

Next Generation Health Care from Personal Genomics

Min Seob Lee Ph.D.

President & CEO
Diagnomics Inc, San Diego, CA

Executive Director

Personal Genomics Institute, Korea





Forward-Looking Statements

TIME

This presentation contains forward-looking statements about: expectations regarding the expansion and commercialization of Diagnomics' programs; the ability of Diagnomics process; and the ability of Diagnomics to apply our technologies to the development, marketing and commercialization. The actual results may differ materially from those projected in the forward-looking statements due to numerous risks and uncertainties that exist in Diagnomics' operations, development efforts and business environment, including without limitation: the progress or our research and development activities, the ability to enter into future collaborations with other companies, the acceptance of our technologies by the clinical industry, as well as other known risks and uncertainties



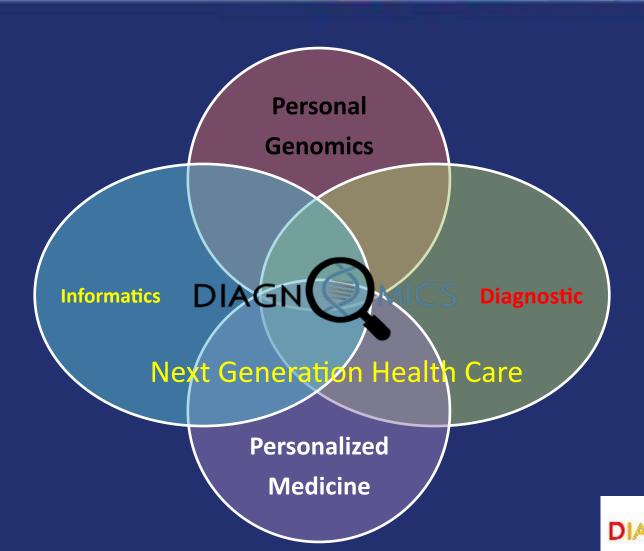
nature Nev neek

Diagnomics' Mission



Next Generation Health Care derived from Personal Genomics

Diagnomics' Business Strategy



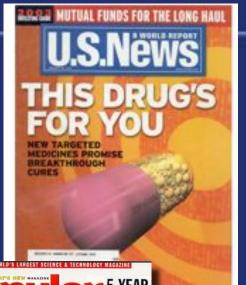
Next \$100 Billion Technology?



- ➤ A technology that will change mankind (Dr. James Watson)
- ➤ Google and PayPal's founders have invested in it
- Major scientific journals are talking about it
- Major companies are scrambling to be a part of it
- Forbes magazine calls it "The next \$100 billion technology market"

Personal Genomics

Personalized Medicine from Genomics



Your Genetic

Card

Your Risk for Deadly Diseases

...and How to Fight Them "Personalized Medicine will radically change the drug develop process."

 "Genomics will dramatically improve success of drug development pipeline"

 "Soon we will be able to get the right drug into the right patient."

 "Applying pharmacogenomics to drug development will cut cycle time to 1.5 - 2 years."

 "Personalized medicine will be able to bring removed drugs back to the market."



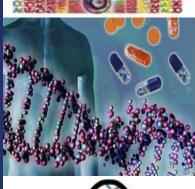


Personal Genomics

- Individual genomes are decoded by genotyping or sequencing, and analyzed by bioinformatics tool
- It can lead to personalized medicine
- Use of Personal Genomics in Predictive Medicine
 - Pharmacogenomics: The study of genome-derived data to predict a body's response to a drug or susceptibility to a disease, i.e., select the most appropriate drug for a patient.
 - ☐ <u>Disease Risk Prediction</u>: Individuals to take steps to prevent or minimize diseases for genetics with increased risk. (Genome Wide Association Studies GWAS)
 - Limitation: Most common medical conditions are multifactoral
 - Combination of genetic and environmental factors
 - Clinical utility is currently limited









Personalized Medicine



- Customization of healthcare
- Decision and practice being tailored to individual patients by evidence
- Mainly involves the systematic use of genetics and other information
- Select or optimize patient's preventative and therapeutic care
- Future of Medicine

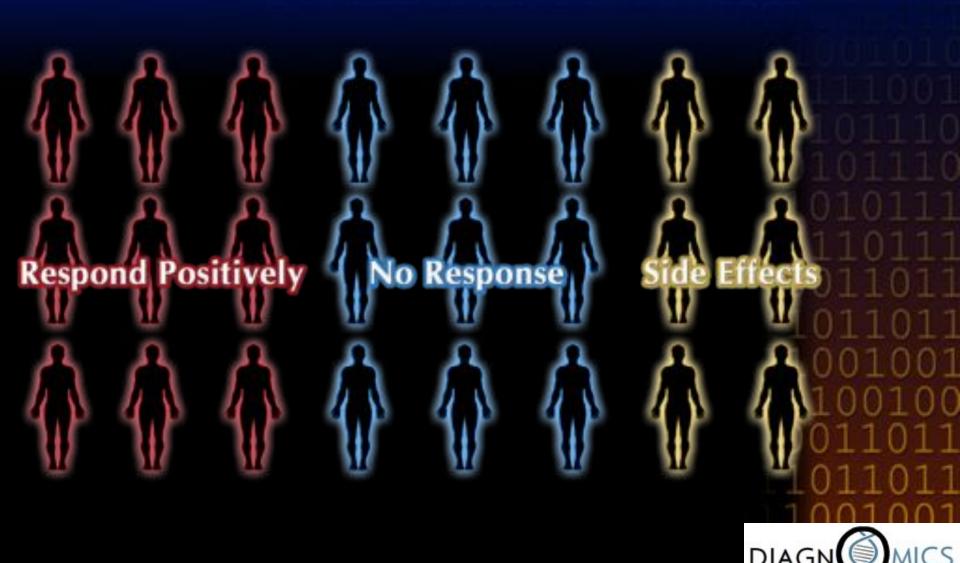


Conventional One-Drug-Fits-All Treatment

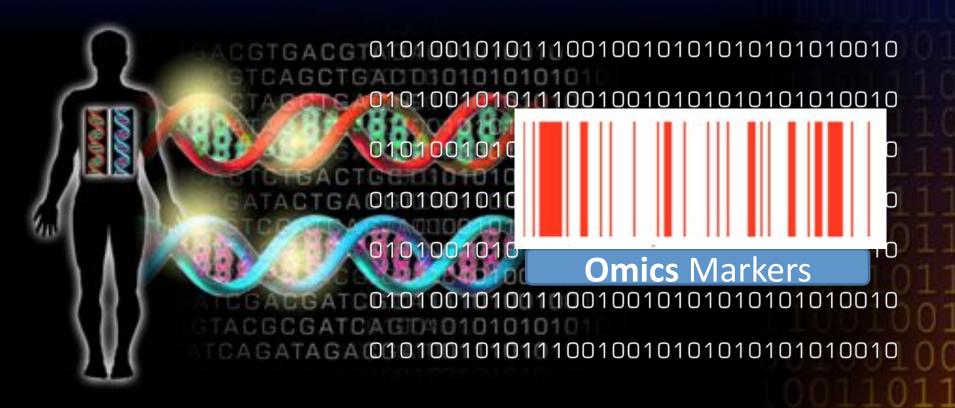




Drug Response Groups

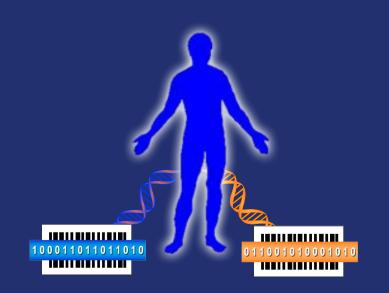


Personalized Medicine from Personalized Genomics

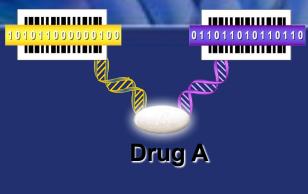


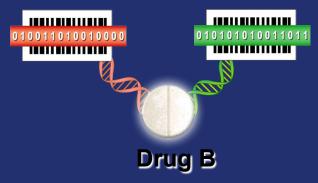
Study of Genomic Difference based on Clinical Outcome

Personalized Medicine from Personalized Genomics





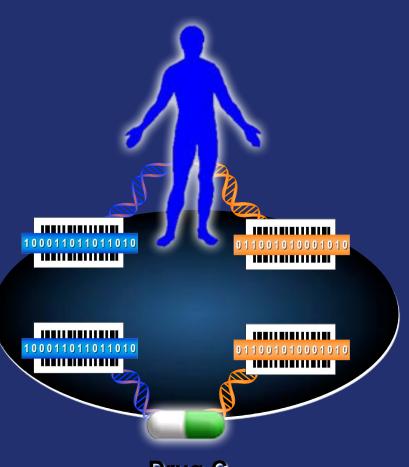






Drug C

Personalized Medicine from Personalized Genomics







Human Genome Project

(1983-2003)



- 20,000 -30,000 genes
- 3 billion base pairs of DNA





-Humans have 99.9% of genome in common -Interested in 0.1% that differs



Personalized Medicine





Although we are 99.9% identical, we are not all the same!





Big or Small Difference?

Difference Among Us is about 1/10 of Chimpanzee





0 0.1%

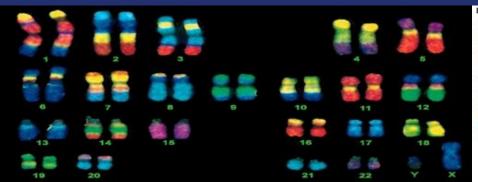
0.5%

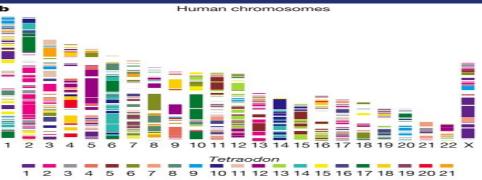
1%

% DNA Difference

Human Genomic Variation

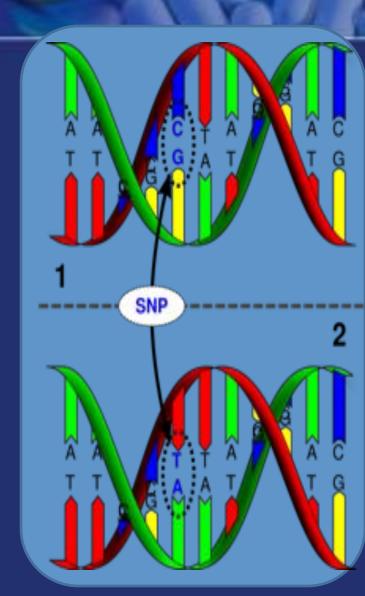
- Human genetic variation in Genome
 - Single nucleotide polymorphisms (SNPs)
 - Copy number differences
 - Insertions
 - Deletions
 - Duplications
 - Rearrangements
- RNA and protein expression differences
- Epigenetic difference



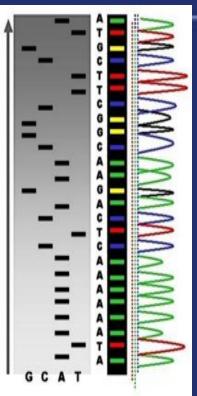


Single Nucleotide Polymorphisms (SNPs)

- Single nucleotide (A,T,C,or G) in the genome sequence is altered, e.g., AAGGCTAA to ATGGCTAA
- Comprise 90% of all human genetic variation
- Exist every 100 to 300 bases in population
- Found every 1000 bases in individual
- Found in both coding (i.e., gene) and noncoding regions of the genome.
- Usually have no effect on cell function, but some could predispose people to disease or influence their response to a drug



Evolution of Sequencing Technology



1953: Watson & Crick - DNA Double Helix

1977: Maxam & Gilbert - Chemical Degradation

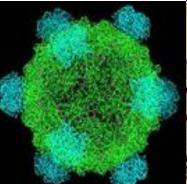
Sequencing, Sanger - Chain Termination, Shotgun Seq.

1977 : First full Genome Sequenced (Phi –X 174)

1987: ABI 370(0) (Automated Sequencing)

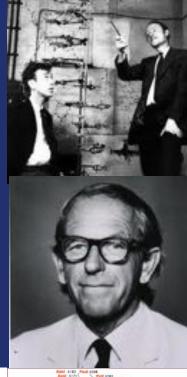
1989 : Human Genome Project Started

2004: Next Generation Sequencing (NGS)









What is Next Generation Sequencing (NGS)?



Massive Parallel Sequencing of Short Nucleotides (Technical Breakthrough for Sequencing)

- Non-Sanger based Sequencing Technology
- Ultimate biological measurement
- Quantitative and Digital output
- Very low noise
- No saturation ceiling
- No probe effects: variable affinity, cross-hyb
- Can be Applied in Novel Biological Applications
- Sample multiplexing
- High throughput / Low cost

Next Generation Sequencing (NGS)

- ►Industrial Revolution: 19th Century "Steam Engine" run by fossil fuel
- ➤ Science & Healthcare Revolution: 21st Century "Genome Engine" run by NGS
 - Applications
 - Cost
 - Time
 - Flexibility
 - Scalable
 - Bioinformatics



Nature Methods 5, 16 - 18 (2008) Published online: 19 December 2007 | doi:10.1038/nmeth1156

Next-generation sequencing transforms today's biology

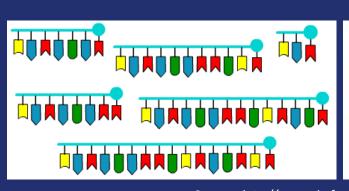
Stephan C Schuster

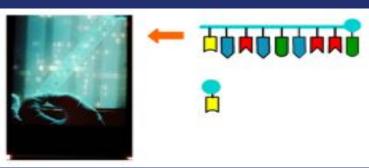
 Stephan C. Schuster is at Pennsylvania State University, Center for Comparative Genomics and Bioinformatics, 310 Wartik Building, University Park, Pennsylvania 16802, USA, e-mail: pca@bx.ppu.edu

A new generation of non-Sanger-based sequencing technologies has delivered on its promise of sequencing DNA at unprecedented speed, thereby enabling impressive scientific achievements and novel biological applications. However, before stepping into the limelight, next-generation sequencing had to overcome the inertia of a field that relied on Sanger-sequencing for 30 years.

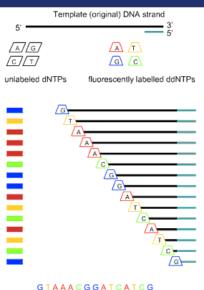
Early Sequencing Technology (1st Gen)

- Sanger sequencing (chain termination)
 - Make millions of different-length copies
 - Print on electrophoresis gel
 - Read lengths small to large
 - Fluorescent indicators denote bases
 - Reassemble small segments with shotgun sequencing
 - First Automated Sequencing Machine ABI 370



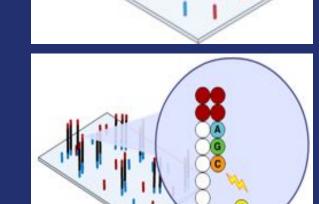






Early Next Generation (2nd Gen) Sequencing Technology

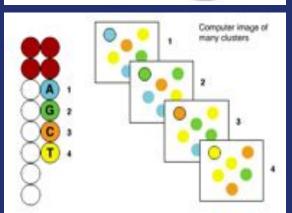
- Illumina Solexa, ABI SOLiD, Roche 454
- Illumina example
 - Attach adaptors to short sequences
 - Amplify by growing clusters (Bridege)
 - Add nucleotides, primers
 - Activate a laser to read bases as incorporated
 - Assemble clusters simultaneously
- Improved read time
 - 10 gigabases/day at \$0.0001 per 1000 bases vs.
 Sanger sequencing (one year at \$1 per 1000 bases)











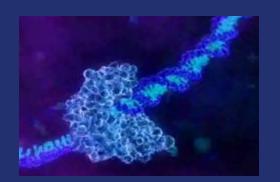
Early Comparison of Existing NGS Methods

62	Feature generation	Sequencing by synthesis
454	Emulsion PCR	Polymerase (pyrosequencing)
Solexa	Bridge PCR	Polymerase (reversible terminators)
SOLiD	Emulsion PCR	Ligase (octamers with two-base encoding)
Polonator	Emulsion PCR Ligase (nonamers)	
HeliScope	Single molecule	Polymerase (asynchronous extensions)

	Cost per megabase	Cost per instrument	Paired ends?	1° error modality	Read-length
454	~\$60	\$500,000	Yes	Indel	250 bp
Solexa	~\$2	\$430,000	Yes	Subst.	36 bp
SOLiD	~\$2	\$591,000	Yes	Subst.	35 bp
Polonator	~\$1	\$155,000	Yes	Subst.	13 bp
HeliScope	~\$1	\$1,350,000	Yes	Del	30 bp

Early 3rd Gen Sequencing

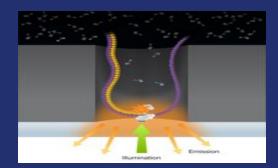
 Sequencing by synthesis pyrosequencing example: Pacific Biosciences SMRT



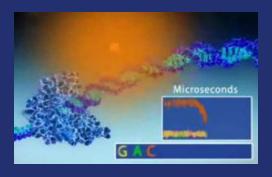
DNA polymerase wrapped around DNA chain



Phospholinked nucleotides



Zero-mode waveguide reads sequence

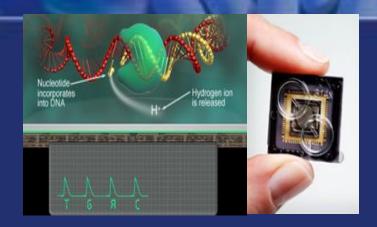


Label fluoresces as cleaved

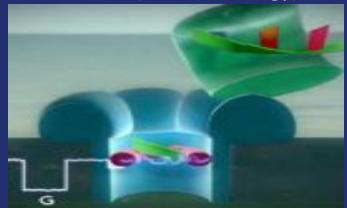


3rd Gen Sequencing Technology

- Electronic sequencing
 - Ion Torrent, NABsys, Oxford
 Nanopore Technologies, Agilent,
 Sequenom, IBM
- Electron microscope reads
 - ZS Genetics, Halcyon Molecular
- George Church's list of next-gen sequencing technologies
 - http://arep.med.harvard.edu/Polonator



Ion Torrent (Life Technology)



Oxford Nanopore Technologies

Genome Becomes Personalized!









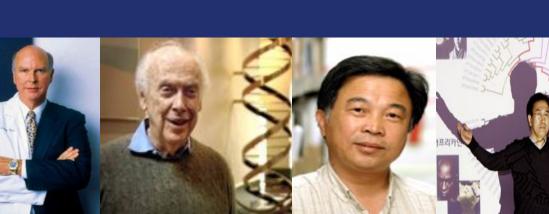






Personal Genome Pioneers

- Reference genome (NCBI-HGP), Pool DNAs 2003
- Craig Venter, Caucasian 2007 (1st)
- James Watson, Caucasian 2008 (2nd)
- Nigerian , African (HapMap) 2008 (3rd)
- Yang, HM, Asian, Han Chinese, 2008 (4th)
- Seong Jin Kim, Asian, Korean 2008 * (5th)
- Rosalynn Gill, Caucasian (PGP9), First Publically Available Published Female Genome2010*
- * Done by PGI Team



























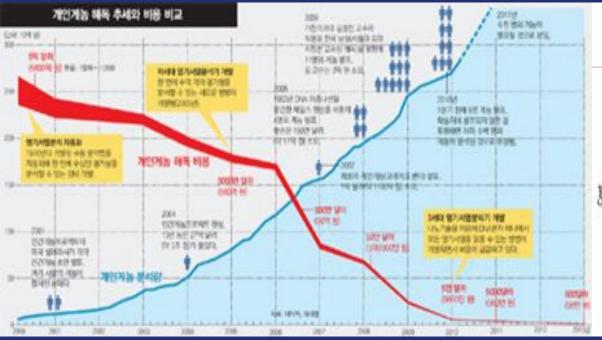
Stephen Quality Da

Personal Genome Analysis (Past, Present, Future)

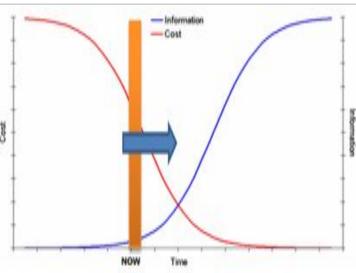
- 2003: HGP: 13 years, \$2.7 billion
- 2007: Craig Venter: 4 years, \$100 million
- 2008: James Watson: 2 years, \$2 million
- 2009: SJ Kim: 6 months, \$170,000
- 2010: 1 month, \$20,000
- 2011: 1 week, \$5,000
- 2015(?): 2 days, <\$1,000







Genome Cost / Genome Information



Personal Genome Sequence **Data Explosion**

- 3 giga base of diploid Genome
- Massive parallel short read of sequence by NGS
- Average 30 X of human genome sequencing (3 GB \times 30 = 90GB)
- 2 3 TB of Raw Data for Solexa, 10 TB of Data for Solid Technology
- Over 100 GB of processed data per individual genome
- 6 billion bases per Person X 6 billion people in the world



DNA sequencing: 10 x / yr improvement

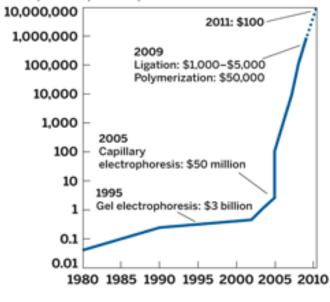


April 1, 2010

A NEW 'MOORE'S LAW'

Improvements in DNA sequencing are driving down the cost of whole genomes

Base pairs sequenced per dollar



NOTE: Dollar figures refer to reagent costs. SOURCE: George Church, Harvard University

Image credit: http://pubs.acs.org/cen/_img/87/i50/8750cover2_law.gif Dec. 14, 2009





Particle Physics
Large Hadron
Collider
(15PB)

http://www.inte
ttp://www.inte
tp://world Wide Webel
p://(~1PB) tel.
://www.intel.co

wiki wiki
iki Wikipedia
iki (10GB) W
ki 100% CAGRVI
i wiki wik

Annual Email Traffic, no spam (300PB+) Internet Archive
(1PB+)

Estimated On-line RAM in Google (8PB)

Personal Digital
Photos / Videos
(1000PB+)
100% CAGR

200 of London's Traffic

Cams
(8TB/day)

2004 Walmart Transaction DB (500TB) Typical Oil Company (350TB+)

Merck Bio
Research DB
(1.5TB/qtr)

UPMC Hospitals
Imaging Data
(500TB/yr)

MIT Babytalk
Speech Experiment
(1.4PB)

Terashake Earthquake

Model of LA Basin

(1PB)

One Day of Instant
Messaging in 2002
(750GB)

Race for Personal Genome **Business Started!**

Who is going to be the Winner?

The \$1,000 Genome and Personalized Medicine





"NHGRI's vision is to...... cut the cost of whole-genome sequencing to \$1,000 or less, which would enable the sequencing of individual genomes as part of medical care. The ability to sequence each person's genome cost-effectively could give rise to more individualized strategies for diagnosing, treating, and preventing disease."

Francis Collins, NHGRI Director, in defense of his FY2006 Budget Request



Complete in a co

Personal

Genome

Project

















Selected Deals in the Next-Generation Sequencing Space

Company	Perteur	Comments		
Applied Boustons	Автохит	6/06 ABI acquired Agenceurt Bencience from Buckman Coulon for \$120 million code.		
Beers	Silve	1/07 Blumms completes sequention of Soleca, enabling entry into next- generation sequencing.		
Roche	454	5/07 Ferolossica of 454 acquisition by Rocke from		
Copies Geomis	-	(0,0		





technologies™





















The First Korean Personal Genome Analysis*

(World 5th, Asia 2nd, Korea 1st)

Downloaded from genome.cshlp.org on August 11, 2009 - Published by Cold Spring Harbor Laboratory Press



The first Korean genome sequence and analysis: Full genome sequencing for a socio-ethnic group

Sung-Min Ahn, Tae-Hyung Kim, Sunghoon Lee, et al.

Genome Res. published online May 26, 2009 Access the most recent version at doi:10.1101/gr.092197.109

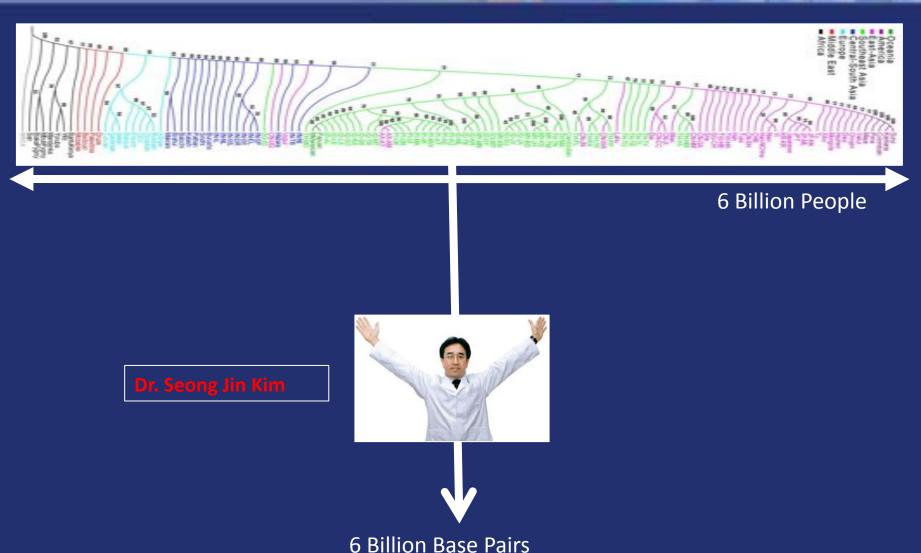


The Karyogram of the donor (Dr. Seong Jin Kim)

No obvious chromosomal abnormalities!



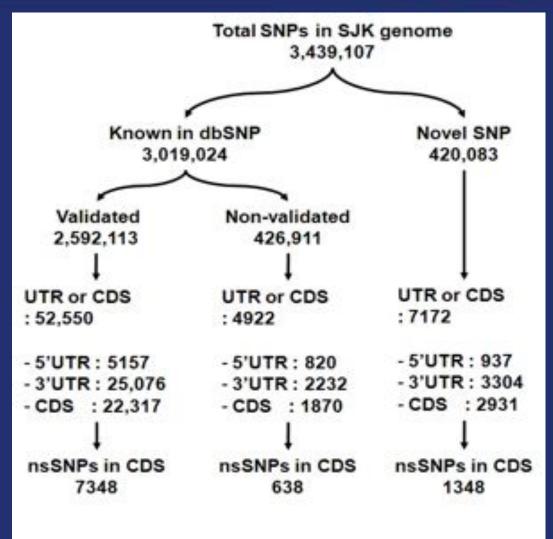
Personal Diversity vs. Population Diversity



The first Korean Genome (SJK)

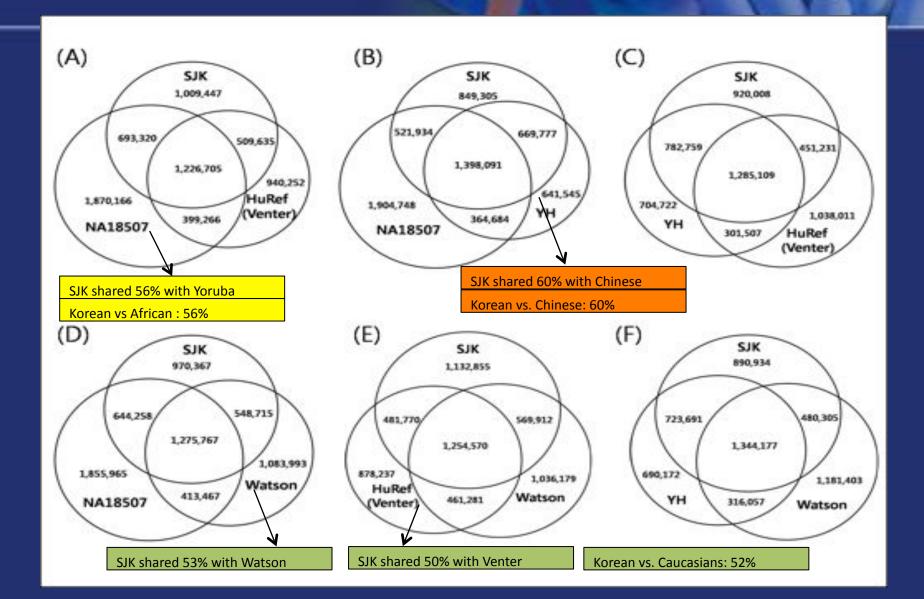
- First annotated and made public on 4th Dec. 2008 (through web and ftp)
- SNP, CNV, indels were analysed
- Automated phenotypic association study was done
- Non-synonymous analysis
- Phylogenetic study of mitochondrial DNA, Y chromosome and autosomes showed Korean relationship to Chinese and Japanese.
- First intra-Asian genome comparison (Chinese and Korean)
- Analyzed at: 7.8, 17.3, 23.5 and 28 x folds
- By Jan. 23.5 fold sequenced and analyzed
- Open freely Available from: http://koreagenome.org

Classification and number of intragenic SNPs



Not represented in dbSNP

Comparison of individual SNPs



Korean Genome Variation Browser

SJK's SNPs

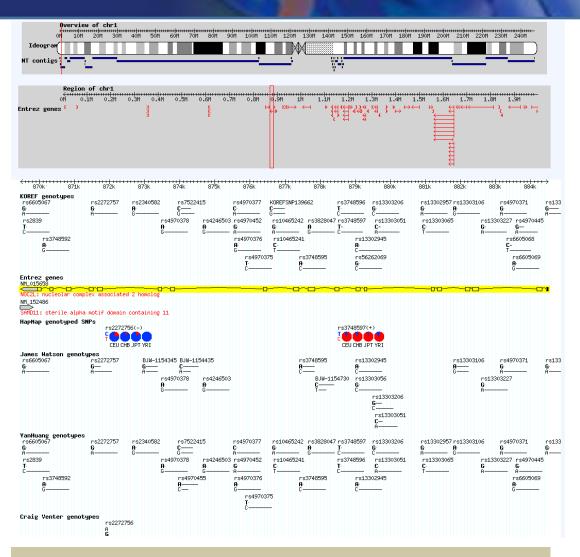
NOC2L gene

Hapmap

Watson's SNPs

YH's SNPs

Venter's SNPs



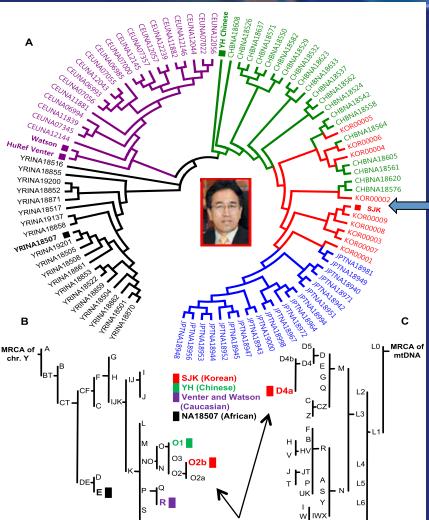
http://koreagenome.org/cgi-bin/gbrowse/kgenome/

SJK's Genetic Lineage



아프리카인







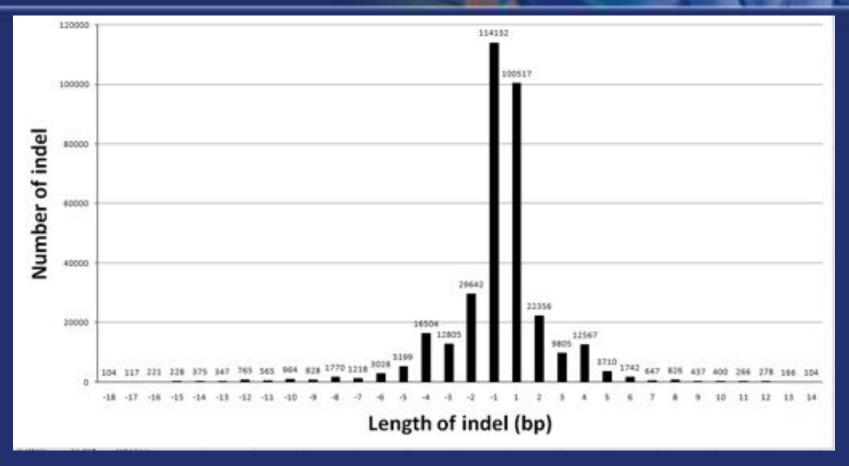
Autosomal phylogenic tree



Chromosome Y haplogroup lineage

mtDNA ethno-geographic lineage

Size distribution and classification of short indels found in SJK

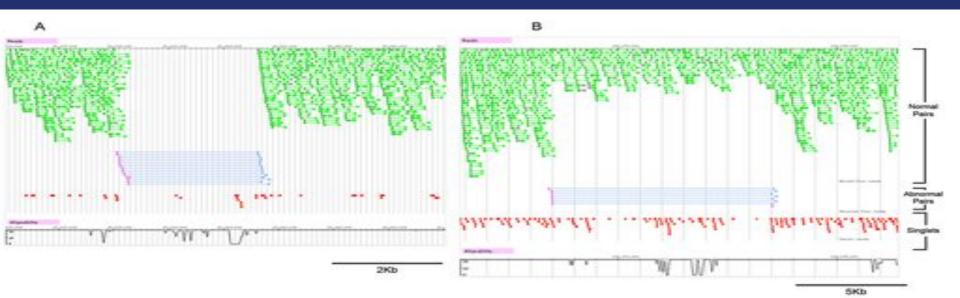


Using MAQ, we identified 342,965 short indels

→ We found that only **247** (0.1%) were validated,**113,287** (33.0%) non-validated, and **229,431** (66.9%) indels were not found in dbSNP

Detection and identification of structural variants

- We found structural variants by using paired-end reads.
 - 1. 2920 deletions (100bp ~ 100kb)
 - 2. 415 inversions (100bp ~ 100kb)
 - 3. 963 insertions (175bp ~ 250bp)
- We found deletion SVs in 21 coding genes.
 - → All heterozygous deletions





Summary of Genetic Disease Susceptibility Results







Your Hello Genome Report

CONFIDENTIAL



Hello Genome

Reference

10111

Highlene

11111 522

Чентори жини

Errotonic Paradigra

11011 122

HOL C

Hello Genome



High level

HIRE BOX

Higher height

HIRE STREET

Type 1 possibility



High levels

11111 123

Week against

11111 28

Normal Possibility

BOAR MARK

High possibility

0110

High your billy

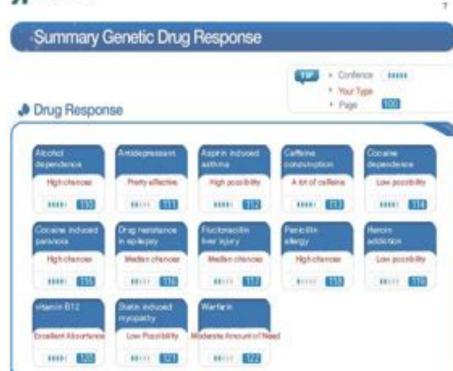
11111

hyroid Function

High TSH Level

11111 2003

briganity



Medium lines

(110) ES

wir ian aging

High Possibility

6618) MOS

Liter Lincol

enter \$103

etum nos



Genetic Disease Susceptibility Results

Age-related macular degeneration

The macuse is very important and is responsible for what we see straight in front of us, allowing see fine data if for activities such as reading and writing, as well as our ability to see of Sometimes the delicate calls of the mecula become damaged and stop working, and there are different conditions which can cause this. If it occurs later in life, it is called "age-related midegeneration", also often known as AMD.

> Related Genes : (3)



Butopean # Asian

SNF	Genotype	Relative I	Risk			Freq	percy	
rs2230199	CC	08 127	202	1	: es: : es: : es:	lı.	+ 0007 + 0007 + 0007 + 0008	1
rs2230200	GG	0.87 0.87 1.2	1.05		1 008 1 109 1 118	.9.	+ 1004 + 105,000 + 114	
rs2241393	GG	077 0%	1,18	11.	- 100 E.A.T. - 100 E.A.T. - 100 E.A.T.	all	- 00 to 00 t	.11

Prevalence of Your Genetic Variant





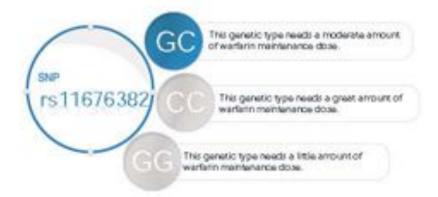
ello Genome

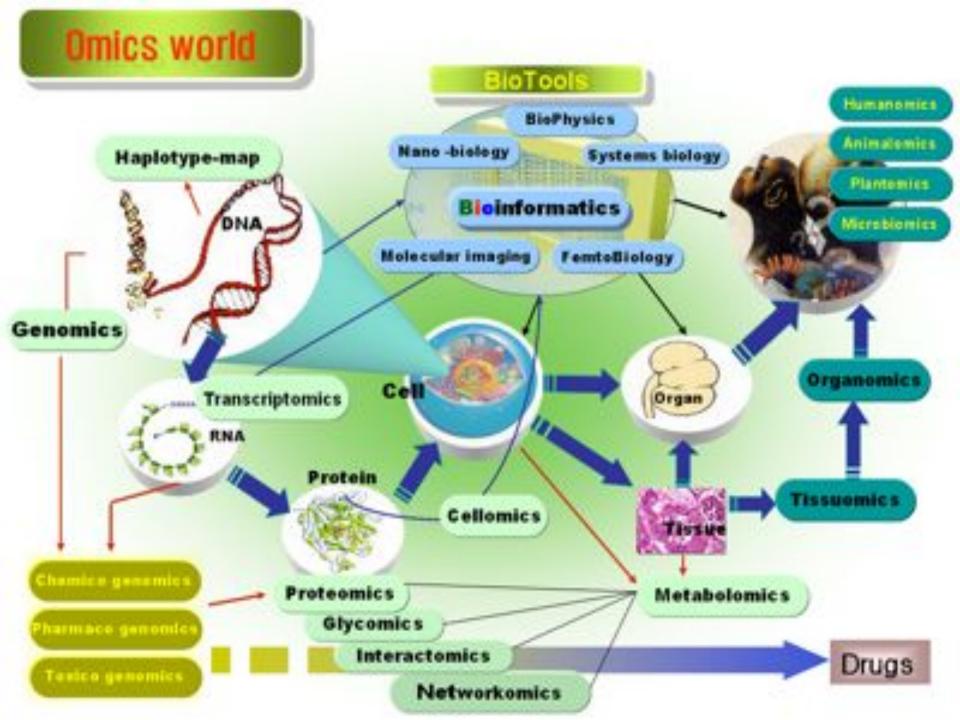
Genetic Drug Response

Warrenin maintenance dose

The most widely prescribed anticoagulant drug in North America. Dosing of warfarin is complicated by the fact that it is known to interact with many commonly-used medications and even with chemicals that may be present in certain foods. These interactions may enhance or reduce warfarin's anticoagulation effect, in order to optimize the therapeutic effect without risking dangerous side effects such as bleeding, close monitoring of the degree of anticoagulation is required by blood testing (NRI).

Your Genetic Variant?





Totalomics by NGS

- Genomics
- Transcriptomics
- Epigenomics

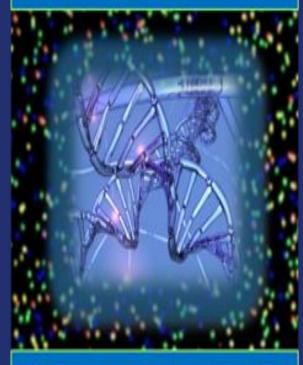
Development of Next Generation Sequencing (NGS) Technology enables study of Totalomics



NGS Applications in Totalomics

- DNA mixtures from diverse ecosystems, metagenomics
- Resequencing previously published reference strains
- Identification of all mutations in an organism
- Errors in published literature
- Expand the number of available genomes (de novo)
- Comparative studies
- Deciphering cell's transcripts at sequence level without knowledge of the genome sequence
- Sequencing extremely large genomes, crop plants
- Detection of cancer specific alleles avoiding traditional cloning
- Chip-seq: interactions protein-DNA
- Epigenomics
- Detecting ncRNA
- Genetic human variation : SNP, CNV (diseases)

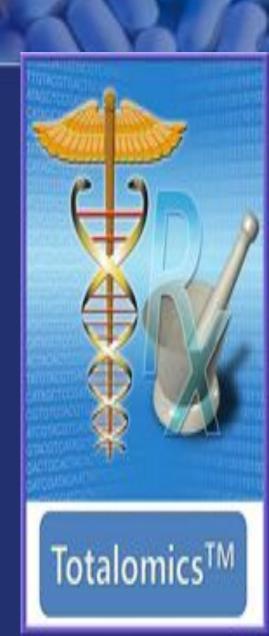
N.G. Sequencing





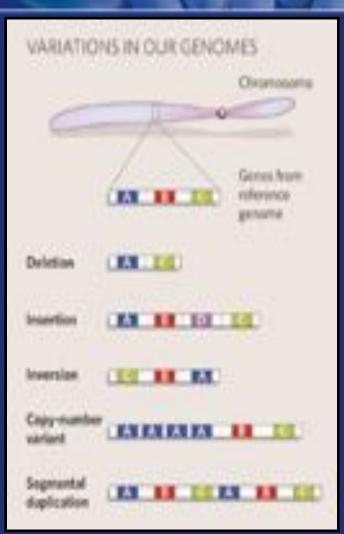
Totalomics Biomarkers

- DNA (Mutation) : Disease causing genes
- RNA (mRNA Expression): Quantitative study of mRNA expression by a cell or a tissue in response to stimuli or clinical outcome
- Genomics (DNA Variation) Identification of multiple DNA variability associated with a functional impairment, disease predisposition or drug response
- Transcriptomics (RNA Variation) Identification of alternative splicing, Gene fusion, RNA editing, Allele Expression, Small RNA
- Epigenomics (Methylation) Methylation provides information on gene regulation that is specific for cell types and diseases.



Human Genome Sequencing (Whole Genome Resequencing)

- Common variants have not yet completly explained complex disease genetics → rare alleles also contribute
- Single Nucleotide Polymorphysm (SNPS)
- Also structural variants, large and small insertions and deletions
- Accelerating biomedical research



De Novo Sequencing

De Novo Sequencing

Draft Map

Fine Map

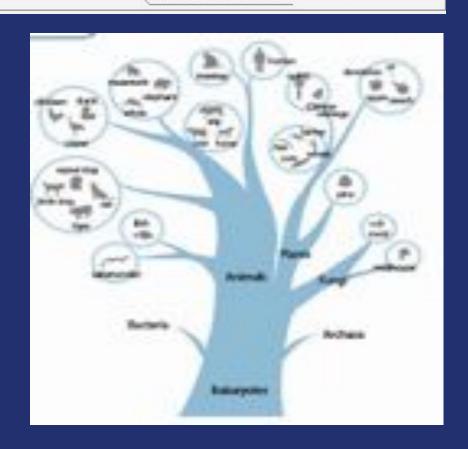
Genome Contents

Comparative & Evolution Genomics

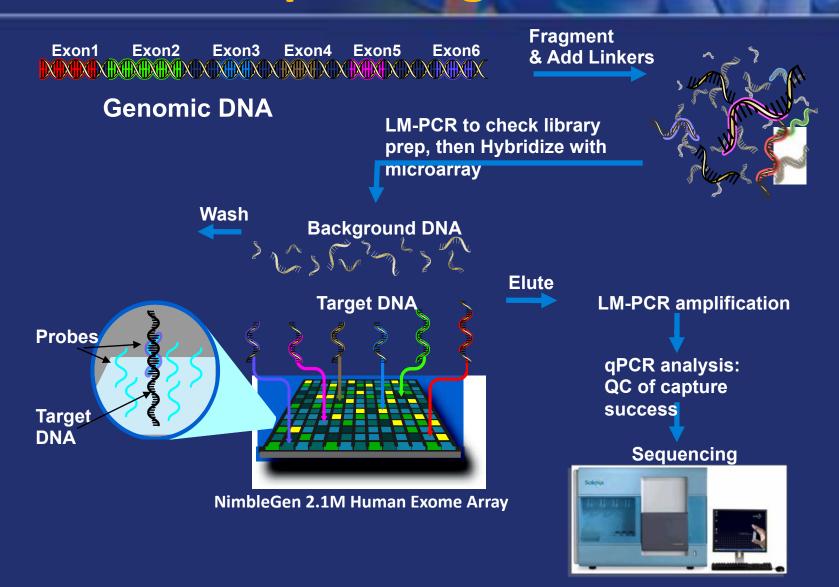
Database Construction

Gene Annotation

- De Novo Sequencing by NGS
- Building a new reference genome
- High Throughput, Low-cost
- Supercomputing / Bioinformatics
- Panda, cucumber, tiger
- 1000 Plant and Animal Reference Genome Project



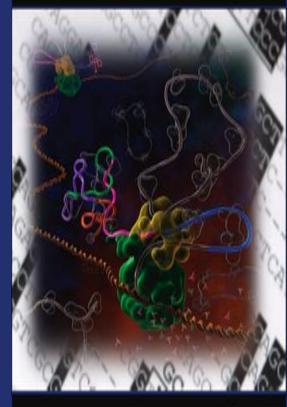
Targeted Sequencing: Exom Sequencing



Transcriptome Analysis RNA Seq

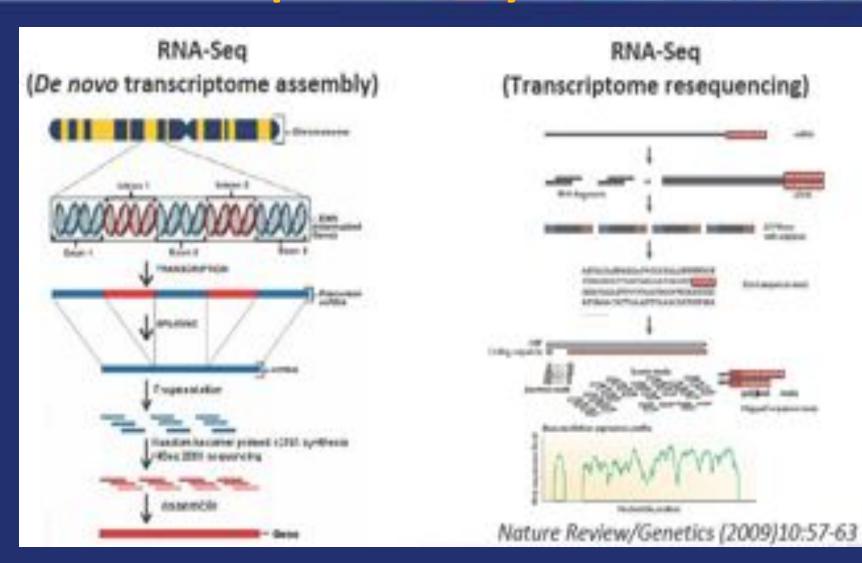
- Whole Transcriptom Shotgun Sequencing (WTSS) by NGS
- The Transcriptome is the complete set of transcripts in a cell, and their quantity, for a specific developmental stage or physiological condition
- Understanding the transcriptome is essential for
 - Interpreting the functional elements of the genome
 - Revealing the molecular constituents of cells, tissues
 - Understanding development and disease

NGS RNA-Seq



TotalomicsTM

RNA Sequencing: Transcriptom Analysis



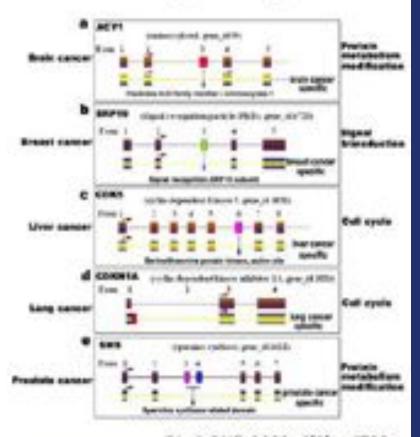
RNA-Sequencing Applications: Alternative Splicing

92%~94% of human multi-exon genes undergo AS



Nature, 2008 Nov 27

Many cancer-associated genes are regulated by AS



PLoS ONE. 2009; 4(3): e4732.

RNA Sequencing Applications: Mutation detection

ORDGOWAL ARTICLE

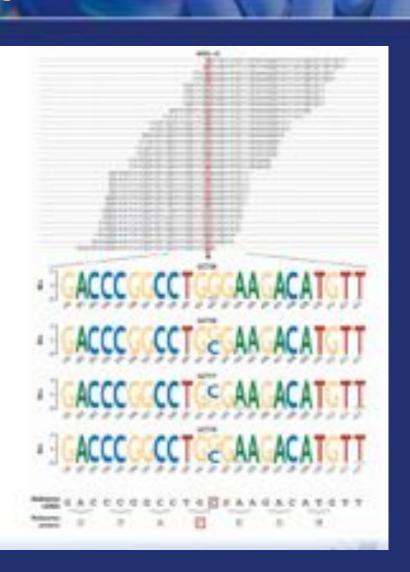
Mutation of FOXL2 in Granulosa-Cell Tumors of the Ovary

- Four adult-type GCTs
- 11 epithelial ovarian tumors

All four GCTs had a missense point mutation, 402C→G(C134W), in FOXL2.

The mutation was absent in 49 SCSTs of other types and in 329 unrelated ovarian or breast tumors.

N Engl J Med. 2009 Jun 25



RNA Sequencing Applications: RNA Editing (post translational)

Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution

Sohrab P. Shah "", Ryan D. Morin", Jaswinder Khattra", Leah Prentice", Trevor Pugh", Angela Burleigh", Allen Delaney", Karen Gelmon", Ryan Guliany, Janine Senz", Christian Steidi. Robert A. Holt", Steven Jones", Mark Sun", Gillian Leung", Richard Moore", Tesa Severson", Greg A. Taylor", Andrew E. Teschendorff", Kane Tse ", Gulisa Turashvill", Richard Varhol", Renè L. Warren", Peter Watson", Yongjun Zhao", Carlos Caldas", David Huntsman ", Martin Hirst", Marco A. Marra" & Samuel Aparicio (1).

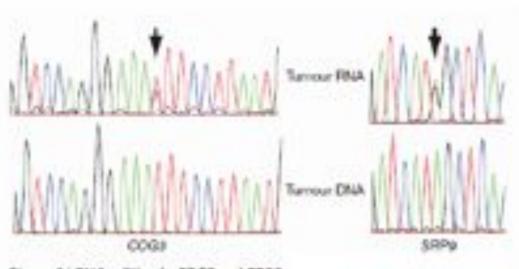


Figure 1 | RNA editing in COG3 and 5RP9.

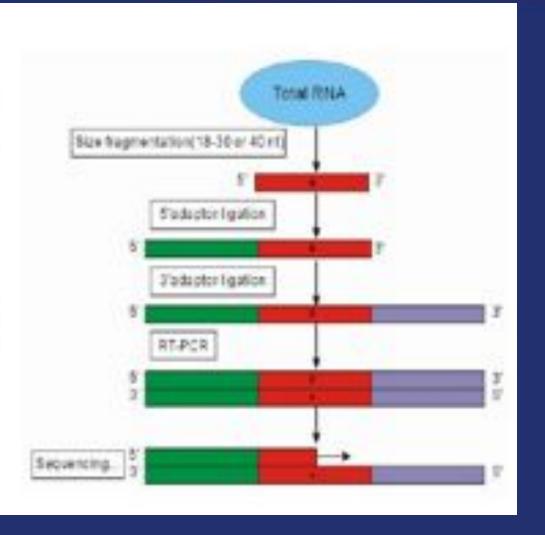
RNA Sequencing Applications: Gene Fusions

Transcriptome Sequencing Identifies New Drug Targets in Prostate Cancer, Gastric Cancer, and Melanoma. Elérd'é Personal resis (10 uMs DOMEST Pointed-and reads SLC-SLAS STAFF Chr. Ngt /

RNA Sequencing Applications: Small RNA detection

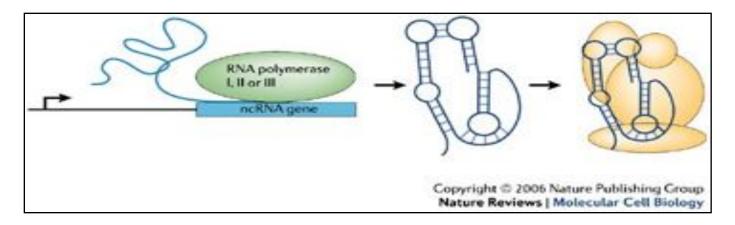


- Small RNA classification and identification
- Small RNA evolutionary analysis
- Roles in cell differentiation and development
- > Roles in biological regulation
- > Small RNA and disease
- Biomarkers identification and therapeutics target detection



RNA Sequencing Applications: ncRNA detection

- ➤ Non-coding RNAs (ncRNA) are very abundant in cells with many unknown function.
- > It is difficult to predict by computational methods
- ➤ Detecting expression level changes that correlate with changes in environmental factors, with disease onset and progression, complex disease set or severity
- > They constitute a new class of biomarkers.



RNA Sequencing vs. Microarray

Quantitate whole-genome gene expression using newgeneration sequencer Illumina HiSeq 2000

Comparing with DNA microarray

- Reproducibility highly reproducible
- Sensitivity detect 50% more genes, wider dynamic range
- Specificity more accurate, no cross-hybridization
- Compatible with data-analysis software

Samples: FirstChoice® Human Brain Reference RNA (HBRR)

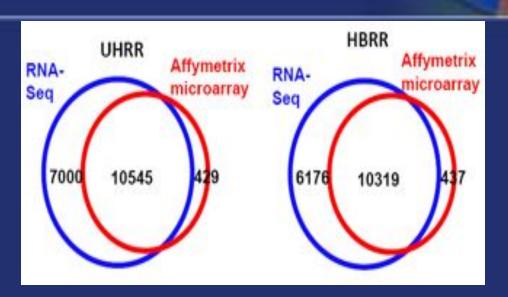
Stratagene Universal Human Reference RNA (UHRR)

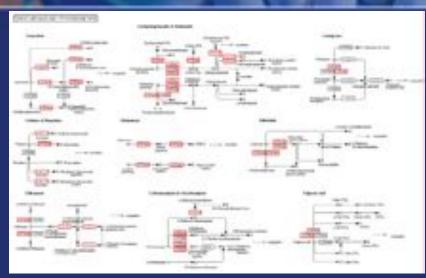
	RNA-Seq (quantification)	Affymetrix microarray			
Technique	New generation sequencing	Hybridization			
Detected signal	Digital	Analogue			

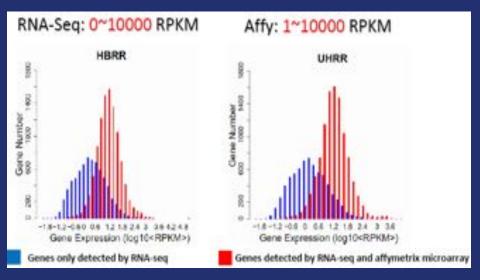


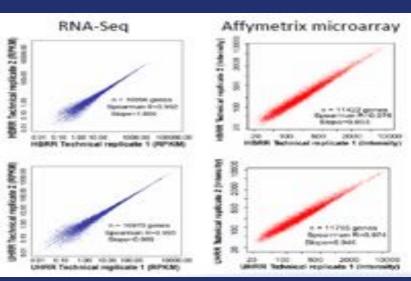


RNA Sequencing vs. Microarray









Summary of Transcriptome RNA Sequencing

- Does not require existing genomic sequence
- Very low background noise
- Can detect novel transcripts
- Resolution up to 1bp
- High-throughput analysis
- Can reveal sequence variations (SNPs), alternative splicing, gene fusion, RNA editing, allele specific expression,
- May replace microarray technology
- Innovative tools for biomarker discovery

NEWS AND VIEWS

Nature Methods 5, 585 - 587 (2008) doi:10.1038/nmeth0708-585

The beginning of the end for microarrays?

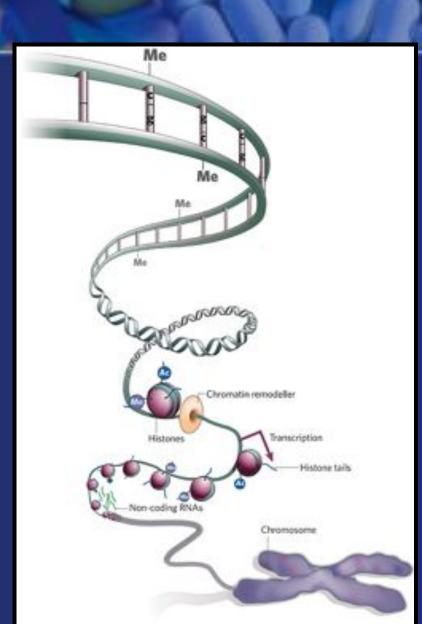
Jay Shendure

 Jay Shendure is in the Department of Genome Sciences, University of Washington, Seattle, Washington 98195, USA. email: shendure@u.washington.edu



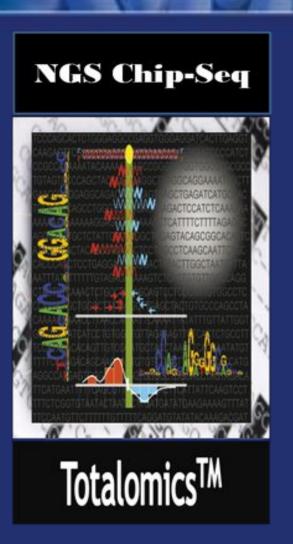
Epigenomics

- Enable of genome-wide patterns of methylation and how this patterns change through the course of an organism's development.
- Enhanced potential to combine the results of different experiments, correlative analyses of genome-wide methylation, histone binding patterns and gene expression, for example.



Chip-Seq (chromatin immunoprecipitation sequencing)

- Key part in regulating gene expression
- Chip: technique to study DNA-protein interaccions
- Recently genome-wide ChIP-based studies of DNAprotein interactions
- Readout of ChIP-derived DNA sequences onto NGS platforms
- Insights into transcription factor/histone binding sites in the human genome
- Enhance our understanding of the gene expression in the context of specific environmental stimuli
- Potential to detect mutations in binding-site



Chip-Seq (chromatin immunoprecipitation sequencing)

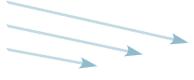
UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly move <<< << > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x position/search_chr21:33,020,037-33,033,335 jump clear size 13, 299 bp. gene OWS: (922-51) 21012 21011 21011-2 21021-1 CHING SIGNLS SIGNATURE dica le DOMEST ARREST Shippers. OWNER I 1204 MyChiphand_normalized MyCh isnaes_hones t 10 . UCSC Denes Based on Refteg, Unifrot, Denbank, CCDS and Comparative Dengalics SCHOLARS M \$001 Merchinistra (Co. RefSea Genes Auriton Secus Hussan saffrets: from Gendlank Human selfrets Human ESTs That Nave Seen Spliced Selices ESTs Placevital Hannal Basevise Conservation by PhyloP Fishball Cond Multiz Alignments of 46 Vertebrates E Normanit Opposit stum Chicken x_brosicalis Seprar Lon DMFK (EDE) 118 **Repost Naster**

Metagenomics

- Characterizing the biodiversity found on Earth
- The growing number of sequenced genomes enables us to interpret partial sequences obtained by direct sampling of specif environmental niches.
- Examples: ocean, acid mine site, soil, coral reefs, human microbiome which may vary according to the health status of the individual

THE METAGENOMICS PROCESS





Extract all DNA from microbial community in sampled environment



DETERMINE WHAT THE GENES ARE (Sequence-based metagenomics)

- Identify genes and metabolic pathways
- Compare to other communities
- and more...

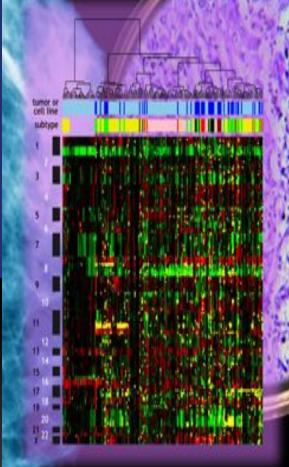
DETERMINE WHAT THE GENES DO(Function-based metagenomics)

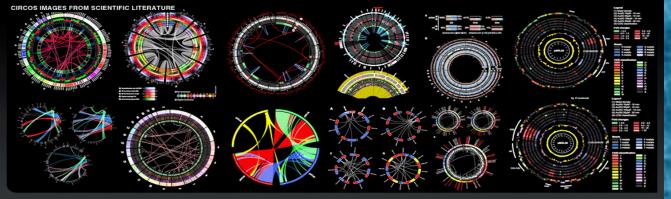
- Screen to identify functions of interest, such as vitamin or antibiotic production
- Find the genes that code for functions of interest
- and more...

Canceromics

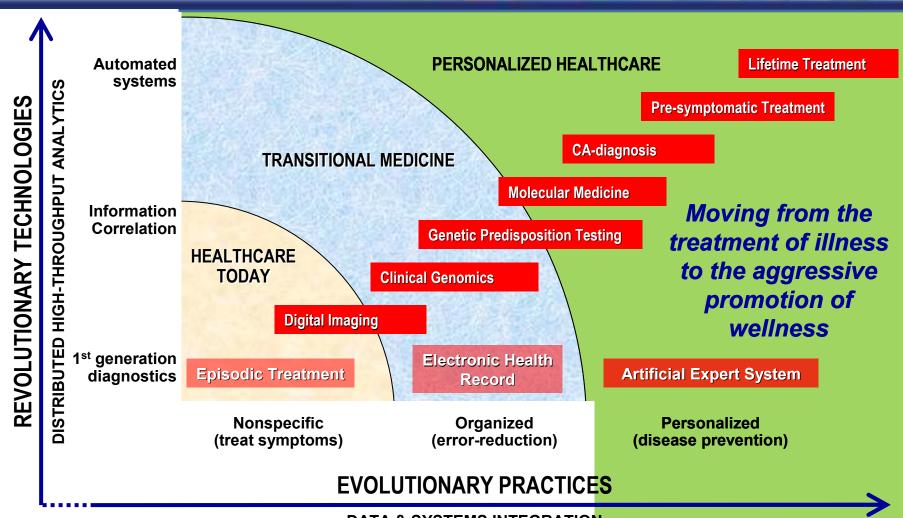
- Cancer is the most complex disease of the genome
- Acquisition of somatic mutation and genome alteration
- The genome records a history of disease
- The genome information provides prognosis of the disease
- Enable personalized treatment and diagnosis







Practice of medicine is moving from the treatment of illness to the aggressive promotion of wellness



Application of Personal Genomics



Thanks!



