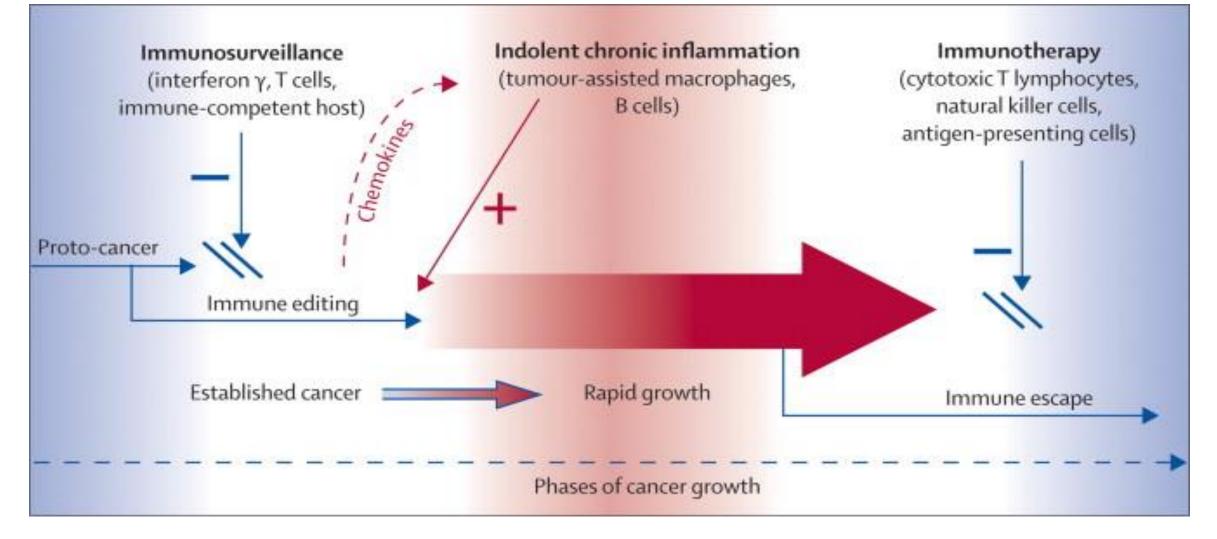
Transcriptome analysis in Immune oncology

Ena Wang

Transcriptome or Gene expression profiling

- Gene expression profiling is the measurement of the expression of thousands of genes at once, to create a global picture of cellular function.
- These profiles can distinguish between cells that are actively dividing, or show how the cells react to a particular treatment.
- Technologies such as RT-PCR (Fluidigm), NanoString, cDNA microarray, RNAseq, can measure an entire genome simultaneously in a given tissue or cell.

Cancer is a Process, not a State

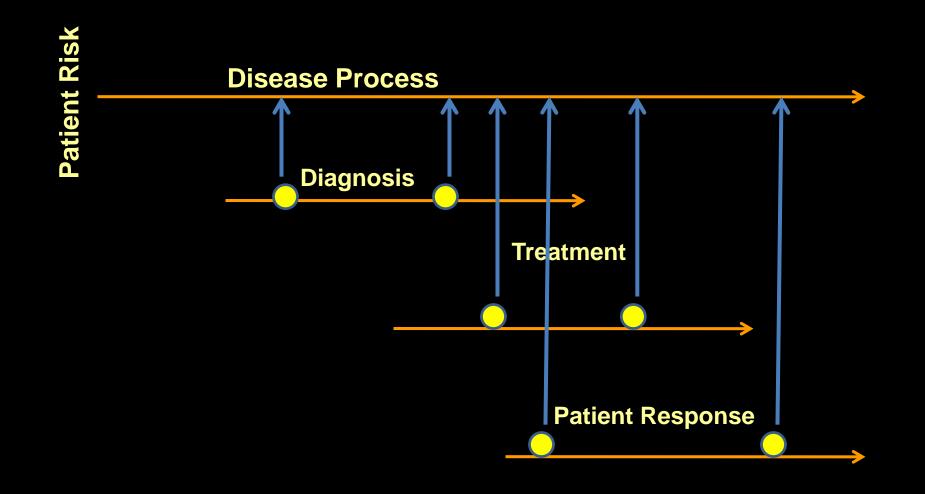


Tumour immunity: effector response to tumour and role of the microenvironment

Alberto Mantovani, Pedro Romero, A Karolina Palucka, Francesco M Marincola

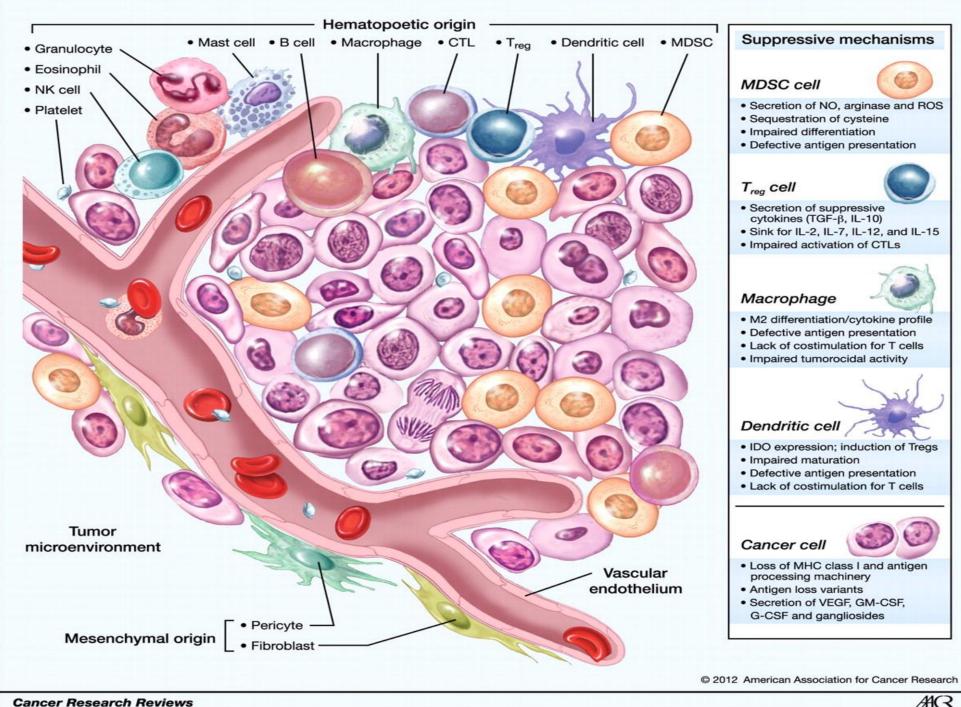
Lancet 2008; 371: 771–83

Disease is a Process, not a State



Time

Liebman M. Strategic Medicine, 2010

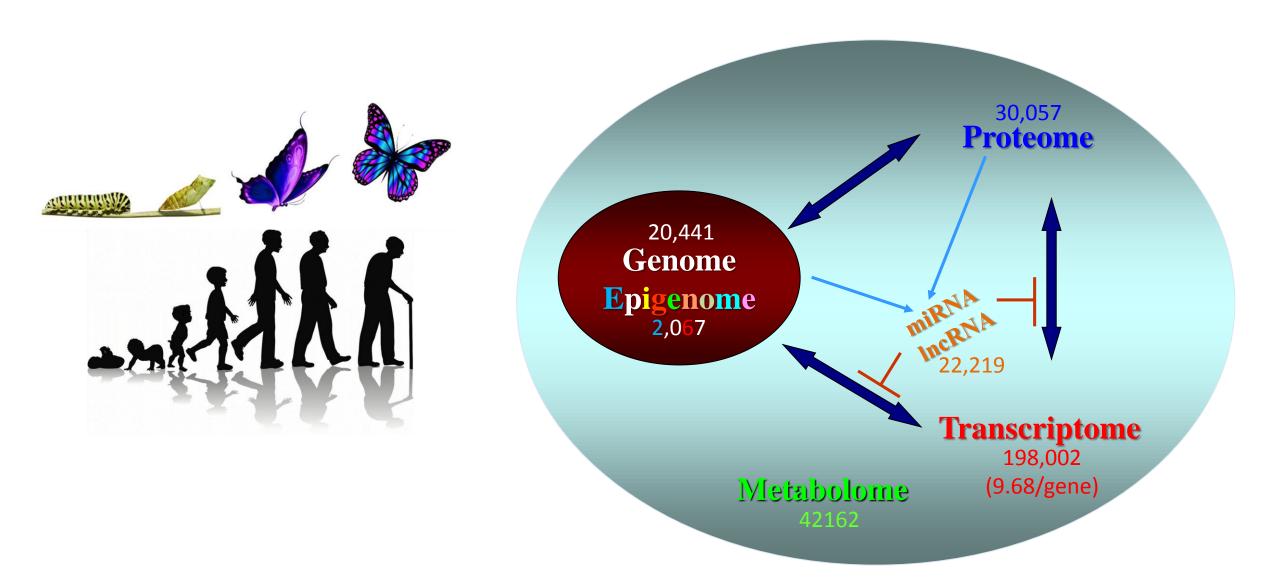


AAR

The immune system is too complicated to proceed in the usual onevariable-at-a-time manner

- Cells type
 - Tumor cells
 - Immune cells
 - Dendritic cells
 - CD8 T cells
 - CD4 T cells
 - Regulatory T cells
 - Stromal cells
- Cytokines and soluble factors

- Cell states
 - Tumor cells
 - MHC expression
 - Neo-antigen repertoire
 - Mutation load
 - Y-IFN pathway status
 - CD8 T cells
 - Location
 - Maturity
 - CD25 signaling
 - CTLA signaling
 - PD1 signaling



Lecture content:

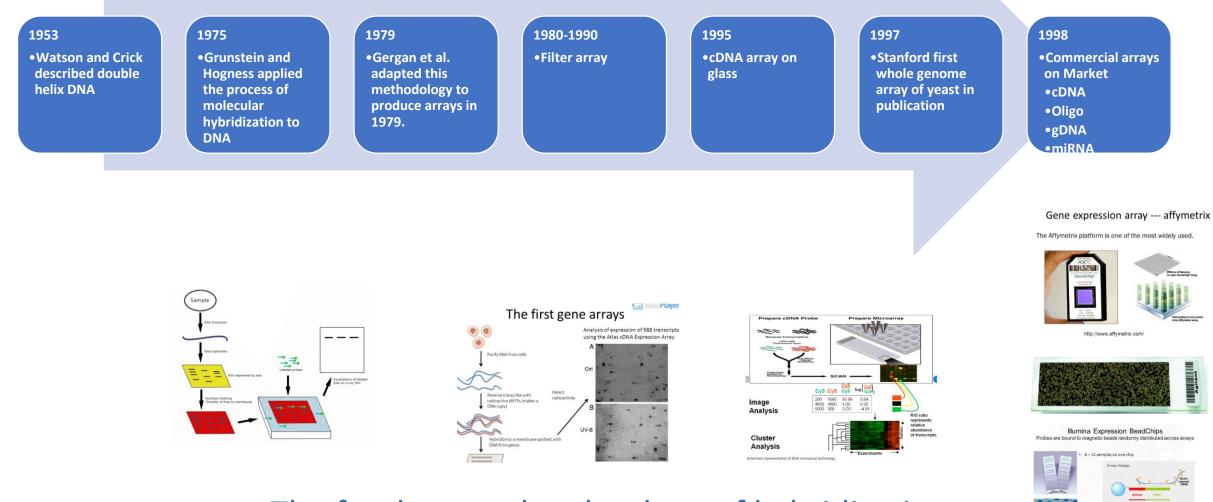
- Technologies for gene expression analysis
- State of art RNAseq and Single Cell gene expression
- Sample collection and handling impact in data quality
- Utility and limitations of gene expression analysis.
- Peripheral vs tissue specific gene expression
- Bulk sample vs single cell vs immune repertoire gene expression
- Challenge in data throughput, analysis and interpretation
- Clinical applications in MOA, prediction of clinical outcome and prediction of treatment response

Lecture content:

Technologies for gene expression analysis

- State of art RNAseq and Single Cell gene expression
- Sample collection and handling impact in data quality
- Utility and limitations of gene expression analysis.
- Peripheral vs tissue specific gene expression
- Bulk sample vs single cell vs immune repertoire gene expression
- Challenge in data throughput, analysis and interpretation
- Clinical applications in MOA, prediction of clinical outcome and prediction of treatment response

History of micro array technology in gene expression analysis



The fundamental technology of hybridization

15 – 30 replicate beads per array target on the averag Most genes are represented by a single probe, some I was probes for different isoforms of the sene

Lecture content:

- Technologies for gene expression analysis
- State of art RNAseq and Single Cell gene expression
- Sample collection and handling impact in data quality
- Utility and limitations of gene expression analysis.
- Peripheral vs tissue specific gene expression
- Bulk sample vs single cell vs immune repertoire gene expression
- Challenge in data throughput, analysis and interpretation
- Clinical applications in MOA, prediction of clinical outcome and prediction of treatment response

Gene expression profiling can:

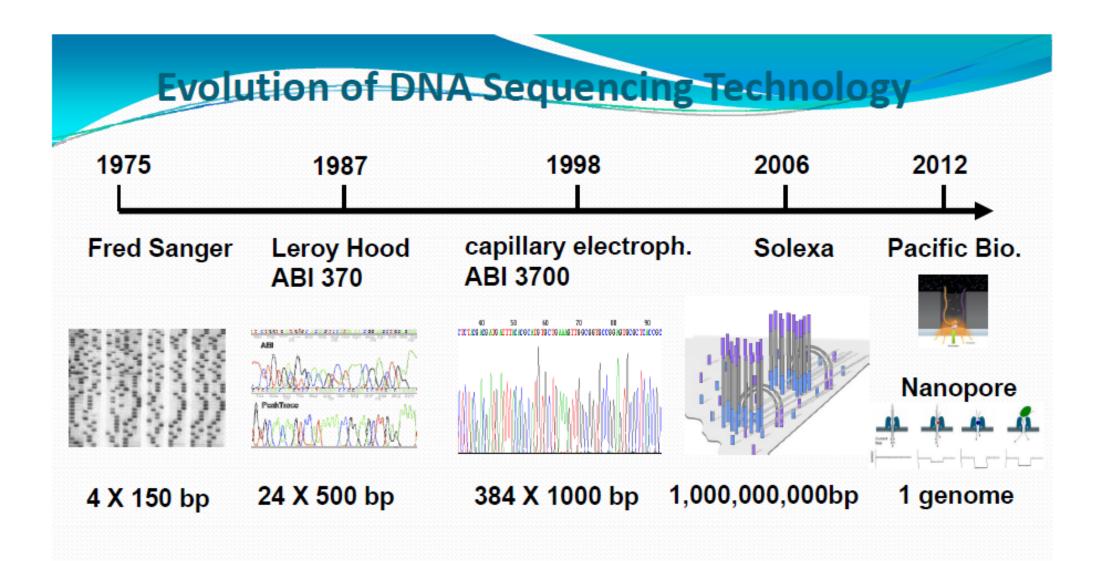
- Identify predictors (pre-treatment samples) of immune or treatment induced responsiveness
- Uncover MOAs of therapeutics (on treatment samples)
- Understand and predict mechanism(s) of adverse events
- Understand the cause of recurrence and the underlying mechanisms adopted by cancer cells to escape from immune recognition (post treatment biopsies)

Next generation sequencing (NGS) NextGen Deepseq

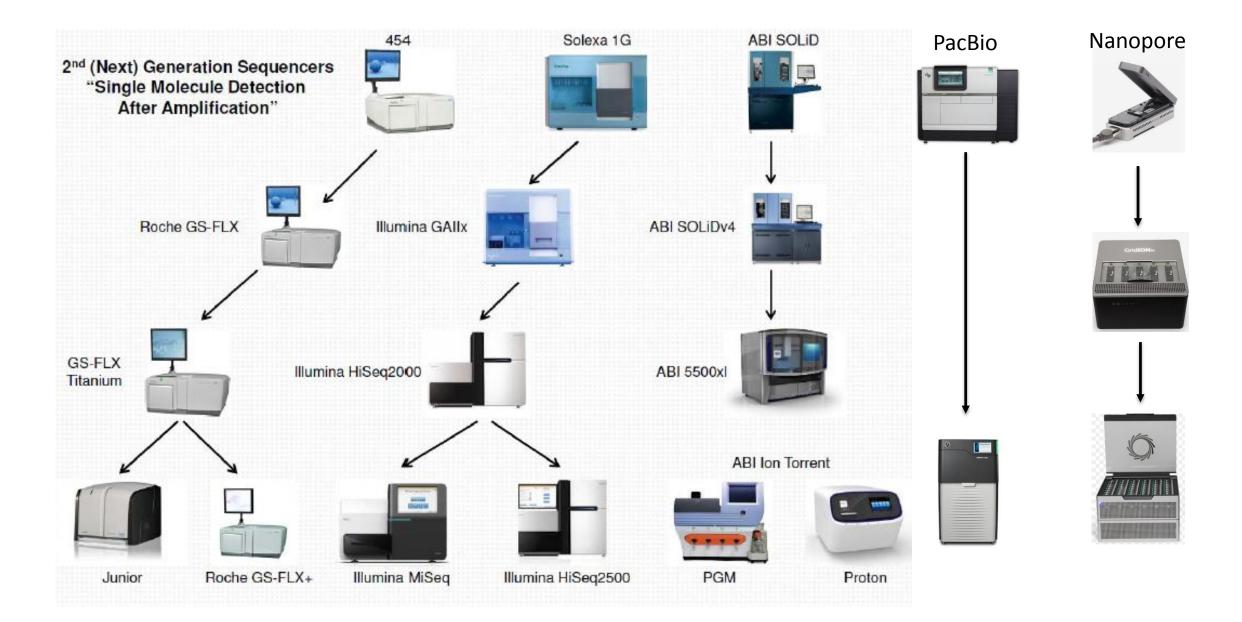
cDNA Array vs RNAseq

cDNA micro array	RNAseq
	Survey the entire Transcriptome, including novel un-annotated regions and non protein coding RNA (miRNA, snRNA, snoRNA,
Profiling mRNA and some ESTs and miRs	piRNA)
	Can determine gene structure by mapping read density across
Profiling the prob region only	transcript or gene
Probe are transcript specific not including splicing	
junction in design	Can use reads to map splice junctions to confirm isoform variants
	Can determine variation at nucleotide level and therefore can detect
Do not detect SNP variants	SNP and indels
Could not differential allelic or direction of	
transcription	Allele specific seq and directional seq
Cost low (reagent and chips)	cost relative high
instrument cost low	Insturment cost high
Data analysis relative traight forward	Need bioinformatics support
Data storage no challenge	much higher throughput
Anotation easy	Anotation challenge when explore beyound expression
	When reads low, lower abundant transcripts could be missed or
Can detect low anoundance transcripts	filtered out.

History of sequencing technology for gene expression analysis



Evolution of NGS platforms



Sequencer comparison

Platforms	Primary advantage	Primary disvantage
Illumina MiniSeq	MiSeq Nano & Micro kits	Limited to PE150; High output PE150 kits are relatively expensive; newness creates unknowns
Illumina MiSeq	Moderate cost instrument and runs; Low cost per Mb for a small platform; Fastest Illumina run times and longest Illumina read lengths	Relatively few reads and Higher cost per Mb compared to NextSEq or HiSeq
Illumina NextSeq 500	Easy to use; Moderate instrument and run costs; Flow cells with room to grow	Version 2 of new chemistry still not as good as older chemistry (used on MiSeq & HiSeq)
Illumina HiSeq 2500	Low cost per MB of data; can run high output (8 lane) and rapid run (2 lane) flow cells with several possible read-length configurations	High instrument cost; high cost per run; requires highly trained personnel; ~20% downtime; can't run one rapid run and one high output flow cell at the same time
Illumina HiSeq 4000	Patterned flow cells; low cost per read & per MB	Similar to HiSeq 2500, but even more expensive; not as flexible as the HiSeq 2500
Illumina HiSeq X Five (or Ten)	Current lowest cost per read or per MB	Requires ≥\$6M plus more data storage than imaginable; only authorized for ≥30x genome resequencing
Illumina NovaSeq	Lowest cost and highest throuput	Need large volume of sample per run. More for service center use
lon Torrent – PGM	Low cost instrument upgraded through disposable chips (the chip is the machine); very simple machine with few moving parts; three chips available with varying numbers of reads	More hands-on time and fewer reads at higher cost per Mb relative to MiSeq; smaller user community; libraries made for 400 base reads don't work well on Proton
lon Torrent – Proton	Moderate cost instrument for medium throughput applications; similar cost to NextSeq, but PII to actually deliver similar reads as NextSeq is years overdue	More hands on time, and fewer overall bases of data than Illumina; Higher cost per MB of data than most Illumina instruments; smaller user community
Oxford Nanopore minION	SMALL PORTABLE Instrument; It is a USB device; extremely low-cost instrument; extremely long reads feasible (multiple kb)	Weird quasi-commercial release; error-rates only now approaching what was promised in 2012; biased errors; High cost per read
Oxford Nanopore PromethION * (forecast)	Should be similar to an array of MinIONs to increase throughput & decrease costs	Similar to or same as MinION; Not yet available
PacBio – RS II	Single molecule real-time sequencing; Longest available read length among well established platforms; Ability to detect base modifications; Short instrument run time; Random error profile; Modest cost per sample Excellent for full genome assembly with sufficient coverage	High error rates so high coverage is necessary; Low total number of reads per run; High cost per Mb; High capital cost; Many methods still in development; past company performance not very good (though markedly improved recently)
PacBio – Sequel	Lower cost for instrument, data & service contract relative to RS II; otherwise should be similar to RSII	New, thus limited info from non-company sources; High cost per Mb and per read relative to other platforms; should be similar to RS II

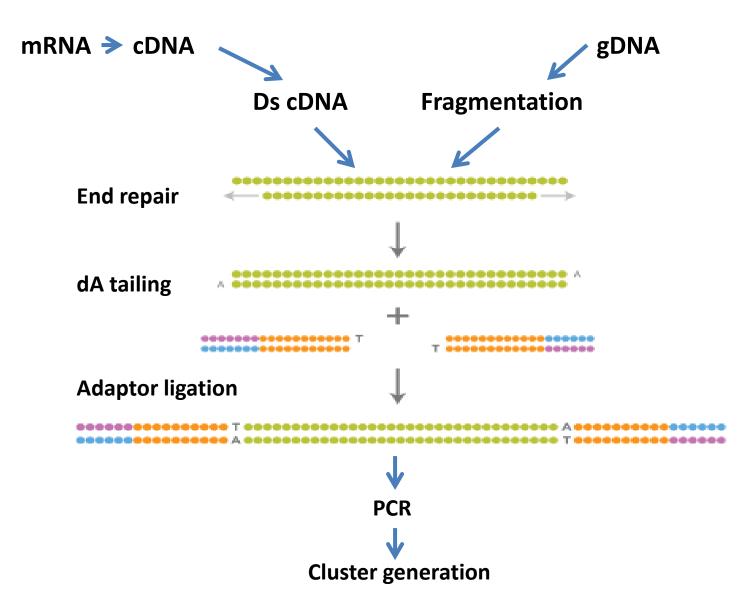
Illumina-GAII, Hiseq, Miseq, NextSeq, NovaSeq

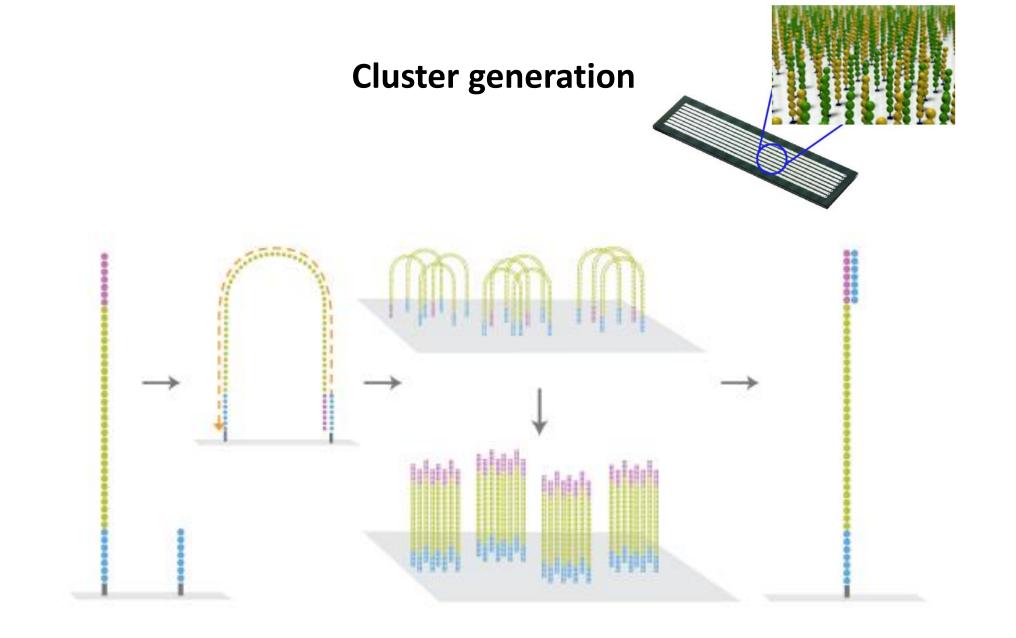
library+cluster generation

CRT (cyclic reversible termination)

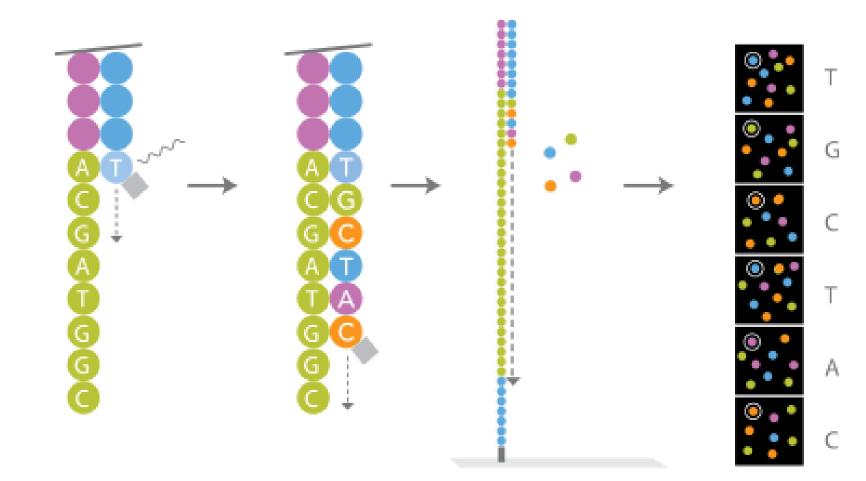
Image capture

Library preparation





Cyclic reversible terminator sequencing and image capture



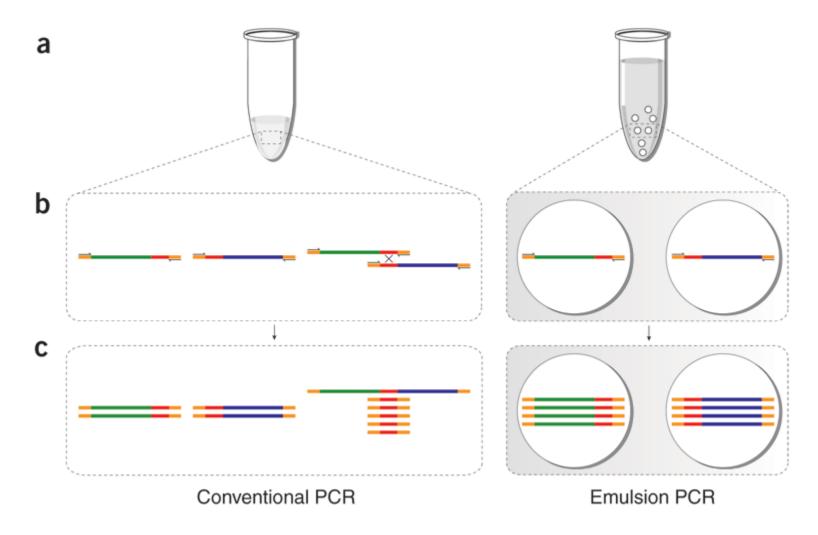
ABI-Ion Torrent, Ion Proton

ePCR

Pyrosequencing

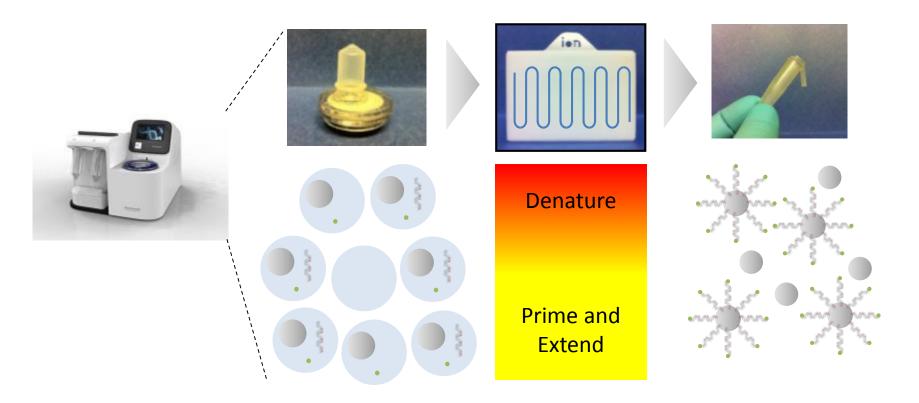
Semiconductor

Library generation by Emulsion PCR



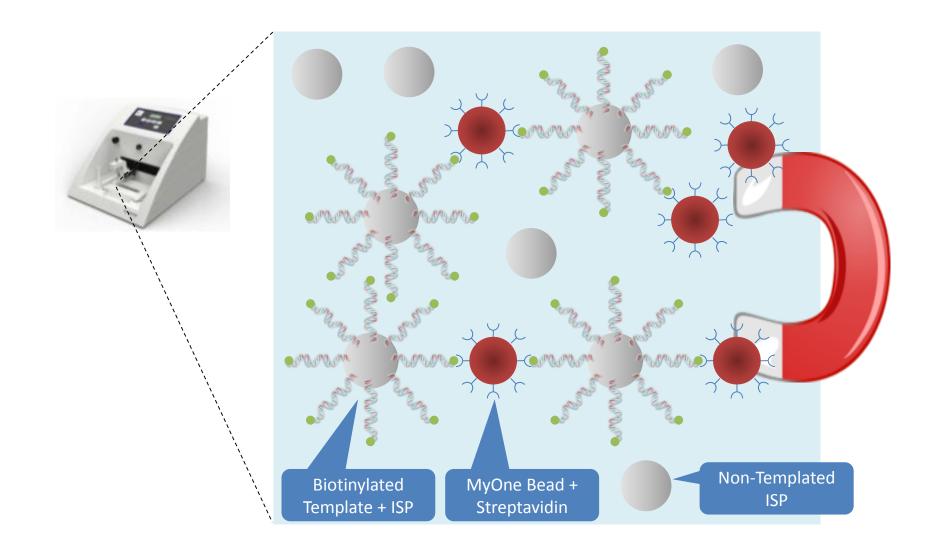
R Williams, et al. Nature methods

Revolutionary In-Line PCR Technology

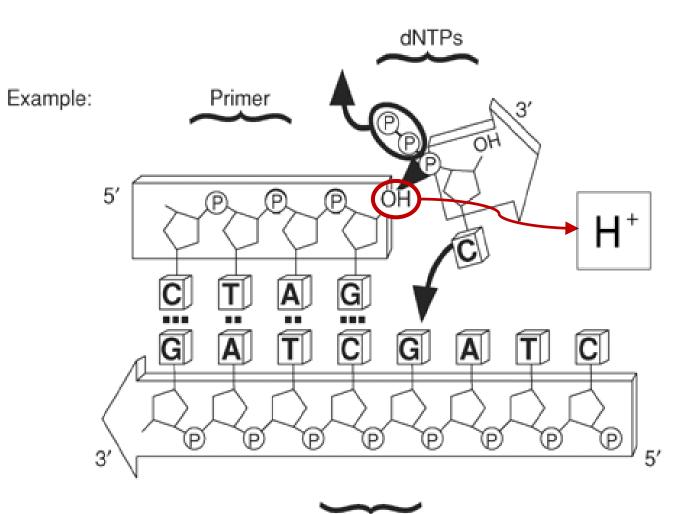


Ion Sphere Particles with 5' primer Template PCR cycle Biotinilate 3' primer

Automated template Enrichment

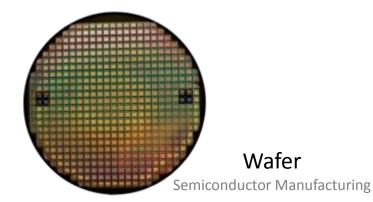


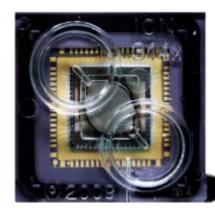
Same chemistry but no enzymes involved in signal generation



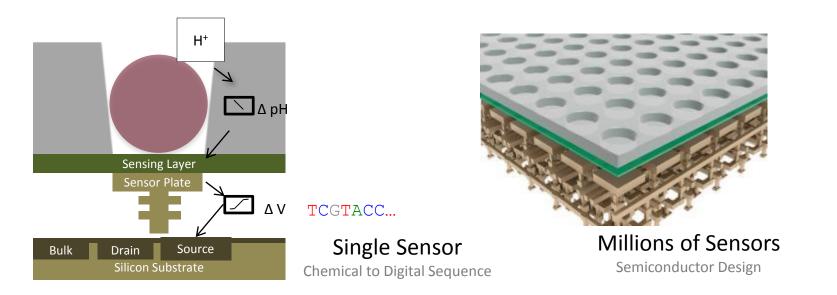
Template

Transformative Technology, Unprecedented Progress Scalable Semiconductor Technology plus Simple Chemistry



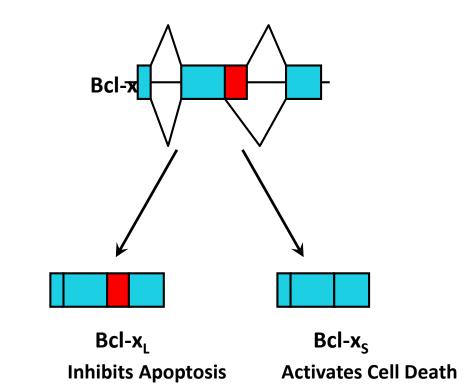


Chip Semiconductor Packaging



Importance of alternative splicing

- Approximately 25,000 protein coding genes in the human genome, but there are >100,000s of transcripts
- ~75% of human genes are thought to be involved in alternative splicing
- ENCyclopedia Of DNA Elements (ENCODE) regions contain an average of 5.4 alternative transcripts/locus
- 20 30% of disease causing mutations are believed to affect premRNA splicing



Deregulation of apoptosis contributes to the development and progression of Cancer

Disease modification and susceptibility mediated through splicing events

Splice event linked with disease modification

Splice event linked with disease susceptibility

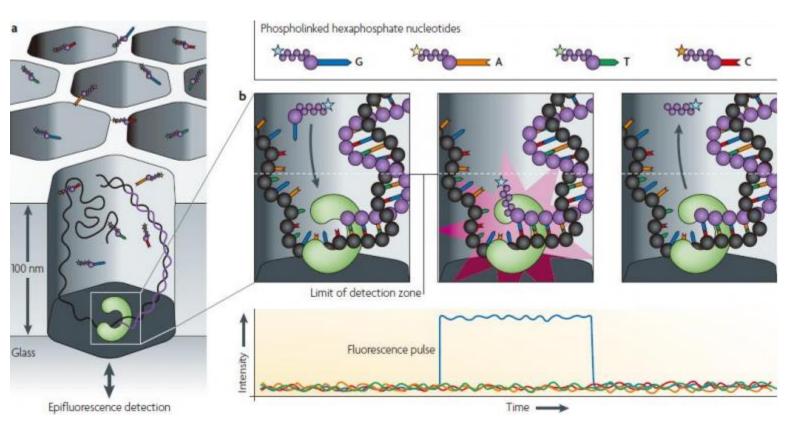
•CFTR	Cystic fibrosis
•MCAD	Medium-chain acyl-CoA
	dehydrogenase deficiency
•SCN1A	Susceptibility to anti-epileptics
•IKBKAP	Familial dysautonomia
•Scn8a	Mouse Neurological disorder

•IRF5	Systemic lupus erythematosus
•CTLA4	Autoimmune diseases
•NCAM1	Bipolar disorder
•ERBB4	Schizophrenia
•OLR1	Myocardial infarction
•OAS1	Type I diabetes
•TNNT2	Cardiac hypertrophy
•GPRA	Asthma
•MAPT	Tauopathies
•PTPRC	Altered immune function &
	Multiple sclerosis
•LDLR	Elevated cholesterol
•SFRS8	Asthma

Wang and Cooper, Nature Review Genetics (2007)



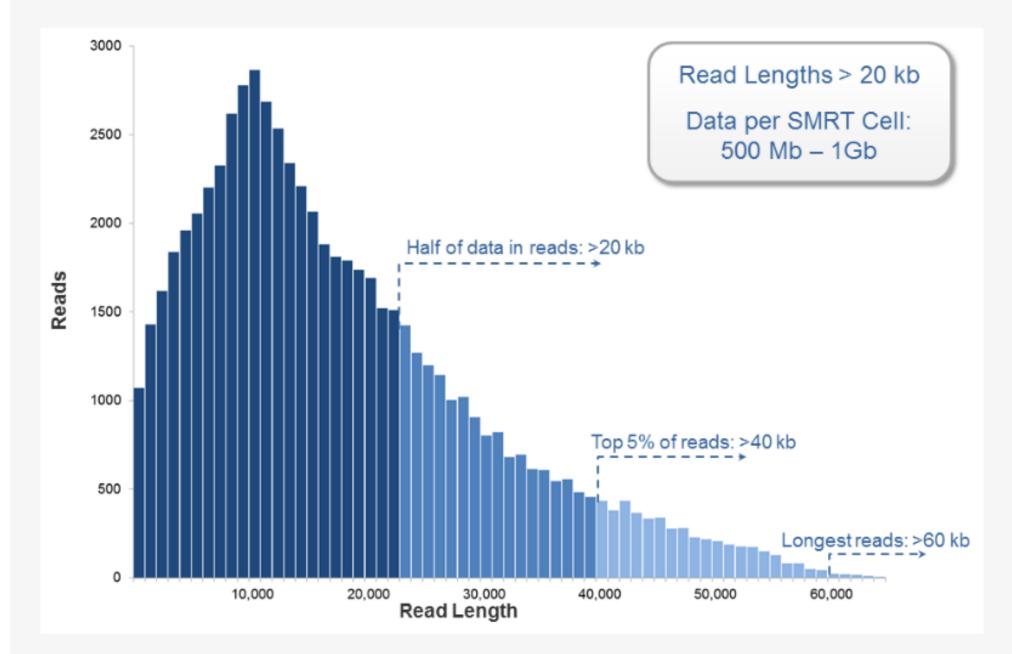
Third Generation Sequencers



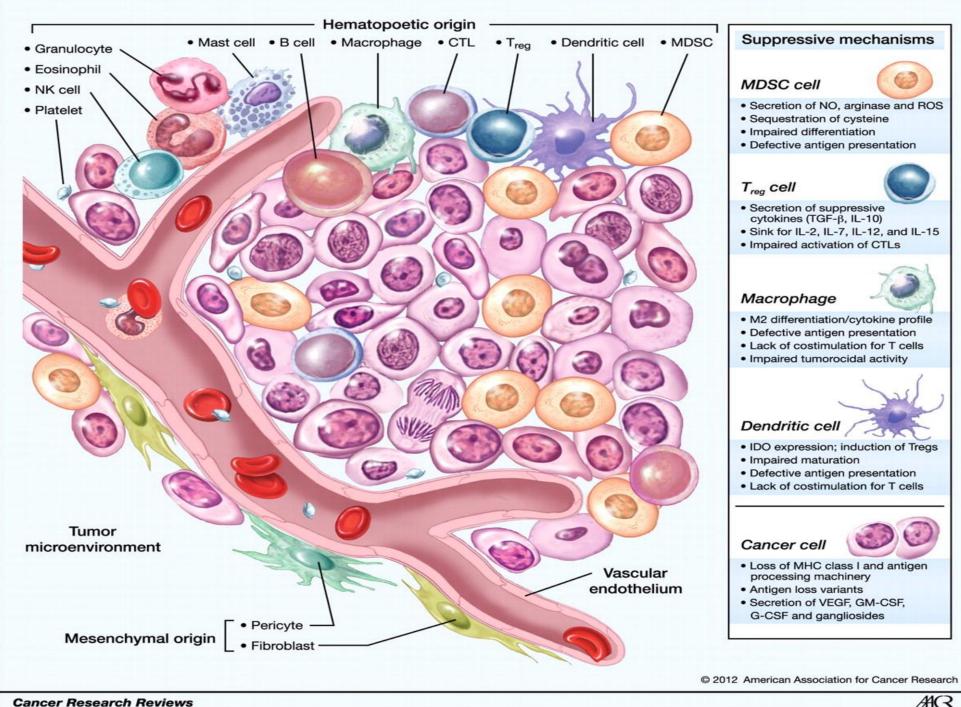
10bp / second

Directly sequence RNA without the need to RT to cDNA

Methylation of the DNA sequence can be determined due to the delayed signal detection



Based on data from a 20 kb size-selected human library using 4-hour movies. Average read lengths for data set shown are 18 kb. Each SMRT Cell generates ~55,000 reads.

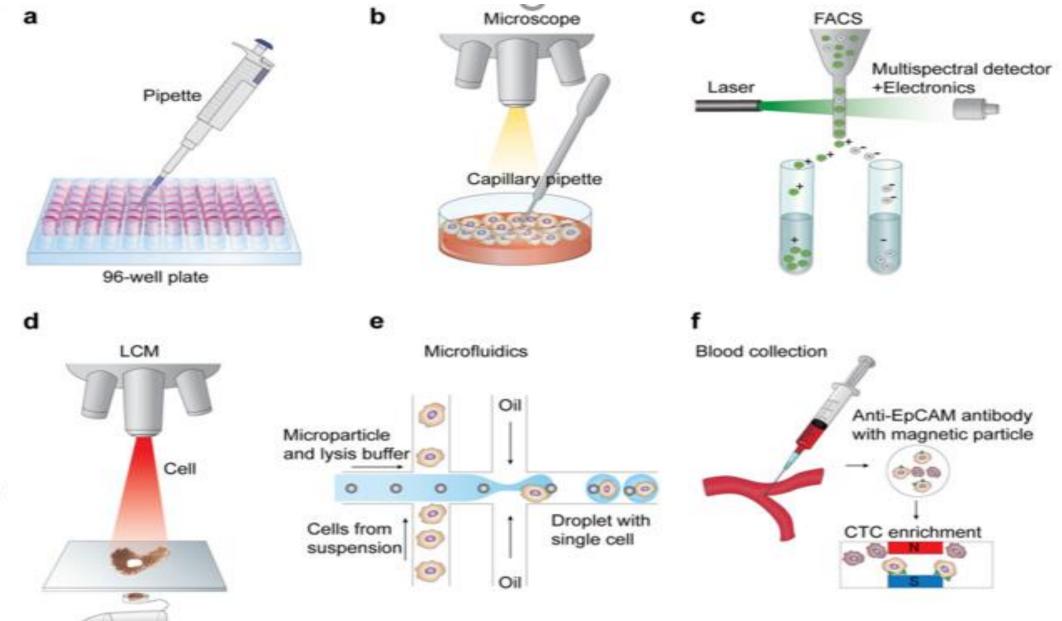


AAR

Lecture content:

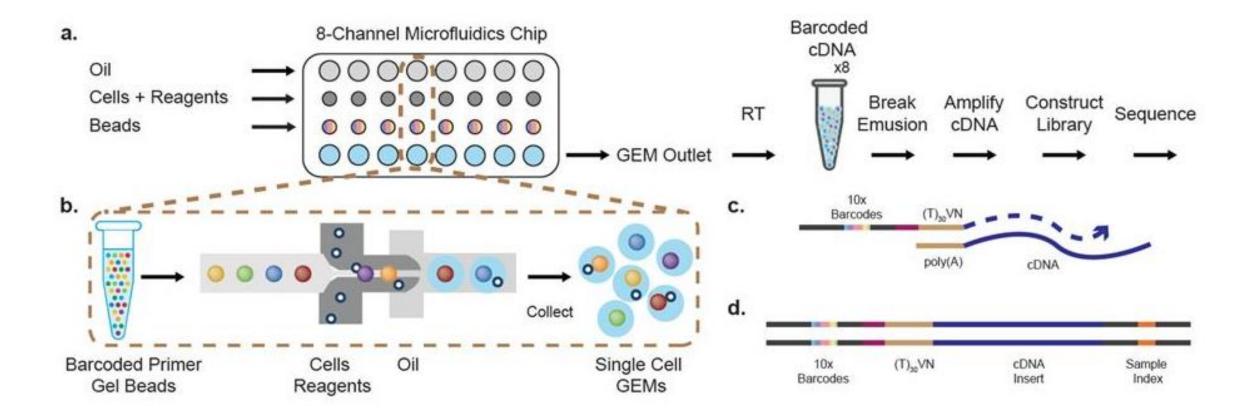
- Technologies for gene expression analysis
- State of the art RNAseq and Single Cell gene expression analysis
- Sample collection and handling impact in data quality
- Utility and limitations of gene expression analysis.
- Peripheral vs tissue specific gene expression
- Bulk sample vs single cell vs immune repertoire gene expression
- Challenge in data throughput, analysis and interpretation
- Clinical applications in MOA, prediction of clinical outcome and prediction of treatment response

Single-cell isolation techniques



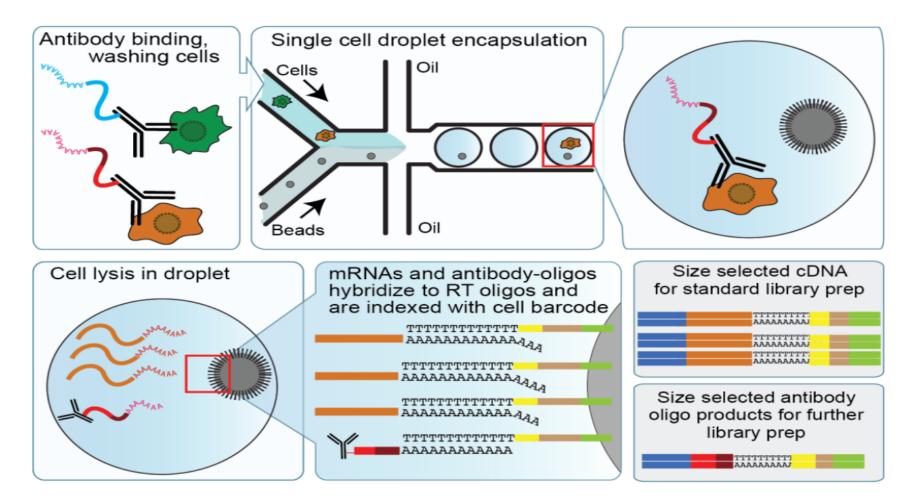
Hwang et al. Experimental & Molecular Medicine (2018) 50:96

Single cell RNAseq – 10X genomics





CITE-seq uses DNA-barcoded antibodies to convert detection of proteins into a quantitative, sequenceable readout. Antibody-bound oligos act as synthetic transcripts that are captured during most large-scale oligodT-based scRNA-seq library preparation protocols (e.g. 10x Genomics, Drop-seq, ddSeq).



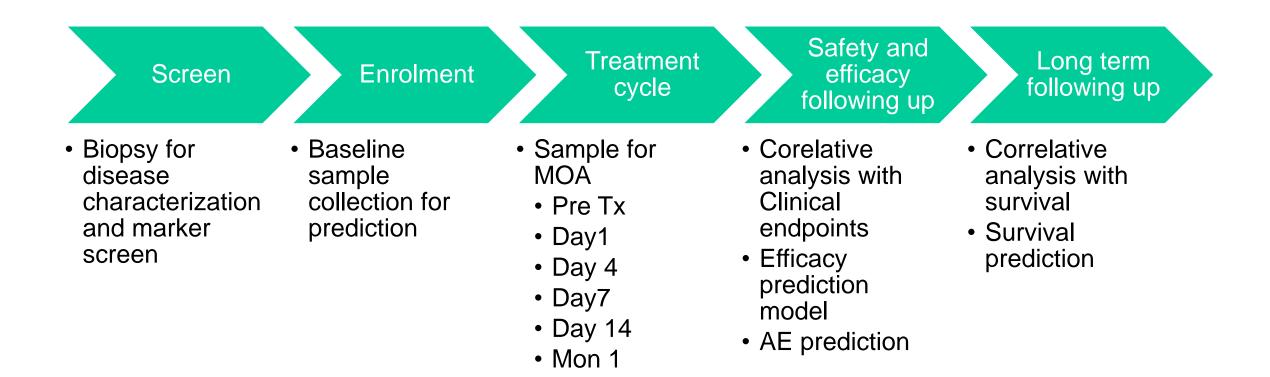
Lecture content:

- Technologies for gene expression analysis
- State of art RNAseq and Single Cell gene expression
- Sample collection and handling impact in data quality
- Utility and limitations of gene expression analysis.
- Peripheral vs tissue specific gene expression
- Bulk sample vs single cell vs immune repertoire gene expression
- Challenge in data throughput, analysis and interpretation
- Clinical applications in MOA, prediction of clinical outcome and prediction of treatment response

Clinical samples

- Collections relevant to disease process
- •Collections at relevant time points
- Proper controls
- Sufficient number of samples
- Storage and handling
- •High quality and sufficient quantity

Schematic of clinical studies



Sample collection for gene expression analysis

- Tempus or PaxGene Blood RNA Tube
- CPT tube for PBMC isolation and cryopreservation
- RNAlater RNA stabilization for tissues biopsy
- Fresh tumor digests

- OCT embedded samples
- FFPE

Lecture content:

- Technologies for gene expression analysis
- State of art RNAseq and Single Cell gene expression
- Sample collection and handling impact in data quality
- Utility and limitations of gene expression analysis.
- Peripheral vs tissue specific gene expression
- Bulk sample vs single cell vs immune repertoire gene expression
- Challenge in data throughput, analysis and interpretation
- Clinical applications in MOA, prediction of clinical outcome and prediction of treatment response

Lecture content:

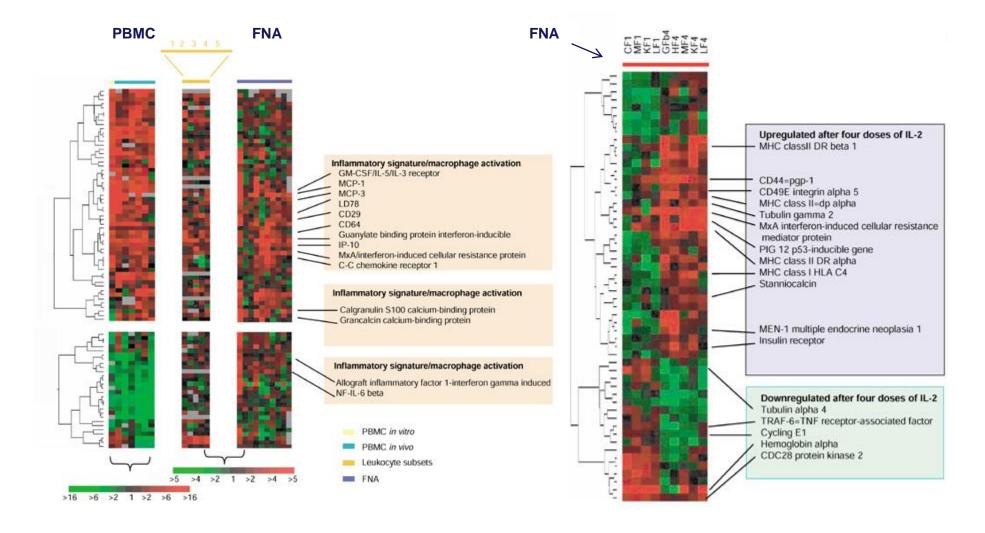
- Technologies for gene expression analysis
- State of art RNAseq and Single Cell gene expression
- Sample collection and handling impact in data quality
- Utility and limitations of gene expression analysis.
- Peripheral vs tissue specific gene expression
- Bulk sample vs single cell vs immune repertoire gene expression
- Challenge in data throughput, analysis and interpretation
- Clinical applications in MOA, prediction of clinical outcome and prediction of treatment response

Research

Gene-expression profiling of the response of peripheral blood mononuclear cells and melanoma metastases to systemic IL-2 administration

Published: 25 June 2002 Genome Biology 2002, 3(7):research0035.1-0035.17

Monica C Panelli^{*}, Ena Wang^{*}, Giao Phan[†], Markus Puhlmann[†], Lance Miller[†], Galen A Ohnmacht[†], Harvey G Klein^{*} and Francesco M Marincola^{*}



Lecture content:

- Technologies for gene expression analysis
- State of art RNAseq and Single Cell gene expression
- Sample collection and handling impact in data quality
- Utility and limitations of gene expression analysis.
- Peripheral vs tissue specific gene expression
- Bulk sample vs single cell vs immune repertoire gene expression
- Challenge in data throughput, analysis and interpretation
- Clinical applications in MOA, prediction of clinical outcome and prediction of treatment response

Where to star?



Hurdles – Choice of analytical software

- Valvet
- BFAST NGS aligner tools
- BLAT
- BWT

•

•

Bowtie No one has all the needed features!

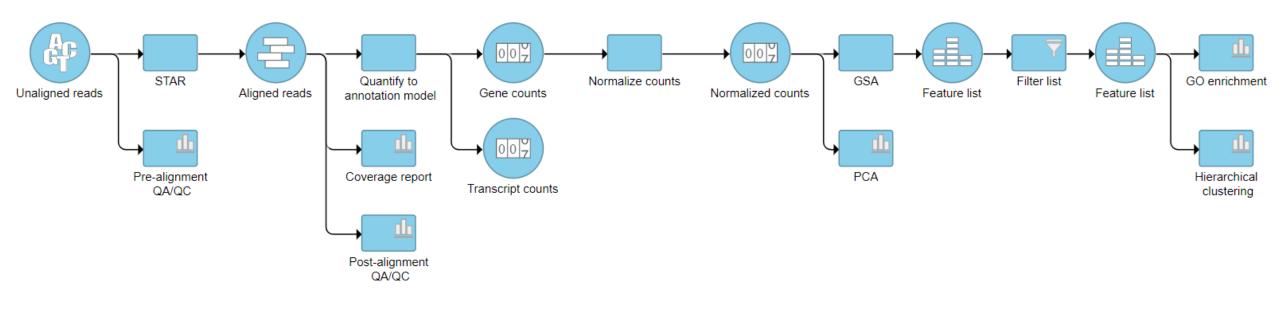
- Tophat
- CLC Bio
- ELAND (CASAVA)
- GMAP
 - MAQ BaseSpace-Illumina Miseq, Hiseq
- NOVAALIGN Genome Modeling tools-http://genomie.
- SOAP....

Wustl.edu

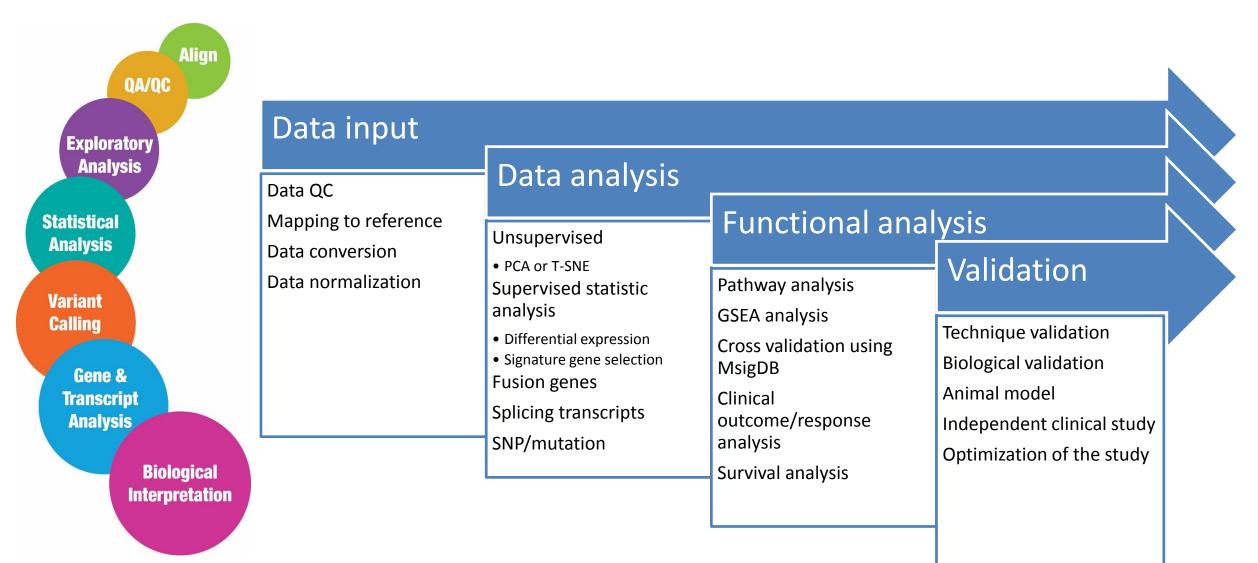
Data analysis and visualization tools

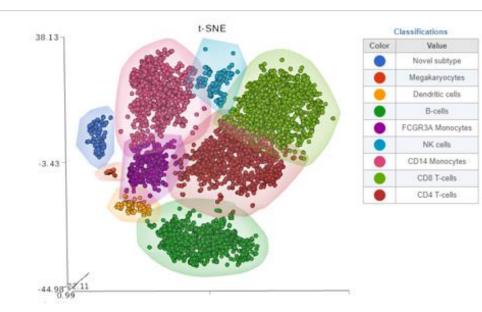
• Partek Genomic Suite, Partek Flow

Point and Click to Build Analysis Pipelines

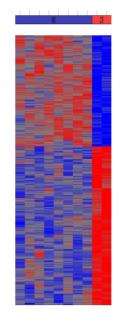


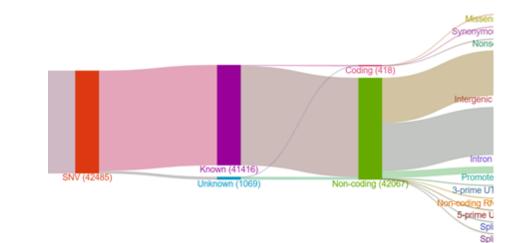
Gene expression analysis

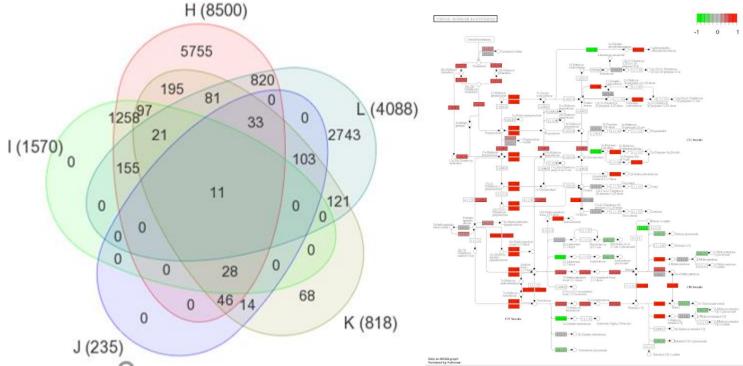


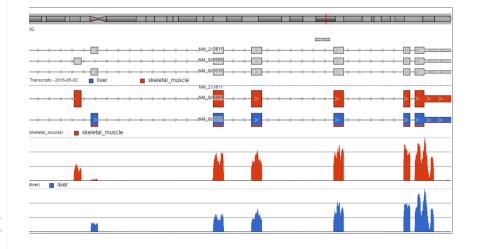


ressed (









Understanding the data

"We are close to having a \$1,000 genome sequence, but this may be accompanied by a \$1,000,000 interpretation." -Bruce Korf, president American College of Medical Genetics

Not only is the cost of sequencing essentially free, but big computers and big storage are cheap, too. what will keep us busy for the next 50 years is understanding the data"

-Russ Altman, chair of Biomedical Engineering at Stanford

Lecture content:

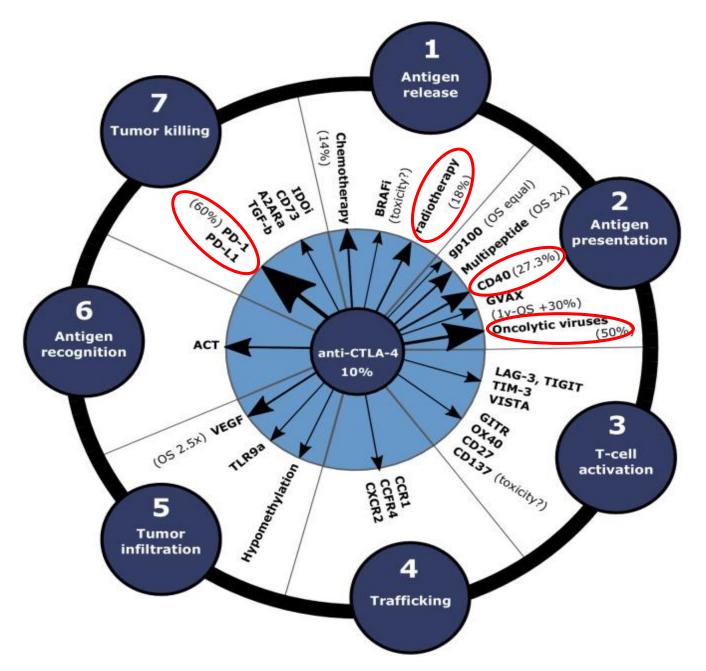
- Technologies for gene expression analysis
- State of art RNAseq and Single Cell gene expression
- Sample collection and handling impact in data quality
- Utility and limitations of gene expression analysis.
- Peripheral vs tissue specific gene expression
- Bulk sample vs single cell vs immune repertoire gene expression
- Challenge in data throughput, analysis and interpretation
- Clinical applications in MOA, prediction of clinical outcome and prediction of treatment response

Why some patient benefit from the treatment while the others do not?

Mechanism of Immunotherapy

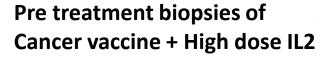
- Cancer vaccine: Directing the immune system to cancer cell by flagging cancer cells for destruction by the immune system
- Cytokine therapy: Increase the killing power of the immune system by enhancing T cell functions
- Oncolytic viro-therapy: Enhance the killing power of the immune system by introducing cancer tropical virus to kill cancer.
- Checkpoint blockade: Releasing the brakes on the immune system
- Adoptive T cell therapy and CAR-T cell therapy: Boosting the killing power of the immune system

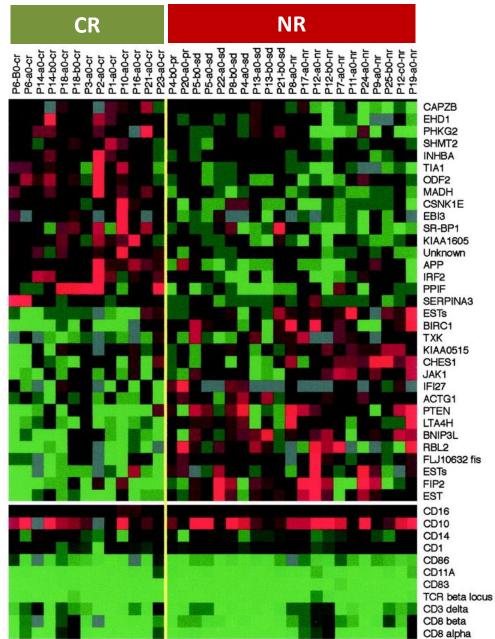
Anti PD1 or PD-L1 combine with CTLA-4 treatment in melanoma



Front Oncol. 2016; 6: 233.

Prediction of response to anti-cancer immunotherapy using gene signatures





Wang et al. Cancer Res. 2002

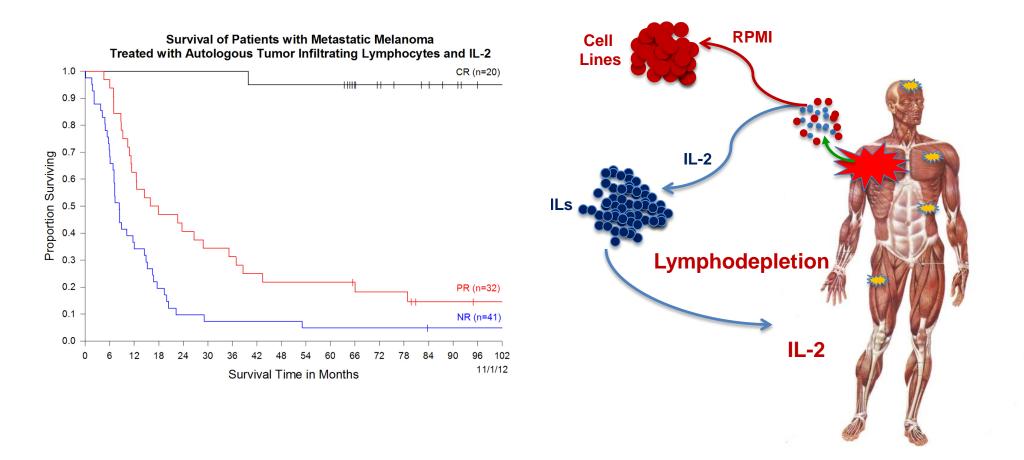
Project overview

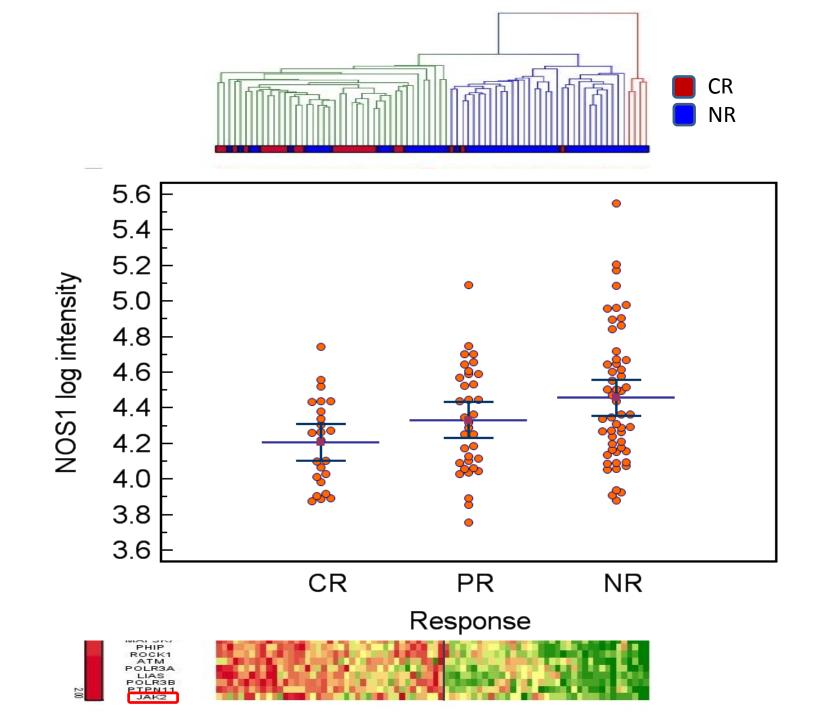
SAMPLES STUDIED

142 TILs from patients enrolled in five adoptive cell therapy trials

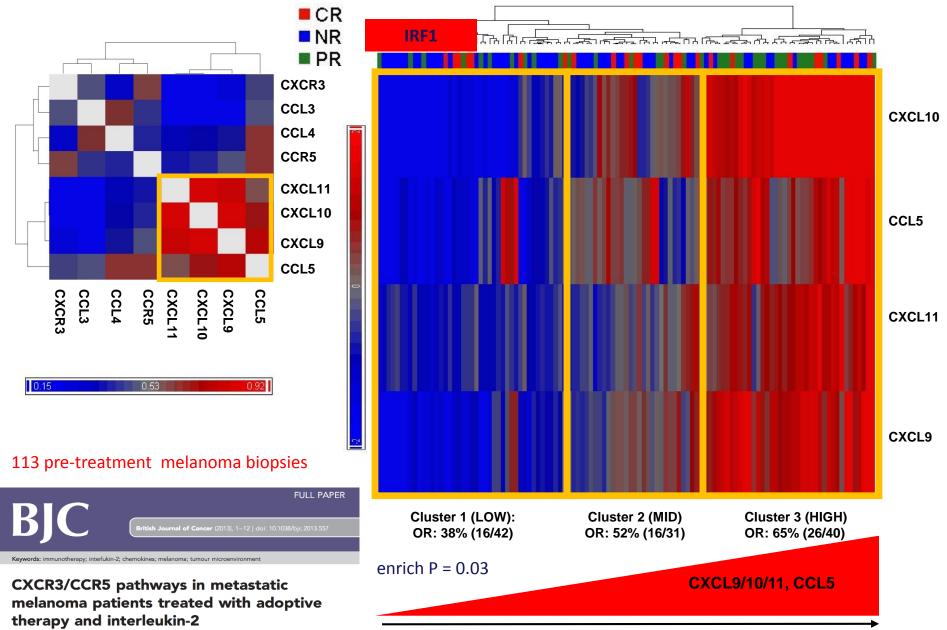
113 parental melanoma metastases

15 melanoma cell lines derived from the 15 melanoma metastases



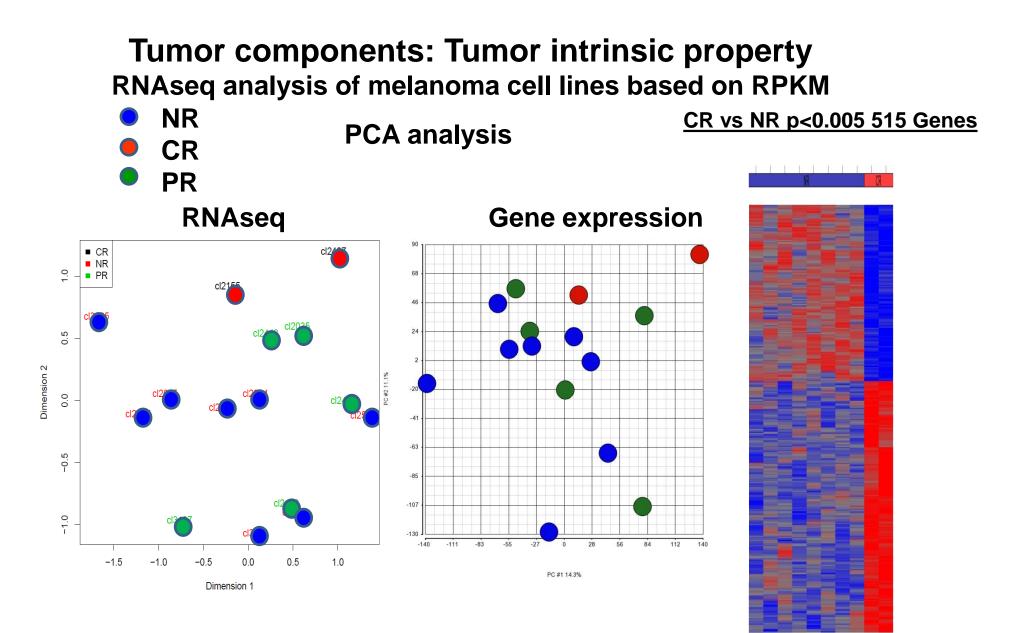


Adoptive therapy + IL-2 (113 pre-treatment melanoma lesions)



D Bedognetti^{*,1,2,3}, T L Spivey^{1,4,5}, Y Zhao⁶, L Uccellini^{1,7}, S Tomei¹, M E Dudley⁸, M L Ascierto^{1,3,9}, V De Giorgi¹, Q Liu¹, L G Delogu¹⁰, M Sommariva^{1,11}, M R Sertol^{2,3}, R Simon⁶, E Wang¹, S A Rosenberg⁸ and F M Marincola^{*,1,2}

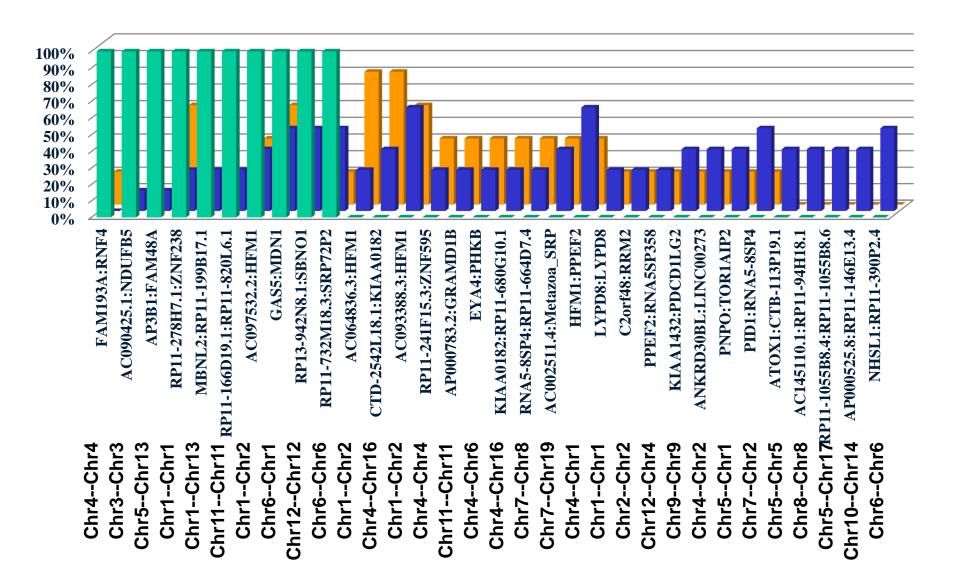
OR Rate: Cluster 1 < Cluster 2 < Cluster 3





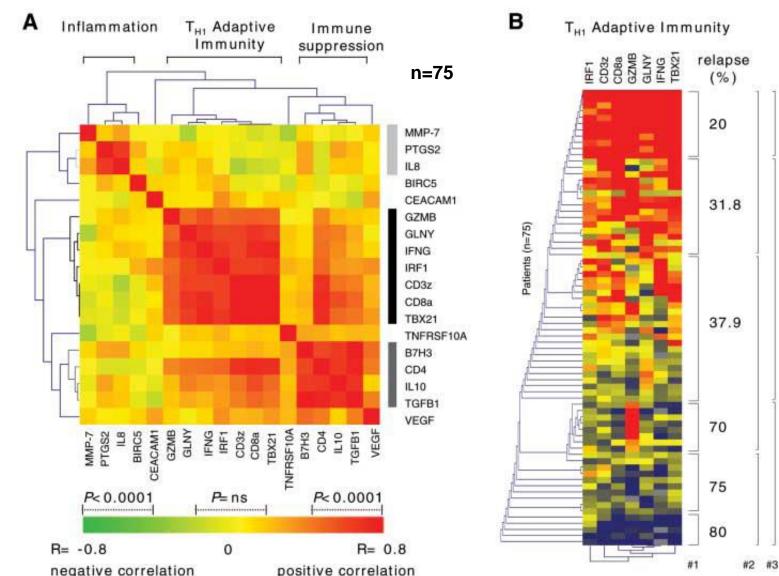
Fusion trascripts identified by RNAseq

■ CR (n=2) ■ NR (n=8) ■ PR (n=5)



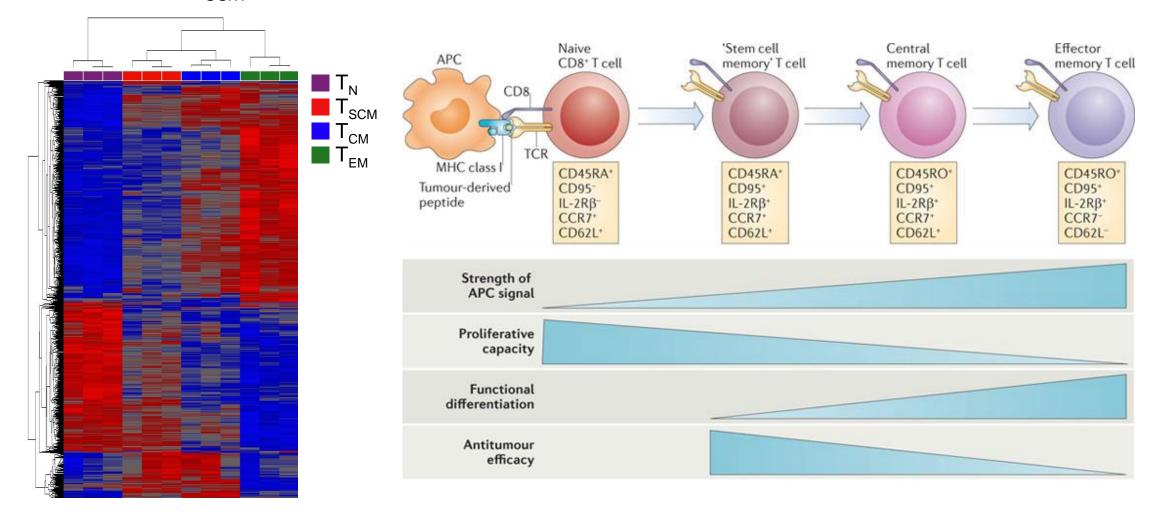
Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome

|érôme Galon,¹*† Anne Costes,¹ Fatima Sanchez-Cabo,² Amos Kirilovsky,¹ Bernhard Mlecnik,² Christine Lagorce-Pagès,³ Marie Tosolini,¹ Matthieu Camus,³ Anne Berger,⁴ Philippe Wind,⁴ Franck Zinzindohoué,⁵ Patrick Bruneval,⁶ Paul-Henri Cugnenc,⁵ Zlatko Trajanoski,² Nolt-Herman Fridman,^{1,7} Franck Pagès^{1,7}†



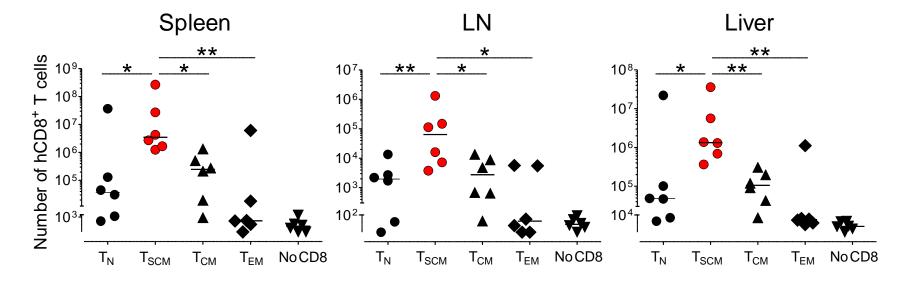
Immune components: Immune cell differentiation status

T_{SCM} are a distinct early T cell memory subset

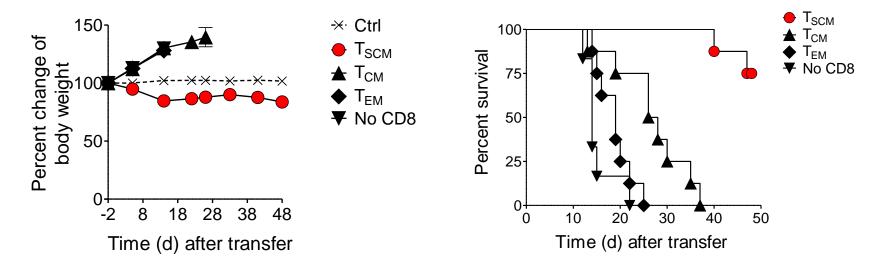


Luca Gattinoni, Nature Med 2011

T_{SCM} have enhanced proliferative and survival capacities



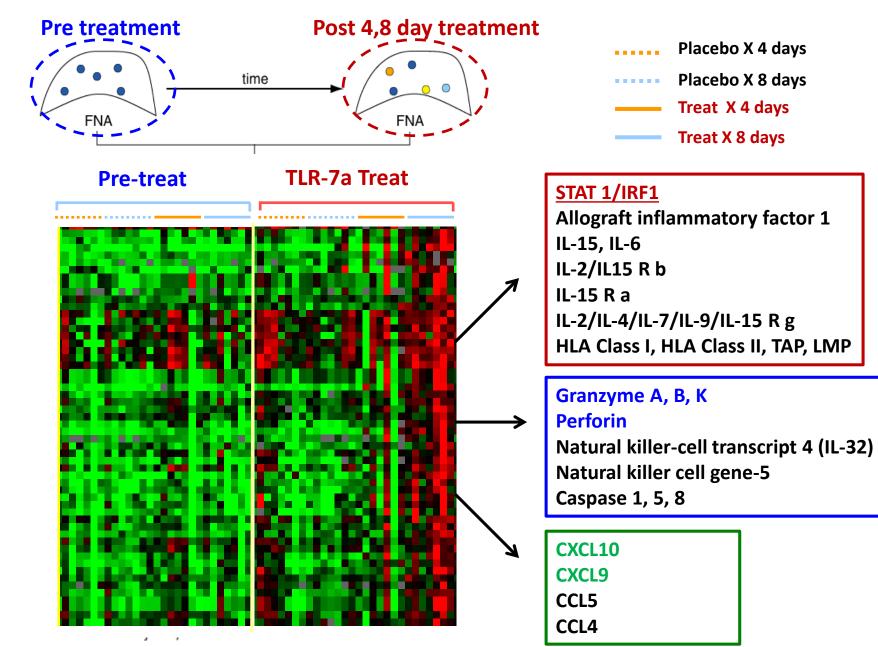
 T_{SCM} have increased anti-tumor activities upon adoptive transfer



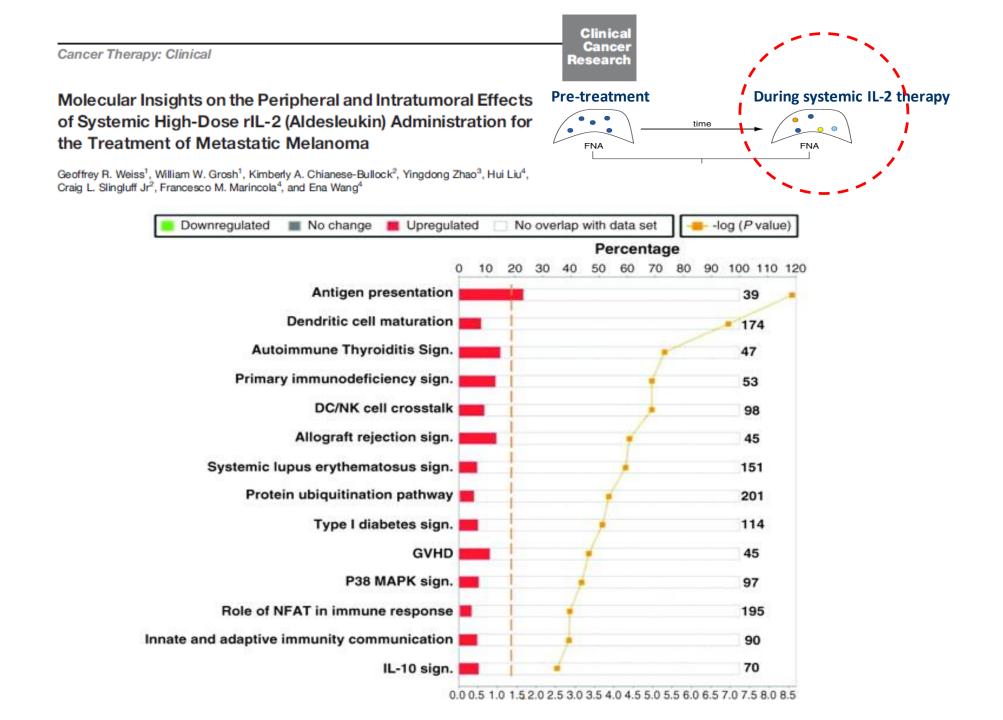
Luca Gattinoni, Nature Med 2011

Mechanism of action: pre vs post treatment

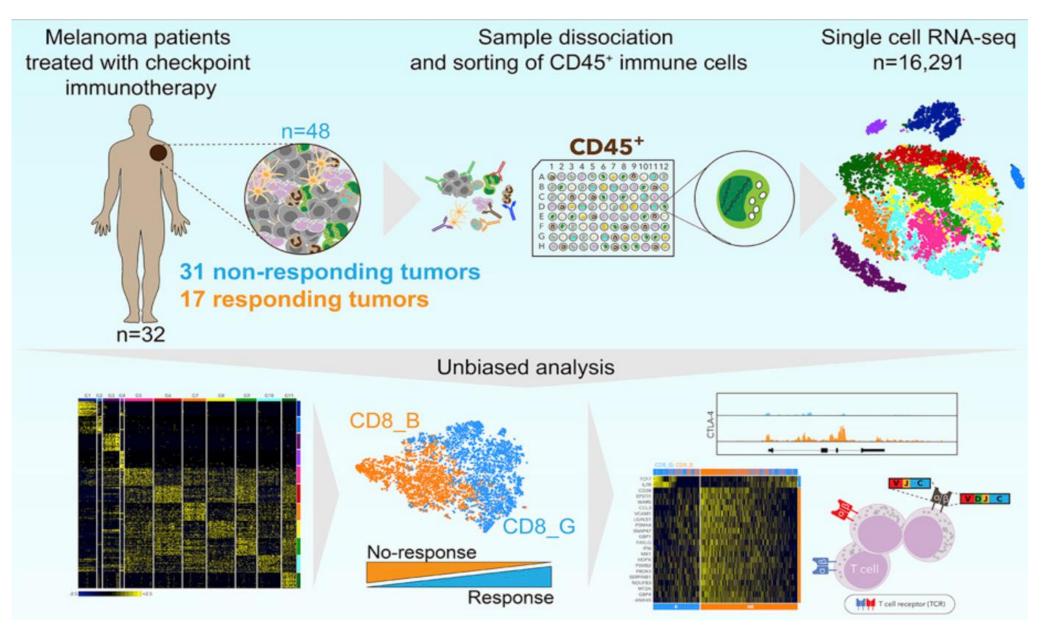
Imiquimod (TLR-7a)-Basal cell Carcinoma



Genome Biology 2007, 8:R8 (doi:10.1186/gb-2007-8-1-r8)

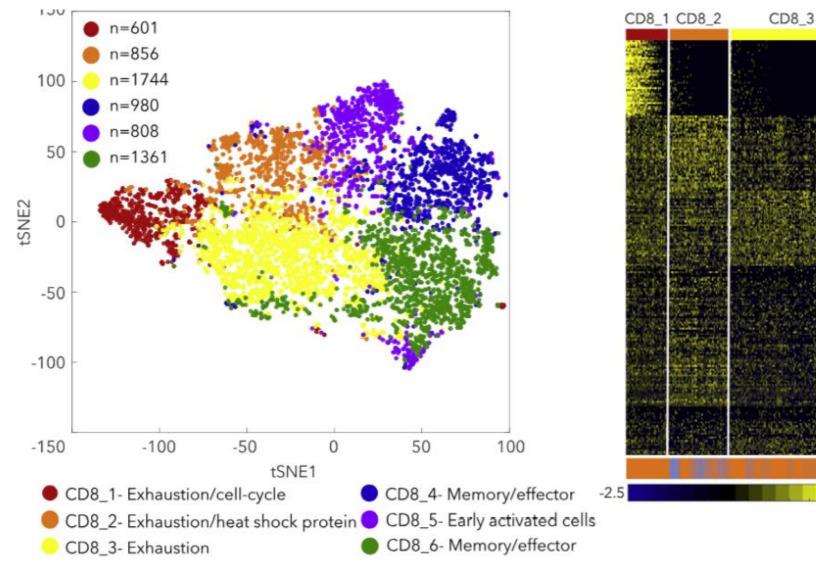


Defining T Cell States Associated with Response to Checkpoint Immunotherapy



Sade-Feldmam, M et al. Cell. 2018 Nov 1;175(4):998-1013.

CD8 subsets identified



Sade-Feldmam, M et al. Cell. 2018 Nov 1;175(4):998-1013

+2.5

CD8_4

CD8_5

CD8_6

CD8_G; CD8_B

CD55 IL7R TCF7 TNF \$100A10

PFKFB3 IL6ST CXCL13 CTLA4

SELL TCF7 LTB IL7R FLT3LG IL16

The heterogeneity of tumor microenvironment

- 1. Broad spectrum of immune infiltration from immune deserted, excluded to inflamed
- 2. Complex components coexisting at the tumor microenvironment
- 3. Diverse architecture and vasculature of tumor
- 4. Balance between immune competent vs immune compromised status
- 5. Insufficient knowledge about immune cell subset functional and activation threshold.
- 6. Potential ways to tip the balance between innate and adaptive immune stimulatory and immune regulatory mechanisms

The complexity of tumor rejection - multi factory phenomena

• Transcription and Abundancy regulations •Translational Chromatin regulation remodeling •Post translation Non coding RNAs modification Circular RNA •Protein analogs Splicing variants •Transporter **Proteomics Epigenomics** Cellomics **Genomics and Environmental Mitochondria** factors Polymorphism Infection Mutation and CNV Microbiome Structure variations Pollution Gene modifiers • Life style Dysfunction Nutrition Culture

Classes of factors influencing immune responsiveness

US NGS commercial providers

Provider Name-Commercial	Provider Name-Commercial	Provider Name-Commercial	Provider Name-Commercial
Accelrys			
ACGT, Inc	DNAStar	Kashi Clinical Laboratories	PrimBio Research Institute
Alpha BioLab			
<u>AltheaDx</u>	Elim Biopharmaceuticals	LC Sciences	Fulcrum Genomics
Ambry Genetics			
Amplicon Express	<u>EpigenDx</u>	Longhorn Vaccines <u>&</u> Diagnostics	Macrogen Clinical Laboratory
Ana-Gen Technologies			
Asuragen	Eureka Genomics	PerkinElmer	Sequencing.com
BGI Americas			
Bio Applied Technologies Joint	<u>Floragenex</u>	Research and Testing Laboratory	Eurofins Genomics
BioReliance			
Centrillion Biosciences	Gene by Gene	<u>Sequetech</u>	Lucigen
Childrens Mercy Hospitals			
City of Hope Integrative Genomics Core	<u>Gene Link</u>	<u>SeqXcel</u>	<u>Medgenome, Inc.</u>
Claritas Genomics			
Cleveland Genomics	GENEWIZ, Inc	Taueret Laboratories	Ashion
Cofactor Genomics			
Complete Genomics	Genome International	Zymo Research	CD Genomics
Covance Genomics Laboratory			
Creative Proteomics Lilly Green	Golden Helix	Fulgent Genetics	