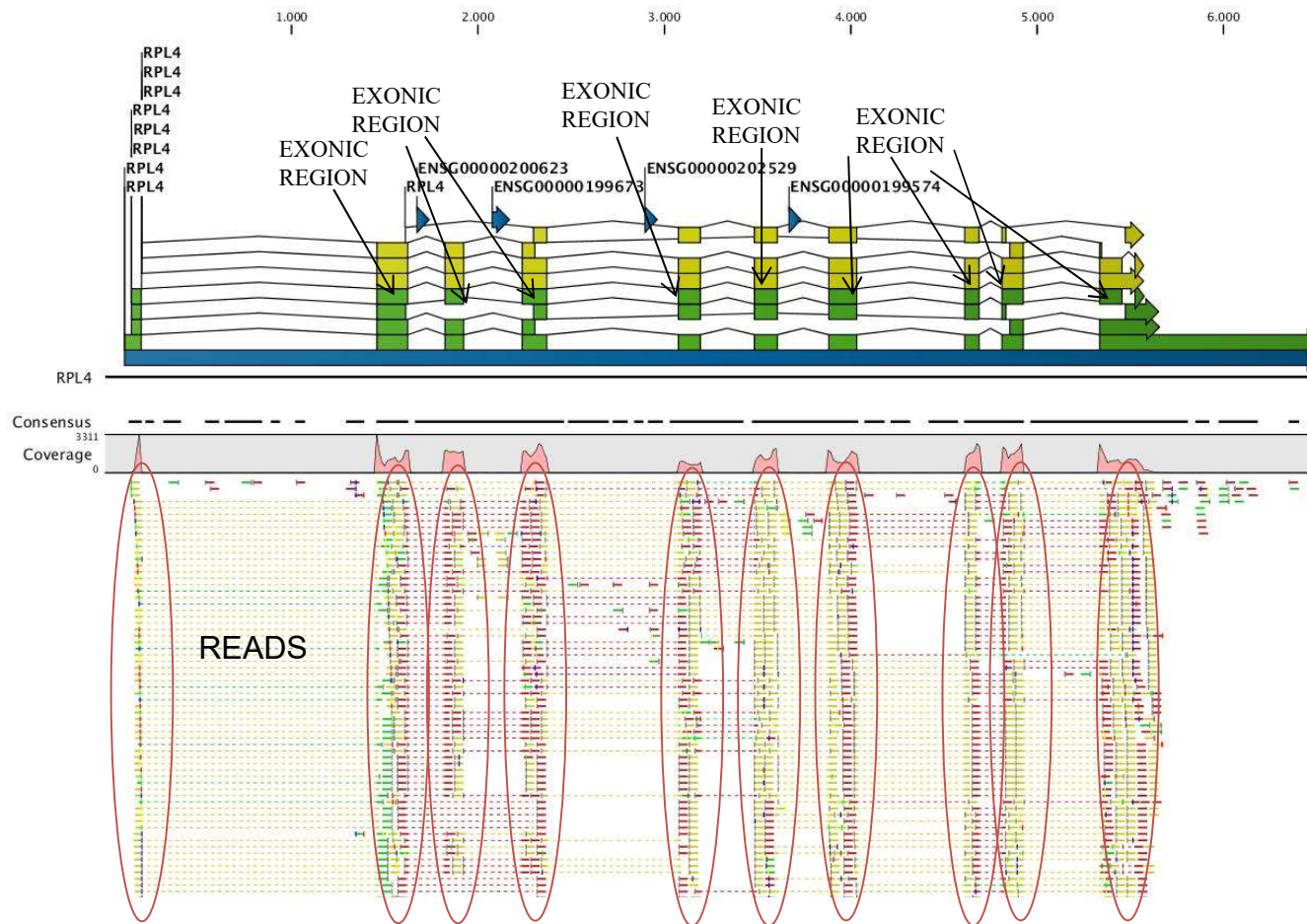
Below the chart is section "4 Detailed mapping statistics" with the following table:

	Uniquely mapped	Fraction	Non-specifically mapped	Fraction	Mapped	% of total mapped
Exon-exon	2.418.539	0,86	401.940	0,14	2.820.479	21,52
Exon-intron	44.783	0,74	15.660	0,26	60.443	0,46
Total exon	10.027.108	0,83	1.989.556	0,17	12.016.664	91,68
Total intron	571.064	0,52	518.762	0,48	1.089.826	8,32
Total gene	10.598.172	0,81	2.508.318	0,19	13.106.490	100,00

Default counting scheme ('Fragment counts'): A intact pair is counted as one, broken pairs are ignored

The interface also shows a navigation area on the left with a tree view of projects and files, and a report settings panel on the right with a table of contents.

2) ALLINEAMENTO

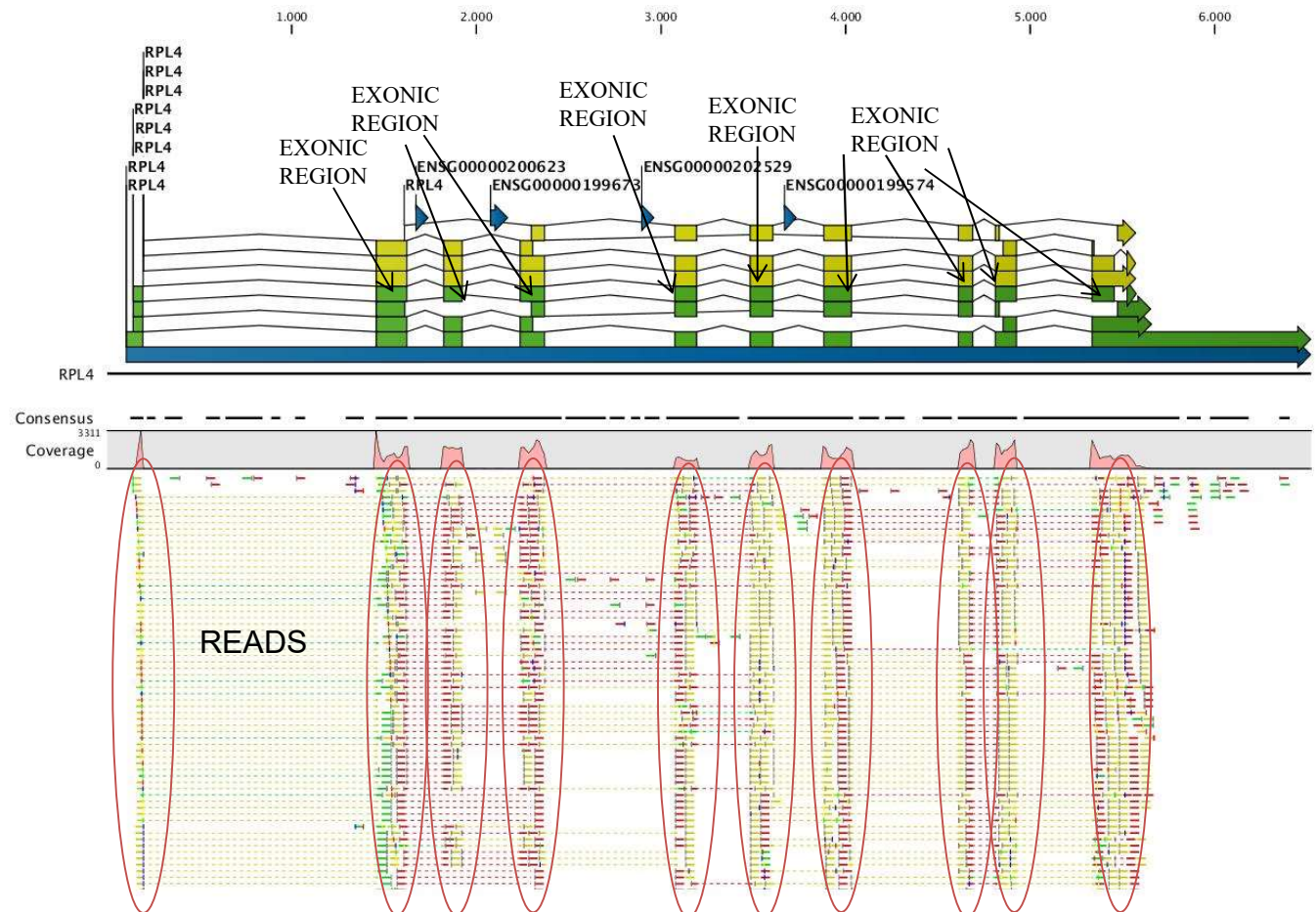


Read: Sequenza fornita in output dal sequenziatore, che identifica l'ordine in cui si susseguono le basi nei frammenti di cDNA. Il numero di basi che la compongono rappresenta la sua lunghezza (misurata in nt o bp).

3) CONTA

Numero di read mappate su una feature biologica di interesse (gene, trascritto, esone). Il numero di reads è proporzionale all'abbondanza della feature, quindi del suo livello di espressione!

RPKM (Reads Per Kilobase per Million mapped reads) :
Indice del livello di espressione
 $RPKM = \frac{\text{total read/reads mappate}[\text{milioni}] * \text{lunghezza gene} [\text{kb}]}{10^6}$



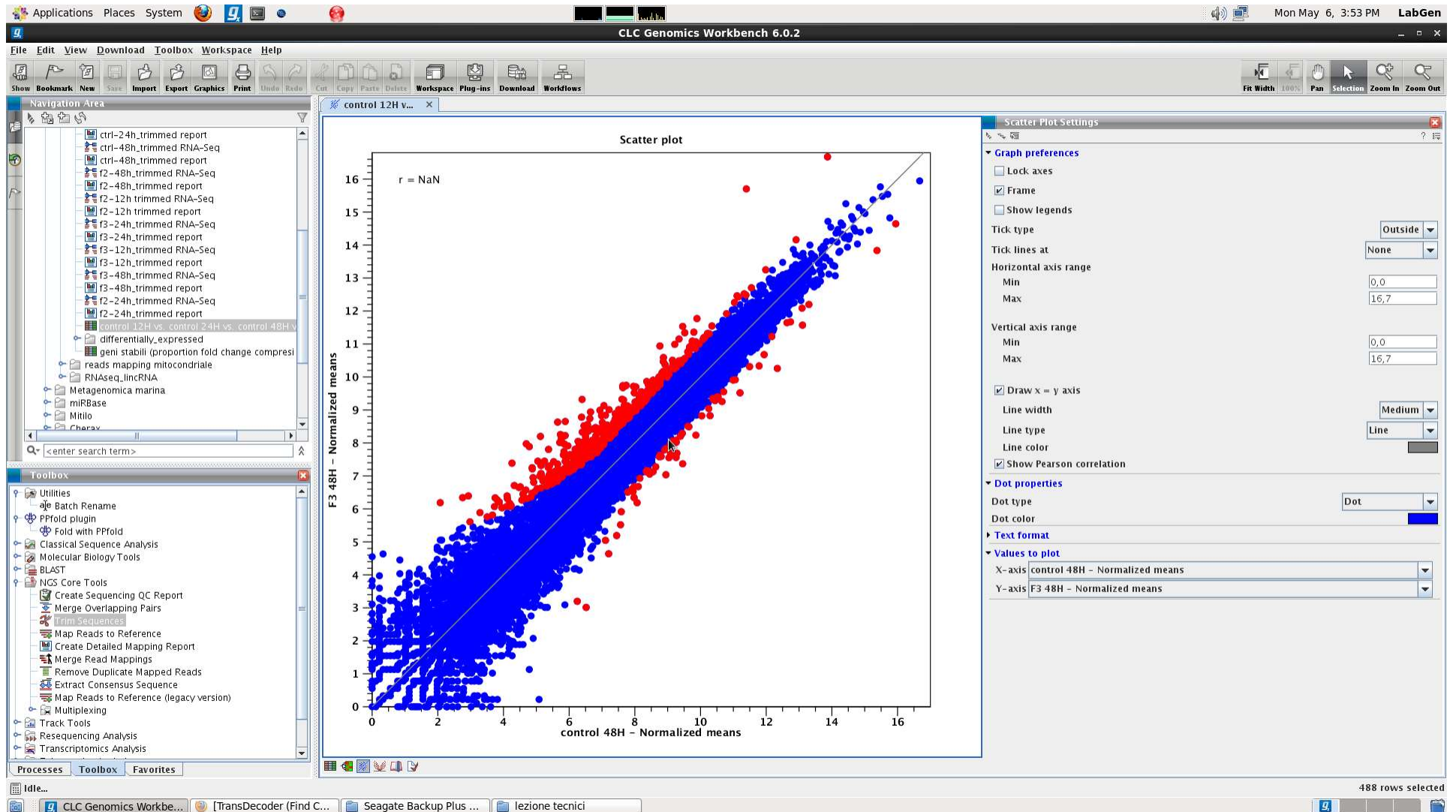
4) ANALISI STATISTICA:

Kal's Z-test (cutoff e-value di 1e-10)

Feature ID	Proportions fold change	FDR p-value correction
RP5-1182A14.1	-1,46	1,19E-12
MLL2	-1,37	0
IFRD1	-1,33	1,48E-12
SEC61A1	-1,32	1,19E-12
DST	-1,28	1,44E-12
MLL3	-1,28	2,02E-12
UBR4	-1,27	0
UGGT1	-1,27	5,78E-13
DHX9	-1,27	0
BCLAF1	-1,26	1,44E-12
DNAJC10	-1,26	1,30E-11
POLR1A	-1,26	2,50E-12
CREBBP	-1,25	1,44E-12
RANBP2	-1,25	0
PRPF19	-1,25	0
HSP90AA1	-1,25	0
TFF1	-1,24	1,84E-12
CCDC47	-1,23	1,57E-12
NPM1	-1,23	8,24E-13
SLC26A2	-1,23	0
HYOU1	-1,22	0
DDX21	-1,22	0
HSPA5	-1,22	0
HSPH1	-1,22	1,57E-12
UTP20	-1,22	7,44E-12
ANKRD11	-1,21	3,15E-12
MKI67	-1,2	0
AMAC1L3	-1,2	2,24E-12
SON	-1,19	0
SRRT	-1,19	1,73E-12
TMED9	-1,19	2,80E-11
SRCAP	-1,19	3,13E-12
PTMA	-1,18	7,48E-13
MTDH	-1,18	0

Feature ID	Proportions fold change	FDR p-value correction
ATP7B		23,65E-13
COX1		2
VEGFB		22,72E-13
SYT15		2,011,89E-11
GABBR1		2,018,80E-14
AC010336.1		2,02
TMPRSS13		2,03
RAB31		2,031,85E-13
SYT7		2,031,31E-12
GBAS		2,04
BRWD3		2,056,49E-13
EHD2		2,05
AF042090.2		2,062,11E-13
GADD45B		2,067,06E-11
RNR1		2,07
MOSPD3		2,084,42E-12
ACSF2		2,081,85E-13
ID2		2,09
TMC4		2,09
NCRNA00263		2,112,09E-11
EPOR		2,141,27E-12
SELENBP1		2,141,03E-12
ZNF704		2,15
CRELD1		2,154,29E-11
ATG16L2		2,16
AHNAK		2,17
CKB		2,171,45E-11
HMGCL		2,192,52E-12
INPP5J		2,21
SLC2A10		2,225,63E-11
DDAH2		2,225,31E-13
LTBP4		2,23
PYROXD2		2,234,74E-14

SCATTER PLOT



5) ANALISI SECONDARIE: CARATTERIZZAZIONE ONTOLOGICA

DAL DATO GREZZO AL SIGNIFICATO BIOLOGICO

Il progetto Gene Ontology fornisce una [ontologia](#) di termini definiti che rappresentano le proprietà dei prodotti dei geni

L'ontologia è suddivisa in tre ambiti:

- *Molecular function;*
- *Biological process;*
- *Cellular component;*

<http://www.geneontology.org/>



http://www.geneontology.org/



Search
gene or protein name

[Downloads](#) [Tools](#) [Documentation](#) [Projects](#) [About](#) [Contact](#)

Welcome to the Gene Ontology website!

The Gene Ontology project is a major bioinformatics initiative with the aim of standardizing the representation of gene and gene product attributes across species and databases. The project provides [a controlled vocabulary of terms](#) for describing gene product characteristics and [gene product annotation data](#) from GO Consortium members, as well as [tools to access and process this data](#). [Read more about the Gene Ontology...](#)

Search the Gene Ontology Database

Search for genes, proteins or GO terms using [AmiGO](#) :

gene or protein name GO term or ID



[AmiGO](#) is the official GO browser and search engine. [Browse the Gene Ontology with AmiGO](#).

The Gene Ontology project very much encourages input from the community into both the content of the GO and annotation using GO. We are very happy to work with others to ensure that the GO is both complete and accurate, and we also very much encourage communities to submit GO annotations for inclusion in the GO database. [Please contact us](#).

The Gene Ontology Consortium is supported by a U41 grant from the National Human Genome Research Institute (NHGRI) [grant HG002273]. [See the full list of funding sources](#). The Gene Ontology Consortium would like to acknowledge the assistance of many more people than can be listed here. Please visit the [acknowledgements page](#) for the full list.

Quick Links

Tools

- [AmiGO browser](#) 
- [Submit GO Annotations](#)
- [OBO-Edit ontology editor](#)
- [Ontology downloads](#)
- [Annotation downloads](#)
- [Database downloads](#)
- [Documentation](#)
- [GO FAQ](#)
- [GO on SourceForge](#) 
- [Contact GO](#)

News

- [GO on Twitter](#) 
- [Finding updates...](#)
- [GO newsdesk](#) 
- [GO news RSS feed](#) 
- [GO on Facebook](#) 

5) ANALISI SECONDARIE: CARATTERIZZAZIONE ONTOLOGICA

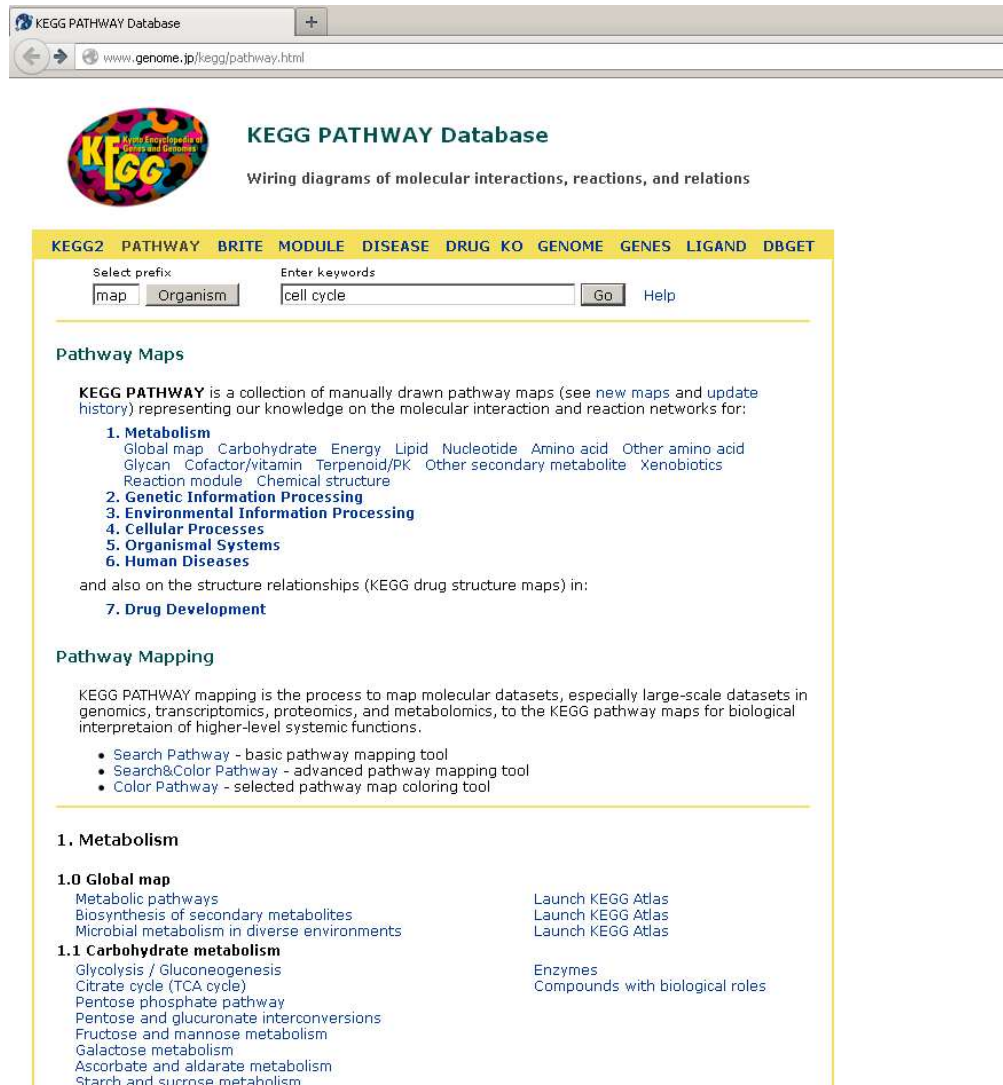
DAL DATO GREZZO AL SIGNIFICATO BIOLOGICO

GO term	Biological Process	F2 - 12H	F2 - 24H	F2 - 48H	F3 - 12H	F3 - 24H	F3 - 48H
	75 cell cycle checkpoint (Reactome:REACT_1538 [TAS])	NS	NS	**	****	***	****
	82 G1/S transition of mitotic cell cycle (PMID:8681378 [TAS])	NS	NS	**	***	***	****
	84 S phase of mitotic cell cycle (PMID:21196493 [IMP])	NS	NS	**	****	***	****
	85 G2 phase of mitotic cell cycle (PMID:9154802 [TAS])	NS	NS	*	*	*	*
	86 G2/M transition of mitotic cell cycle (Reactome:REACT_2203 [TAS])	NS	NS	NS	NS	*	*
	122 negative regulation of transcription from RNA polymerase II promoter (PMID:10973986 [IDA])	NS	NS	NS	NS	NS	*

GO term	Molecular Function	F2 - 12H	F2 - 24H	F2 - 48H	F3 - 12H	F3 - 24H	F3 - 48H
	3723 RNA binding (PMID:3323886 [TAS])	**	**	**	***	***	***
	4003 ATP-dependent DNA helicase activity (GO_REF:0000002 [IEA] InterPro:IPR000212)	NS	*	NS	NS	*	**
	4004 ATP-dependent RNA helicase activity (PMID:7610041 [TAS])	*	*	*	**	*	**
	5515 protein binding (PMID:12681488 [IPI] UniProtKB:Q7Z465)	****	****	***	***	***	***
	5524 ATP binding (GO_REF:0000037 [IEA] UniProtKB-KW:KW-0067)	*	**	***	***	***	****
	5525 GTP binding (PMID:17107948 [IDA])	NS	*	NS	*	NS	*


GO term	Cellular Compartment	F2 - 12H	F2 - 24H	F2 - 48H	F3 - 12H	F3 - 24H	F3 - 48H
	502 proteasome complex (PMID:8811196 [TAS])	*	*	**	****	****	***
	5605 basal lamina (GO_REF:0000019 [IEA] Ensembl:ENSMUSP00000005532)	NS	NS	*	**	*	NS
	5606 laminin-1 complex (PMID:10964500 [NAS])	NS	NS	NS	**	NS	NS
	5634 nucleus (PMID:19188445 [IDA])	*	NS	**	***	**	***
	5641 nuclear envelope lumen (PMID:8245774 [IDA])	NS	*	NS	**	**	**
	5643 nuclear pore (PMID:11024021 [NAS])	NS	NS	*	**	*	**

5) ANALISI SECONDARIE: PATHWAY ALTERATE



KEGG PATHWAY Database

www.genome.jp/kegg/pathway.html

 **KEGG PATHWAY Database**
Wiring diagrams of molecular interactions, reactions, and relations

KEGG2 PATHWAY BRITE MODULE DISEASE DRUG KO GENOME GENES LIGAND DBGET

Select prefix: map Organism Enter keywords: cell cycle Go Help

Pathway Maps

KEGG PATHWAY is a collection of manually drawn pathway maps (see [new maps](#) and [update history](#)) representing our knowledge on the molecular interaction and reaction networks for:

- 1. Metabolism**
 - Global map
 - Carbohydrate
 - Energy
 - Lipid
 - Nucleotide
 - Amino acid
 - Other amino acid
 - Glycan
 - Cofactor/vitamin
 - Terpenoid/PK
 - Other secondary metabolite
 - Xenobiotics
 - Reaction module
 - Chemical structure
- 2. Genetic Information Processing**
- 3. Environmental Information Processing**
- 4. Cellular Processes**
- 5. Organismal Systems**
- 6. Human Diseases**

and also on the structure relationships (KEGG drug structure maps) in:

- 7. Drug Development**

Pathway Mapping

KEGG PATHWAY mapping is the process to map molecular datasets, especially large-scale datasets in genomics, transcriptomics, proteomics, and metabolomics, to the KEGG pathway maps for biological interpretation of higher-level systemic functions.

- Search Pathway - basic pathway mapping tool
- Search&Color Pathway - advanced pathway mapping tool
- Color Pathway - selected pathway map coloring tool

1. Metabolism

1.0 Global map

Metabolic pathways
Biosynthesis of secondary metabolites
Microbial metabolism in diverse environments

[Launch KEGG Atlas](#)
[Launch KEGG Atlas](#)
[Launch KEGG Atlas](#)

1.1 Carbohydrate metabolism

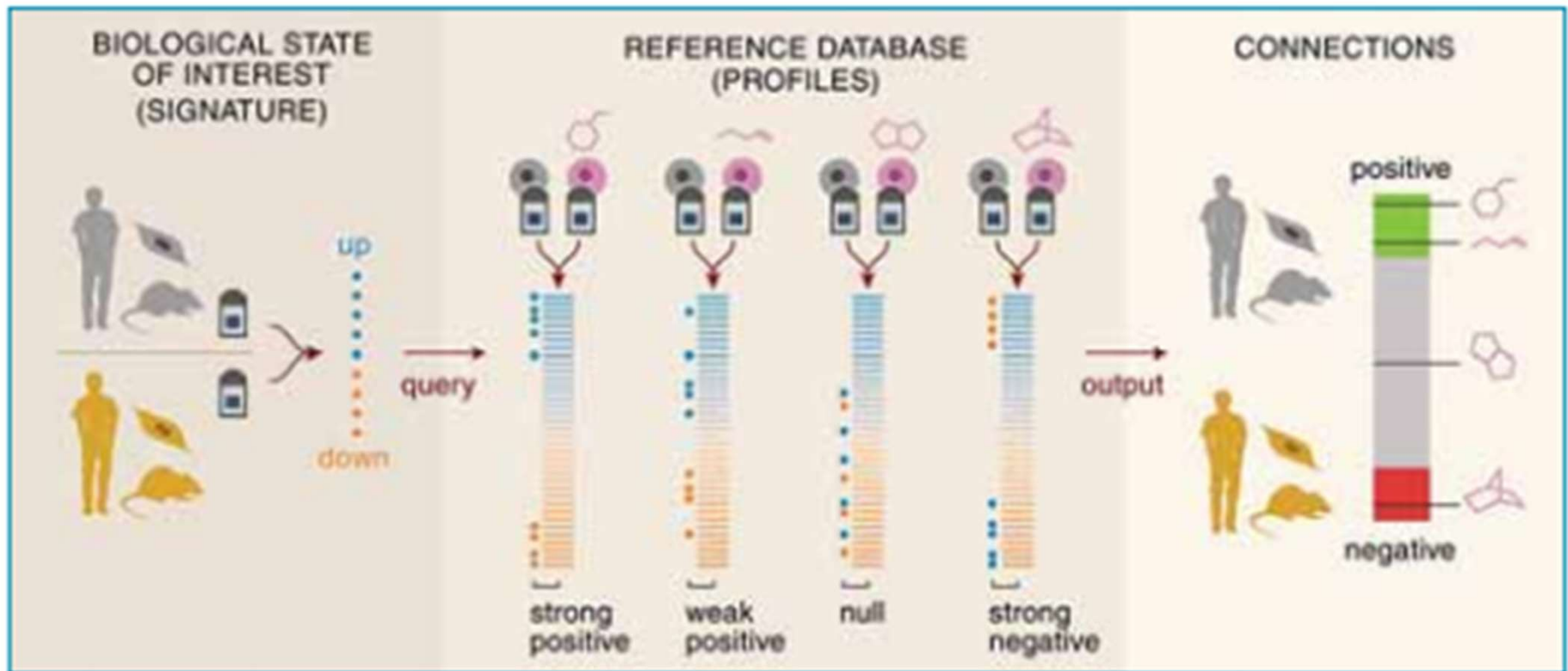
Glycolysis / Gluconeogenesis
Citrate cycle (TCA cycle)
Pentose phosphate pathway
Pentose and glucuronate interconversions
Fructose and mannose metabolism
Galactose metabolism
Ascorbate and aldarate metabolism
Starch and sucrose metabolism

Enzymes
Compounds with biological roles

<http://www.genome.jp/kegg/pathway.html>

5) ANALISI SECONDARIE: CONNECTIVITY MAP

<http://www.broadinstitute.org/cmap/>



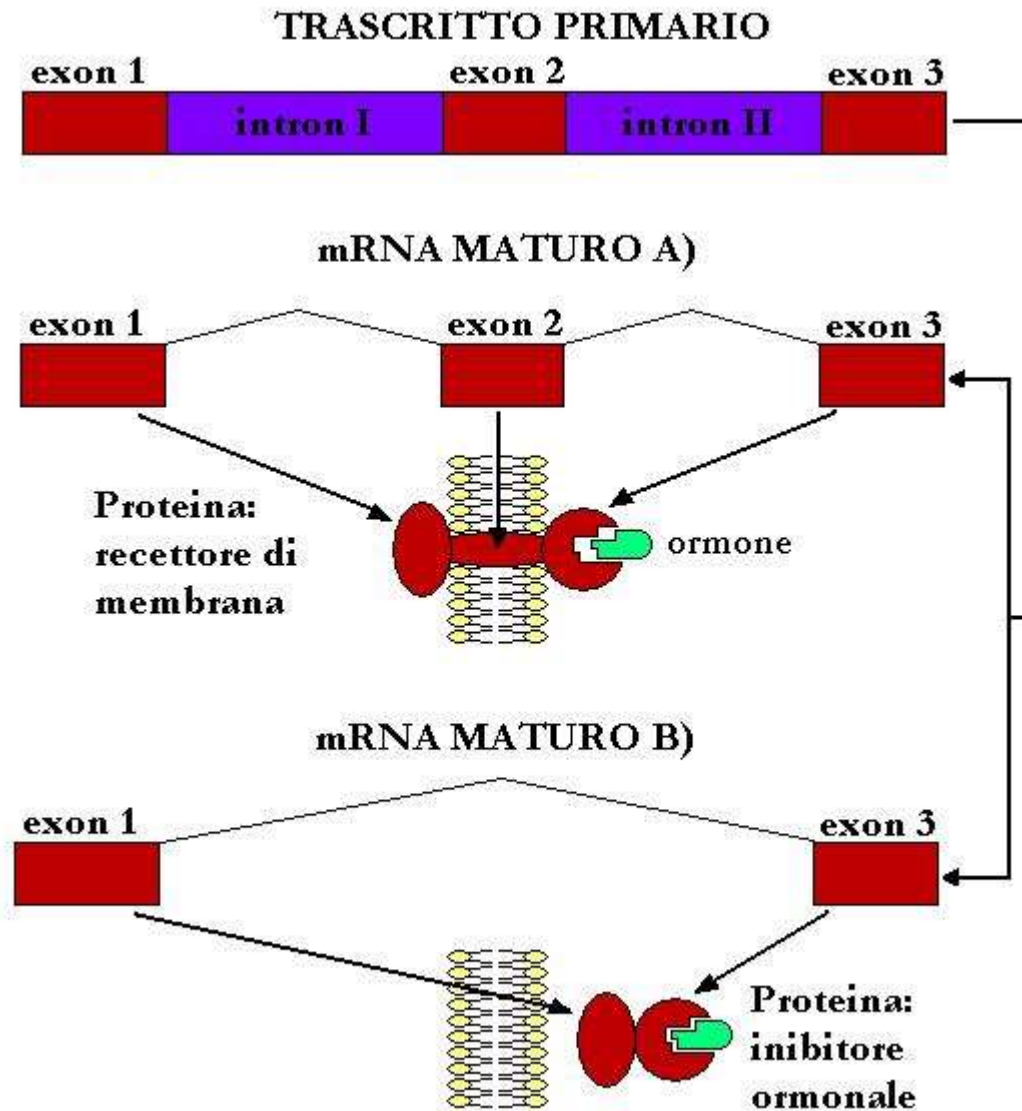
5) ANALISI SECONDARIE: CONNECTIVITY MAP

<http://www.broadinstitute.org/cmap/>

CMAP NAME	SPECIFICITY	P-VALUE						
		F2 - 12 H	F2 - 24 H	F2 - 48 H	F3 - 12 H	F3 - 24 H	F3 - 48 H	
Wortmannin	**	ns	0	0	0	0	0	= inhibitors of PI3K signal transduction pathway;
Trichostatin A	ns	ns	0	0	0	0	0	
Quinostatin	*	ns	0,00404	0	0.00002	0.00002	0	
Syrosingopine	**	ns	0.00258	0.00109	0.00012	0.02386	0.00036	= inhibitors of histone deacetylase enzymes ;
Rescinnamine	*	ns	0.02604	0.00505	0.00382	0.00264	0.00078	
Tonzonium bromide	*	ns	0.03074	0.00419	0.00658	0.00097	0.00074	
Clomifene	*	ns	0.00734	0.03754	0.00197	0.00734	0.00221	= antipsychotic drugs
Fluphenazine	ns	ns	0.05081	0.00183	0.00325	0.00377	0.00423	
Flunarizine	*	ns	0.06261	0.00249	0.01998	0.0041	0.00505	
Homochlorethazine	*	ns	0.04123	0.00366	0.0585	0.00853	0.00207	
Sirolimus	*	ns	0	ns	0	0	0	
Fulvestrant	*	ns	ns	0.00006	0	0	0	
Vorinostat	ns	ns	0.00228	ns	0	0.00018	0	
Trifluoperazine	ns	ns	0.00332	ns	0.0004	0.00689	0.00012	
Amiodarone	*	ns	ns	0.01693	0.00112	0.00326	0.00054	
0297417-0002B	*	ns	0.05036	ns	0.00145	0.00091	0.00052	
Rottlerin	*	ns	ns	0.09434	0.0001	0.0001	0.0001	
Etoposide	*	ns	ns	0.08817	0.00618	0.00563	0.00414	
Pyrvinium	**	ns	ns	0.09941	0.0003	0.00579	0.0001	
LY-294002	*	ns	ns	ns	0	0	0	
Thioridazine	*	ns	ns	ns	0	0.00022	0	
Prochlorperazine	*	ns	ns	ns	0	0.00268	0	
Valproic acid	*	ns	0.0001	ns	0.00556	ns	0.0001	

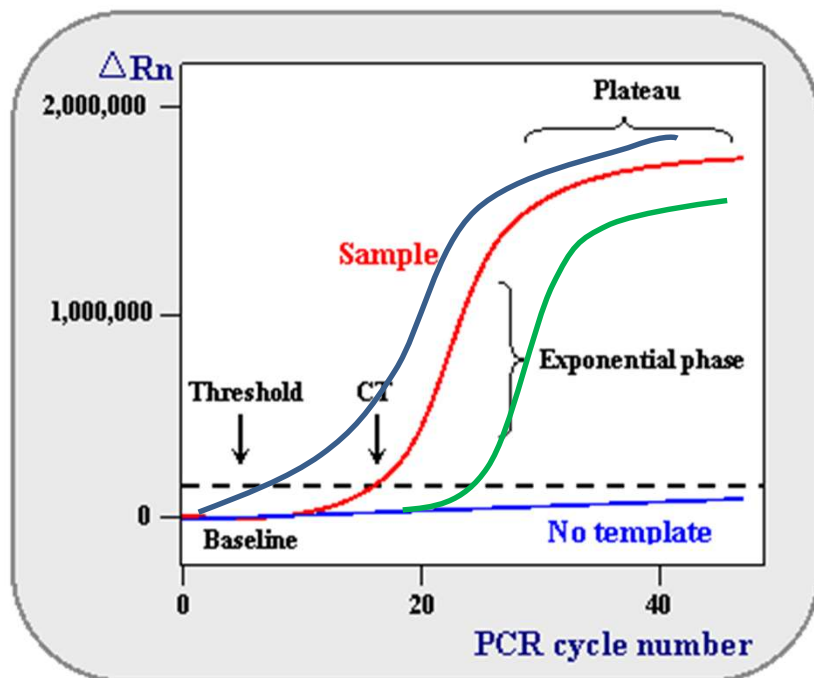
Connectivity Map analysis summary. * = significant specificity (<0.1), ** = highly significant specificity (<0.01).

RNA-sequencing: splicing-alternativo



Quantitative Real-time PCR

Model of real time quantitative PCR plot



Cycle
(40 Cycles)

Baseline

is defined as PCR cycles in which a reporter fluorescent signal is accumulating but is beneath the limits of detection of the instrument.

ΔR_n

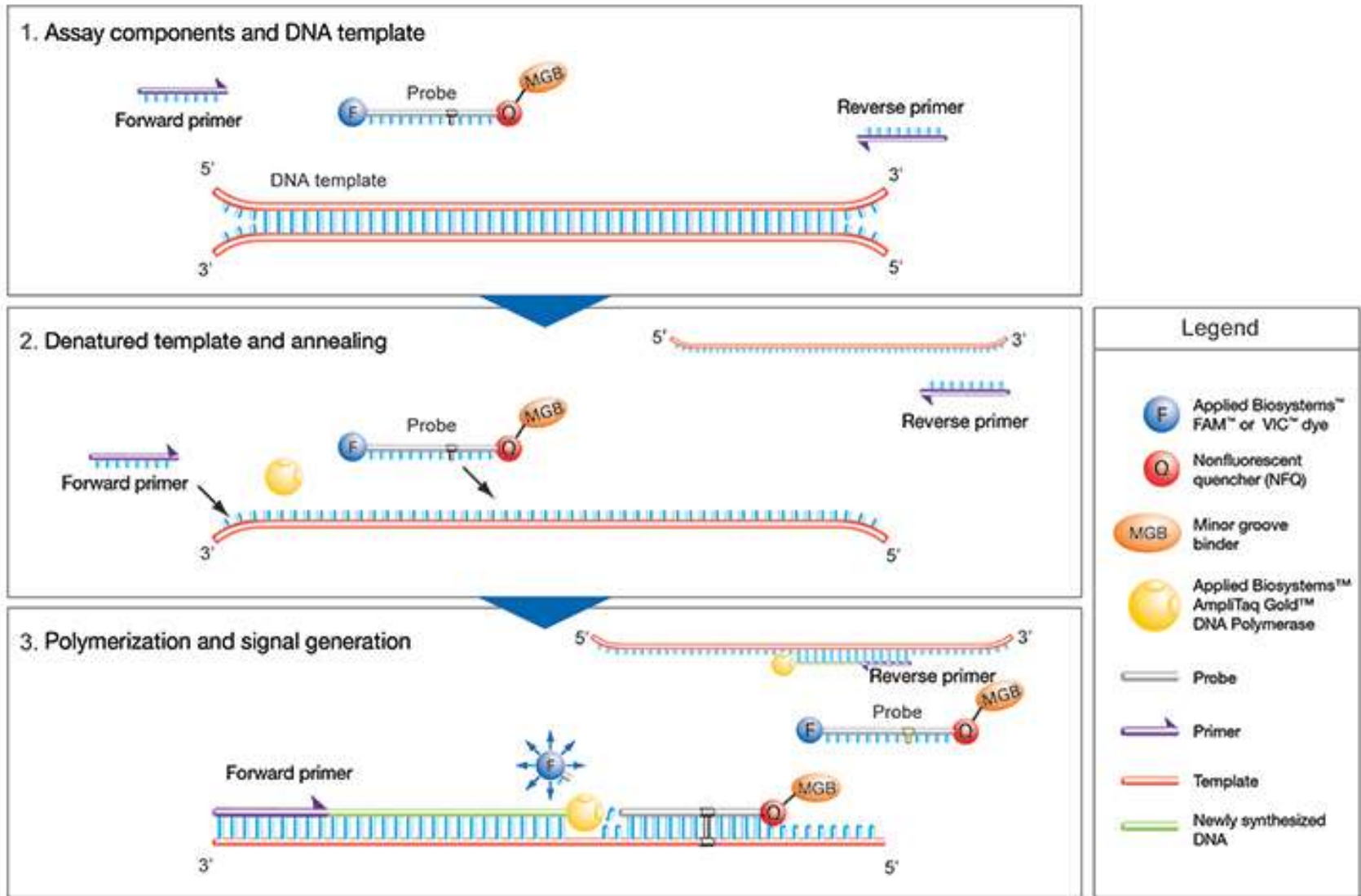
is an increment of fluorescent signal at each time point. The ΔR_n values are plotted versus the cycle number.

Threshold

is an arbitrary level of fluorescence chosen on the basis of the baseline variability. A signal that is detected above the threshold is considered a real signal that can be used to define the threshold cycle (Ct) for a sample. Threshold can be adjusted for each experiment so that it is in the region of exponential amplification across all plots.

Ct is defined as the fractional PCR cycle number at which the reporter fluorescence is greater than the threshold. The Ct is a basic principle of real time PCR and is an essential component in producing accurate and reproducible data.

Gene Expression Analysis Using TaqMan Assays



Quantitative Real-time PCR

SAMPLE	18S Ct	NR3C1 Ct
Patient S1	11,4	23,2
Patient S2	11,3	23,4
Patient S3	10,9	23,5
Patient R1	10,5	28,5
Patient R2	10,4	28,8
Patient R3	10,1	28,4

S= sensitive

R= resistant

The *Livak* method or delta delta CT method

$$\Delta Ct_{PS} = Ct_{target} - Ct_{calibrator} \quad 23,4 - 11,2 = 12,2$$

$$\Delta Ct_{PR} = Ct_{target} - Ct_{calibrator} \quad 28,6 - 10,3 = 18,3$$

$$\Delta \Delta Ct = \Delta Ct_{PR} - \Delta Ct_{PS} \quad 18,3 - 12,2 = 6,1$$

$$\text{Relative expression} = 2^{-(\Delta \Delta Ct)} \quad 2^{-6,1} = 0,015$$



Patients R presented low levels of *NR3C1*