Genomic Medicine

Hsueh-Wei Chang 張學偉
高醫生物醫學暨環境生物學系

changhw@kmu.edu.tw

From Genetic Medicine to Genomic Medicine

The “Old Genetics”

Involves conditions wholly caused by:

An extra or missing complete chromosome
or part of chromosome
e.g., Down syndrome

A mutation in a single gene
e.g., Cystic fibrosis, sickle cell disease

The “Old Genetics”

These conditions:

-- are of great importance to individuals and families with them

-- But, even when added together, are relatively rare.

-- Most people not directly affected

-- Thus, genetics played small role in health care (and in society)

Medicine and the New Genomics

- Gene Testing
- Gene Therapy
- Pharmacogenomics

Anticipated Benefits

•improved diagnosis of disease
•earlier detection of genetic predispositions to disease
•rational drug design
•gene therapy and control systems for drugs
•personalized, custom drugs

Human Genome Program, U.S. Department of Energy, Genomics and Its Impact on Medicine and Society: A 2001 Primer, 2001

Genomics

- The study of the structure and function of the genome.
- The experimental study of complete genomes.

Genomic Medicine

Comes largely from knowledge emanating from the Human Genome Project

Impact on Bioinformatics

- Genomics produces high-throughput, high-quality data, and bioinformatics provides the analysis and interpretation of these massive data sets.

- It is impossible to separate genomics laboratory technologies from the computational tools required for data analysis.

Genomics Technologies

- Automated DNA sequencing
- Automated annotation of sequences
- DNA microarrays
  - gene expression (measure RNA levels)
  - single nucleotide polymorphisms (SNPs)
- Protein chips (SELDI, etc.)
- Protein-protein interactions

Genetic Polymorphism
基因多型性

- **Genetic Polymorphism:**

  A difference in DNA sequence among individuals, groups, or populations.

- **Genetic Mutation:**

  A change in the nucleotide sequence of a DNA molecule. Genetic mutations are a kind of genetic polymorphism.

articles

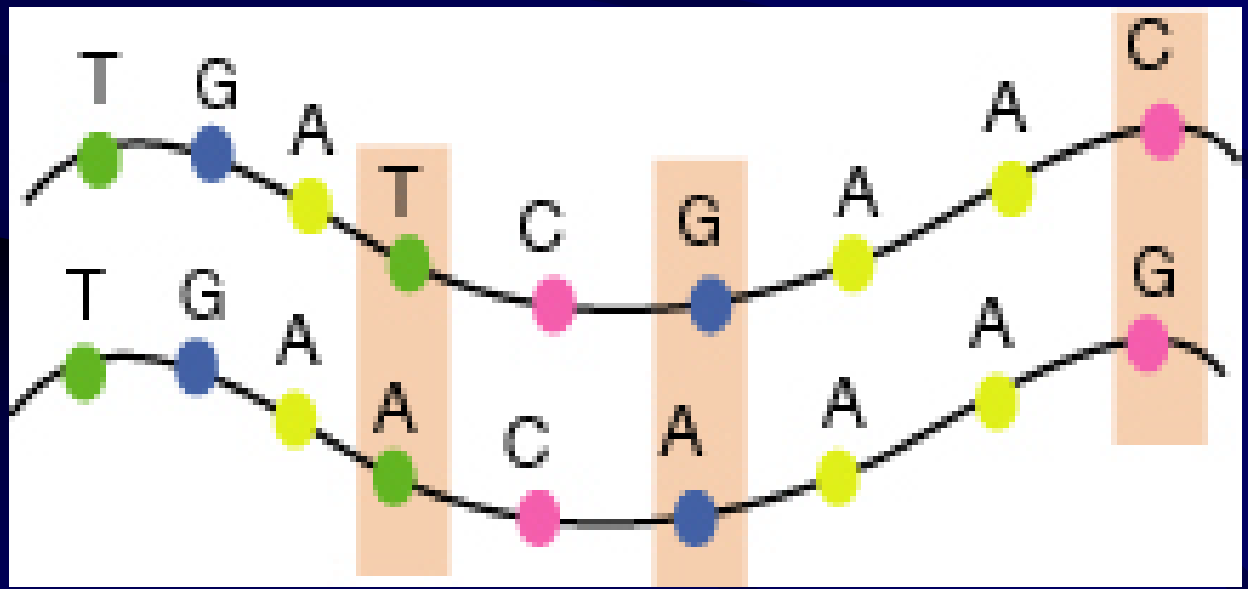
A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms

The International SNP Map Working Group*

* *A full list of authors appears at the end of this paper.*

SNP = Single Nucleotide Polymorphism

(read in SNiP)



Single nucleotide polymorphisms (SNPs) are an abundant form of genome variation, distinguished from rare variations by a requirement for the least abundant allele to have a frequency of 1% or more.

4 states of a SNP are possible (A,G,C,T)

But the diallelic type is the most common SNP (e.g. T->A)

Table 1. Publicly available SNP databases on the internet

Public SNP databases	URL
DBSNP NCBI database of deposited SNPs	www.ncbi.nlm.nih.gov/SNP
The SNP Consortium[a,b]	snp.cshl.org
Cancer Genome Anatomy Project: GAI	lpg.nci.nih.gov/GAI
HGVbase[a,b]	Hgvbase.cgb.ki.se
Database of Japanese Single Nucleotide Polymorphisms[a,b]	snp.ims.u-tokyo.ac.jp

[a]Deposited in dbSNP.
[b]International database.
[c]Search EST sequences for SNPs.
Abbreviations: EST, expressed sequence tag; GAI, Genetic Annotation Initiative; HGV, Human Genome Variation; NCBI, National Center for Biotechnology Information; SNP, single nucleotide polymorphism.

TRENDS in Molecular Medicine Vol.8 No.6 June 2002, page 288.

Single nucleotide polymorphisms (SNPs) are common DNA sequence variations among individual alleles

SNPs will be distributed across the human genome

~ 25,000	non-synonymous cSNPs
~ 50,000	synonymous cSNPs
~ 25,000	regulatory region SNPs
~ 50,000	intragenic non-coding SNPs
~ 50,000	distributed intergenic SNPs

SNP Distribution

- SNP Distribution is not uniform for any of the three categories:
  - Over a complete genome (1/3 in coding, 2/3 in non-coding).
  - Over all the chromosomes (fewer SNPs in sex chromosomes).
  - Over a single chromosome (SNPs often concentrated around a specific location).

SNP markers

- SNPs are very common in the human population.
- SNPs can be found that are *linked* to any disease alleles.
- These mutations are likely to be neutral - they have no direct effect on phenotype
- Linked SNPs can be used as markers for the disease in diagnostic tests.

Classification of SNP

1. Coding SNPs
   - Replacemental polymorphism (change a.a.)
   - Synonymous polymorphism (unchange a.a.)
2. Non-coding SNPs
   - 3‘,5’-NTR (non-transcribed region)
   - 3‘,5’-UTR (untranslated region)
   - Intron
   - intergenic

Coding SNP

SNPs in Functional Proteomics

- SNP related functional proteomics involve the identification of functional SNPs that modify proteins and protein active sites structure and function.
- By identifying the modifications generated by functional (coding) SNPs in disease related proteins, "new compounds can be designed for correcting or enhancing the effects of those mutations in the population."

The Effects of Non-Synonymous SNPs

- **Phenotype**: The observable properties of an individual as they have developed under the combined influences of the individual's genotype and the effects of environmental factors.

- **Genotype**: An exact description of the genetic constitution of an individual, with respect to a single trait or a larger set of traits (sequence of complete set of genes, both dominant and recessive - “SNP scoring”).

- The genotype-phenotype relation forms the basis and motivation for SNP research. If SNPs account for diversity in genotypes, then SNPs also can be mapped to account for diversity in phenotypes.

SNP Topics

1. SNP Genotyping and Haplotyping
   - detection, high-throughput genotyping, haplotyping, the haplotype map
2. Haplotype Mapping
   - the haplotype map
3. SNP Applications
   - pharmacogenomics, diagnostic genomics, functional proteomics and therapeutic genomics

SNP Genotype & Haplotype

SNP Mapping

- Identify SNP sites along the genome to track disease genes.
- A human SNP map would specify the contributions of individual genes to diseases and other phenotypes.

What is an SNP?

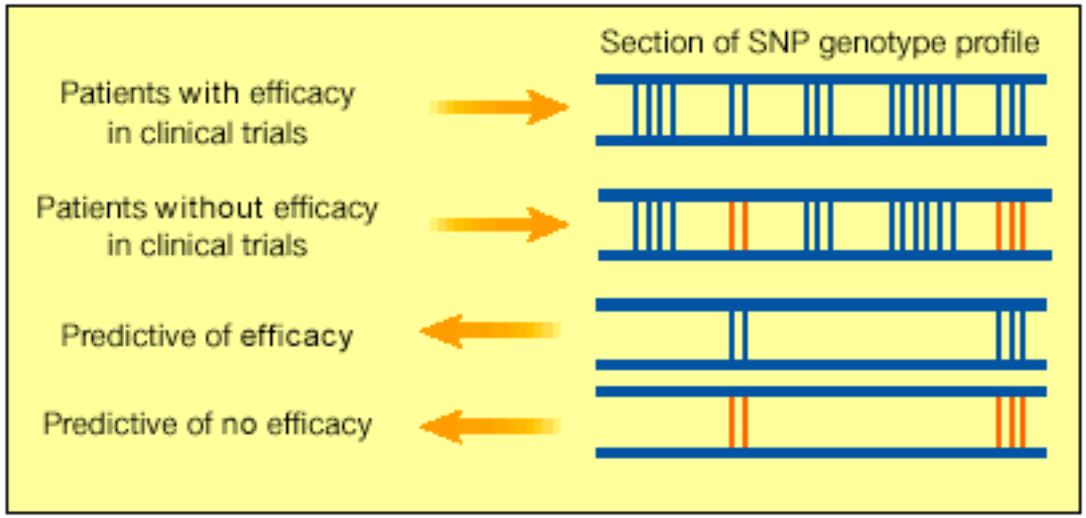
Different people can have a different nucleotide or base at a given location on a chromosome

. . . G G **T** A A C T G . . .

. . . G G **C** A A C T G . . .

What is an SNP map?

How can an SNP map be used to predict medicine response?



Genotype and Haplotype

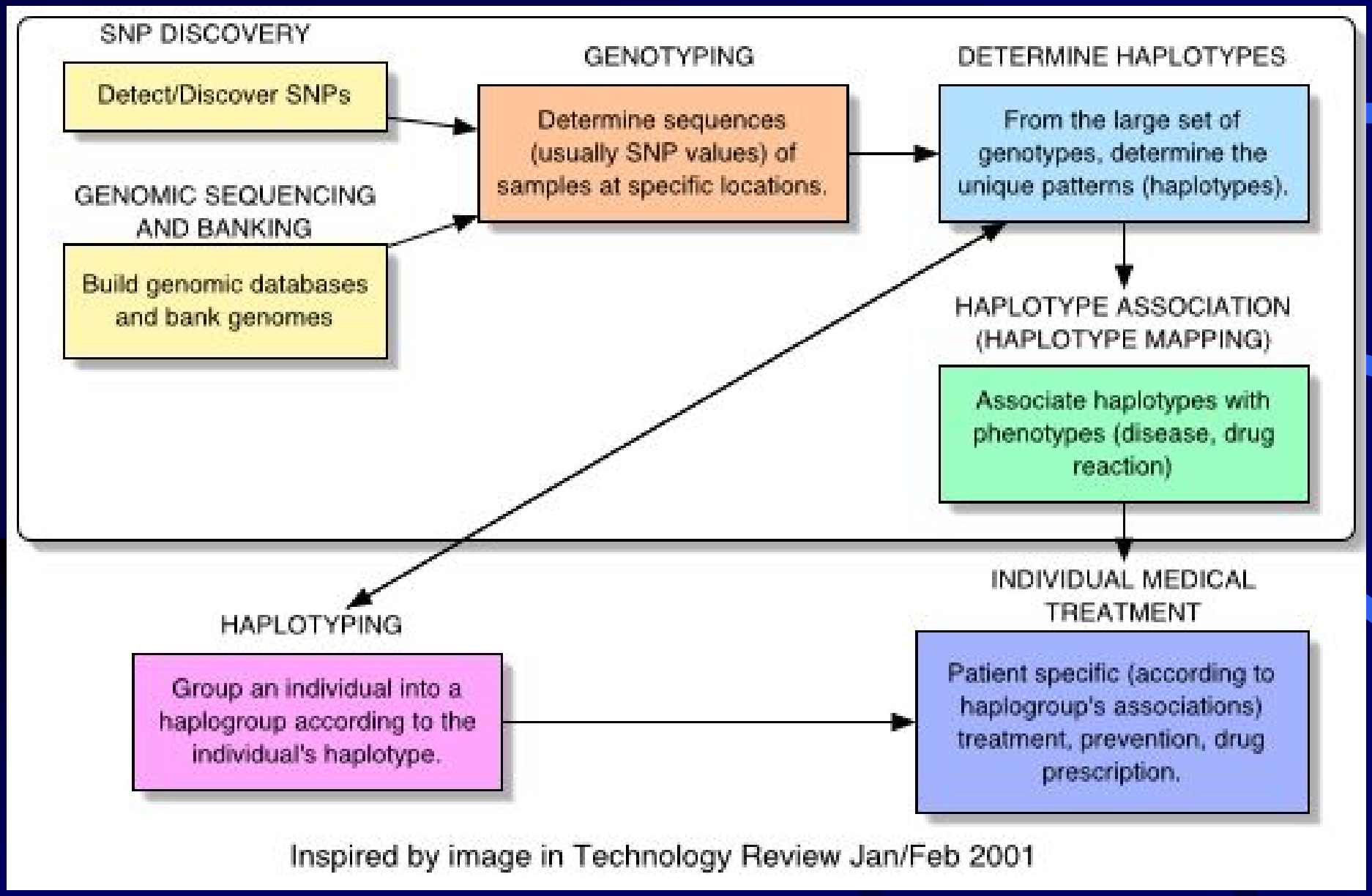
- In the most basic sense, a haplotype is a "haploid genotype".
- **Haplotype:** particular pattern of sequential SNPs (or alleles) found on a single chromosome. These SNPs tend to be inherited together over time.
- **Haplotyping**: involves grouping subjects by haplotypes, or particular patterns of sequential SNPs, found on a single chromosome. There are thought to be a small number of haplotype patterns for each chromosome.
- Microarrays, mass spectrometry and sequencing are all used to accomplish haplotyping.

Haplotype Map

- The map's purpose is to relate human genetic variation with disease predisposition, specifically common or complex disorders.

- For specific genomic locations (chosen to avoid recombination and recurrent mutation), a small number of SNP patterns (haplotypes) were found that account for 80%-90% of the entire human population.

SNP Mapping (cont'd)



Complex promoter and coding region β_2-adrenergic receptor haplotypes alter receptor expression and predict *in vivo* responsiveness

Connie M. Drysdale*†, Dennis W. McGraw*‡, Catharine B. Stack†, J. Claiborne Stephens†, Richard S. Judson†, Krishnan Nandabalan†, Kevin Arnold†, Gualberto Ruano†, and Stephen B. Liggett‡§¶

PNAS | September 12, 2000 | vol. 97 | no. 19 | 10483–10488

Nucleotide:	-1023	-709	-654	-468	-406	-367	-47	-20	46	79	252	491	523
Alleles:	G/A	C/A	G/A	C/G	C/T	T/C	T/C	T/C	G/A	C/G	G/A	C/T	C/A
Haplotype													
1	A	C	G	C	C	T	T	T	A	C	G	C	C
2	A	C	G	G	C	C	C	C	G	G	G	C	C
3	G	A	A	C	C	T	T	T	A	C	G	C	C
4	G	C	A	C	C	T	T	T	A	C	G	C	C
5	G	C	A	C	C	T	T	T	G	C	G	C	C
6	G	C	G	C	C	T	T	T	G	C	A	C	A
7	G	C	G	C	C	T	T	T	G	C	A	T	A
8	G	C	A	C	C	T	T	T	A	C	A	C	A
9	A	C	G	C	T	T	T	T	A	C	G	C	C
10	G	C	G	C	C	T	T	T	G	C	A	C	C
11	G	C	G	C	C	T	T	T	G	C	G	C	C
12	A	C	G	G	C	T	T	T	A	C	G	C	C
Location:	5'	5'	5'	5'	5'	5'	AA19 BUP Cys/Arg	5'	AA16 Gly/Arg	AA27 Gln/Glu	syn	AA164 Thr/Ile	syn

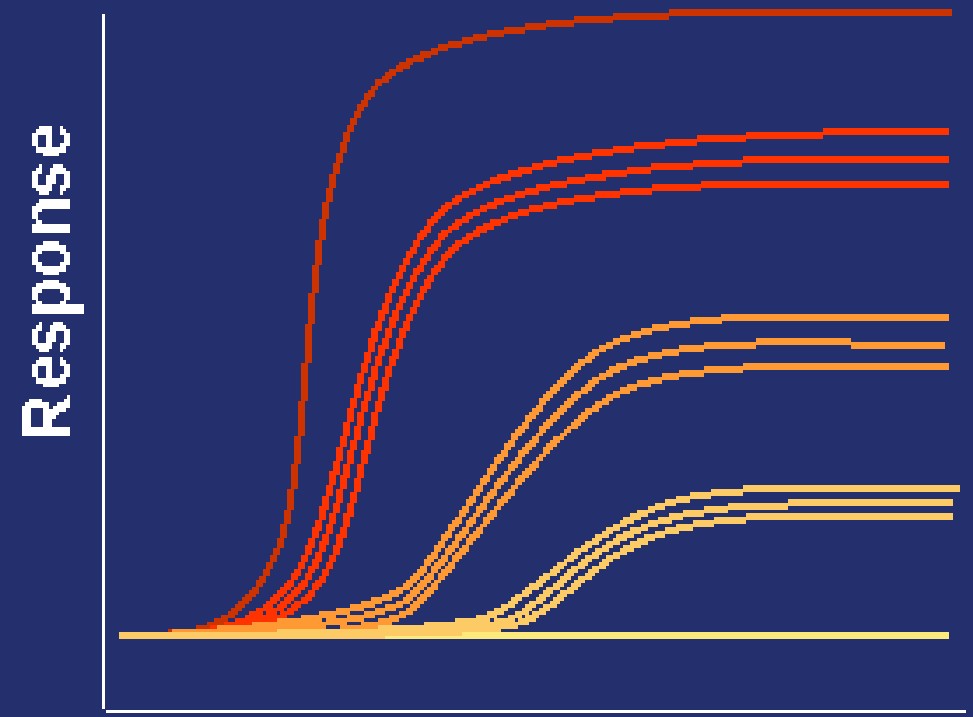
Haplotype Diversity and Linkage Disequilibrium at Human *G6PD*: Recent Origin of Alleles That Confer Malarial Resistance

Sarah A. Tishkoff,[1,2]* Robert Varkonyi,[2] Nelie Cahinhinan,[2] Salem Abbes,[3] George Argyropoulos,[4] Giovanni Destro-Bisol,[5] Anthi Drousiotou,[6] Bruce Dangerfield,[7] Gerard Lefranc,[8] Jacques Loiselet,[9] Anna Piro,[10] Mark Stoneking,[11] Antonio Tagarelli,[10] Giuseppe Tagarelli,[10] Elias H. Touma,[9] Scott M. Williams,[12]† Andrew G. Clark[2]

Haplotype Variation and Linkage Disequilibrium in 313 Human Genes

J. Claiborne Stephens,* Julie A. Schneider, Debra A. Tanguay, Julie Choi, Tara Acharya, Scott E. Stanley, Ruhong Jiang, Chad J. Messer, Anne Chew, Jin-Hua Han, Jicheng Duan, Janet L. Carr, Min Seob Lee, Beena Koshy, A. Madan Kumar, Ge Zhang, William R. Newell, Andreas Windemuth, Chuanbo Xu, Theodore S. Kalbfleisch, Sandra L. Shaner, Kevin Arnold, Vincent Schulz, Connie M. Drysdale, Krishnan Nandabalan, Richard S. Judson, Gualberto Ruaño, Gerald F. Vovis

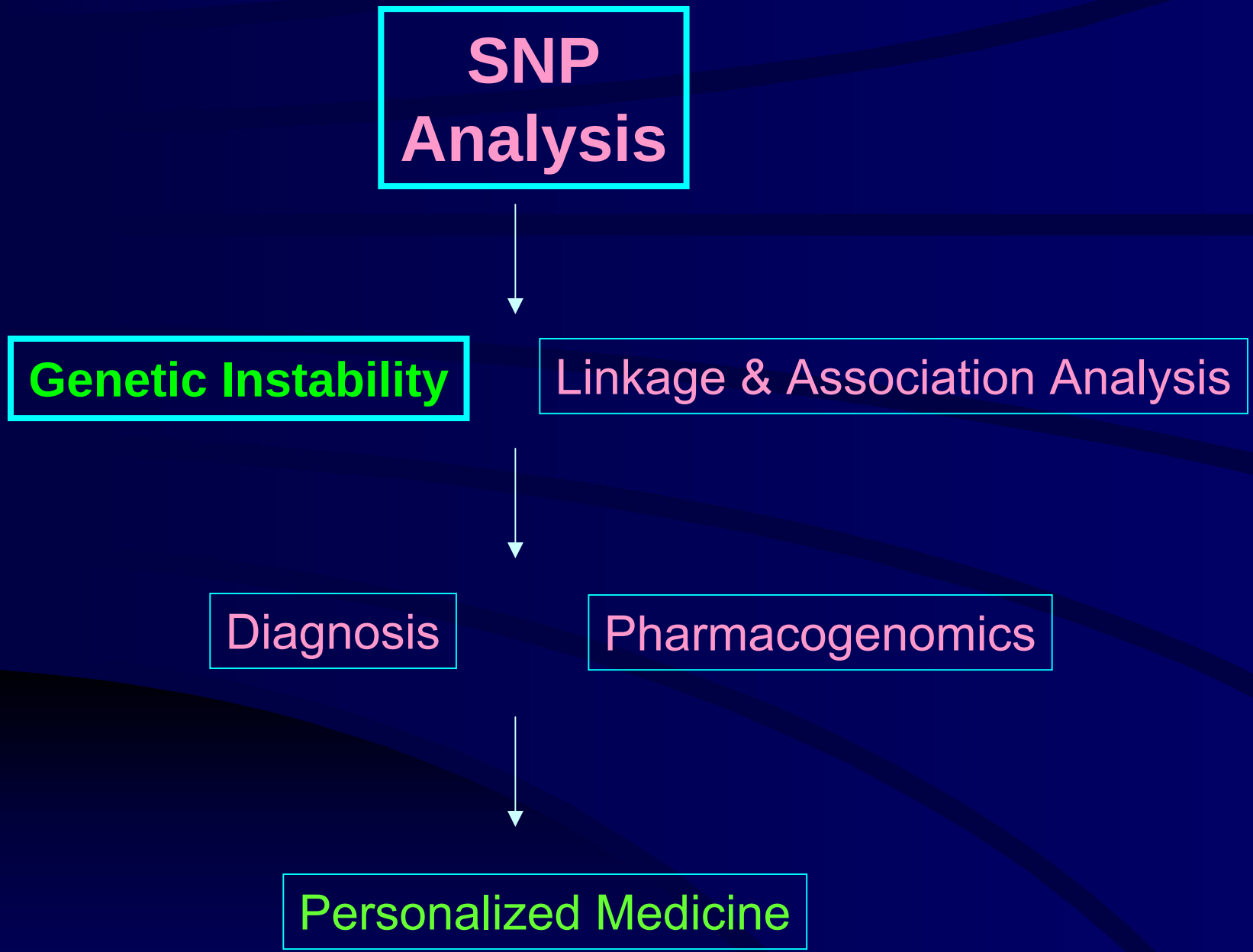
Effects of Gene Variability on Drug Response



Protein Affinity	Effect on Drug Response
Extreme	Undetectable
High	Detectable
Medium	Detectable
Low	Detectable
Negligible	None

GENAISSANCE PHARMACEUTICALS

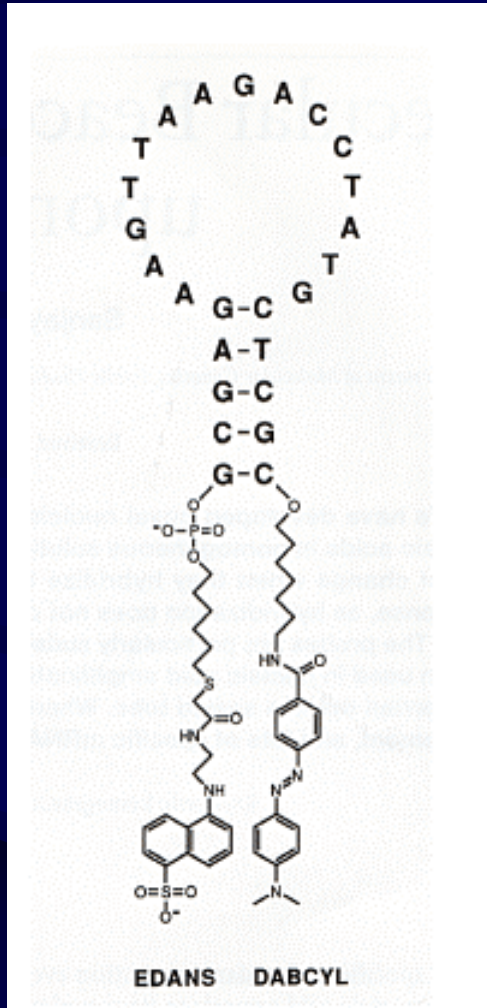
SNP Applications


SNP Analysis
Genetic Instability
Linkage & Association Analysis
Diagnosis
Pharmacogenomics
Personalized Medicine

Digital SNP analysis

→ detect genetic instability of cancer

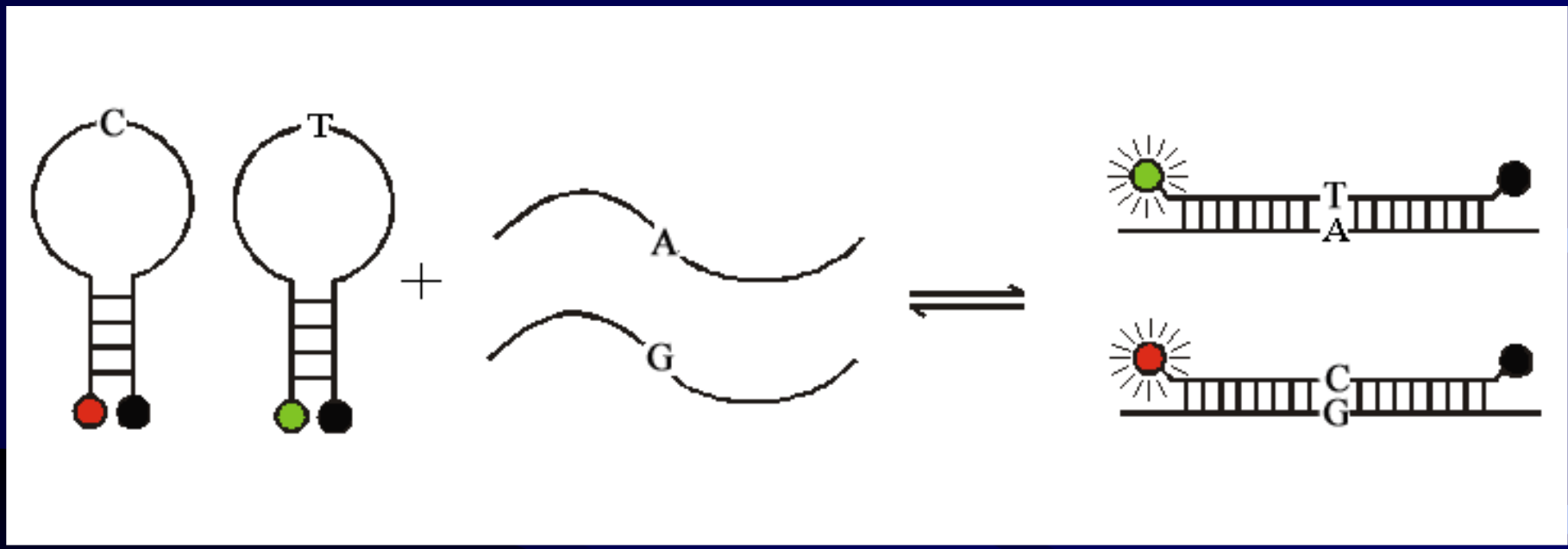
Molecular Beacons : Structure



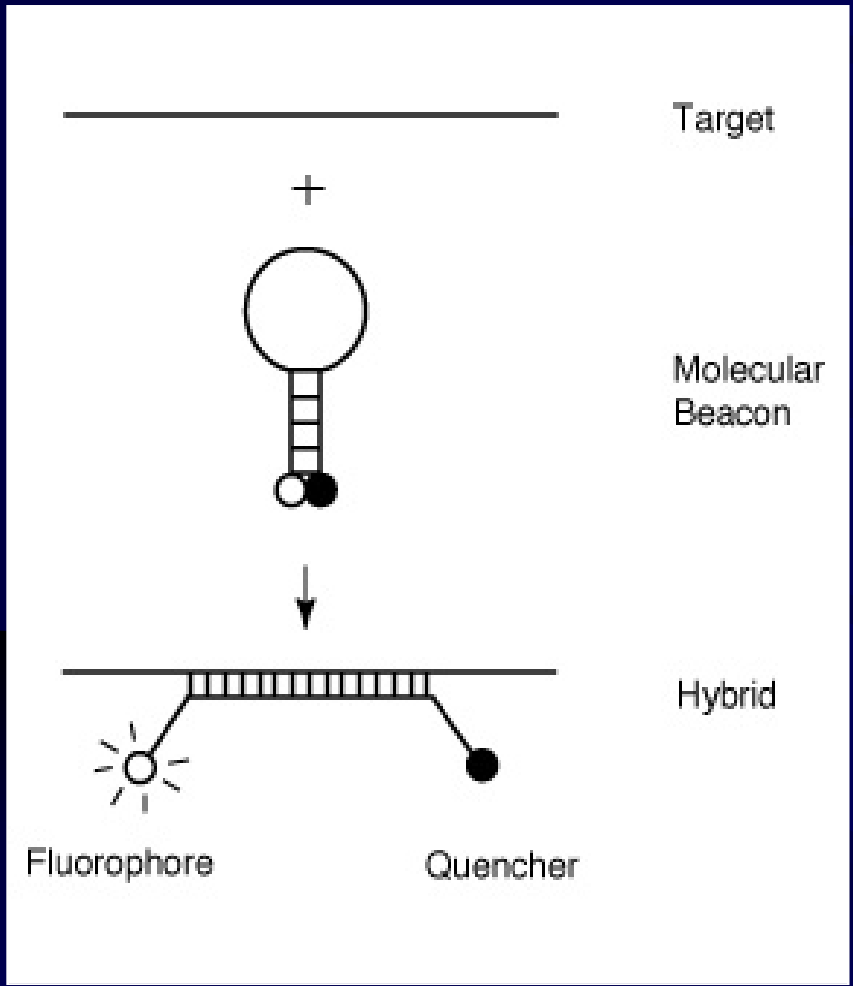
fluorophore quencher

Sanjay Tyagi and Fred Kramer (1996)
Nature Biotechnology 14: 303-308

SNP-specific Molecular Beacon



How do Molecular Beacons work?



- When hybridized with complementary target sequence, the stem structure opens.

- The fluorophore is removed from the quencher, and fluorescent signal can be detected.

DNA diluted to 384-well plate

↓

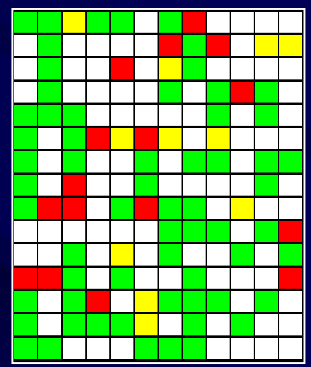
DNA templates become *all-or-none*
in each well of 384-well plate

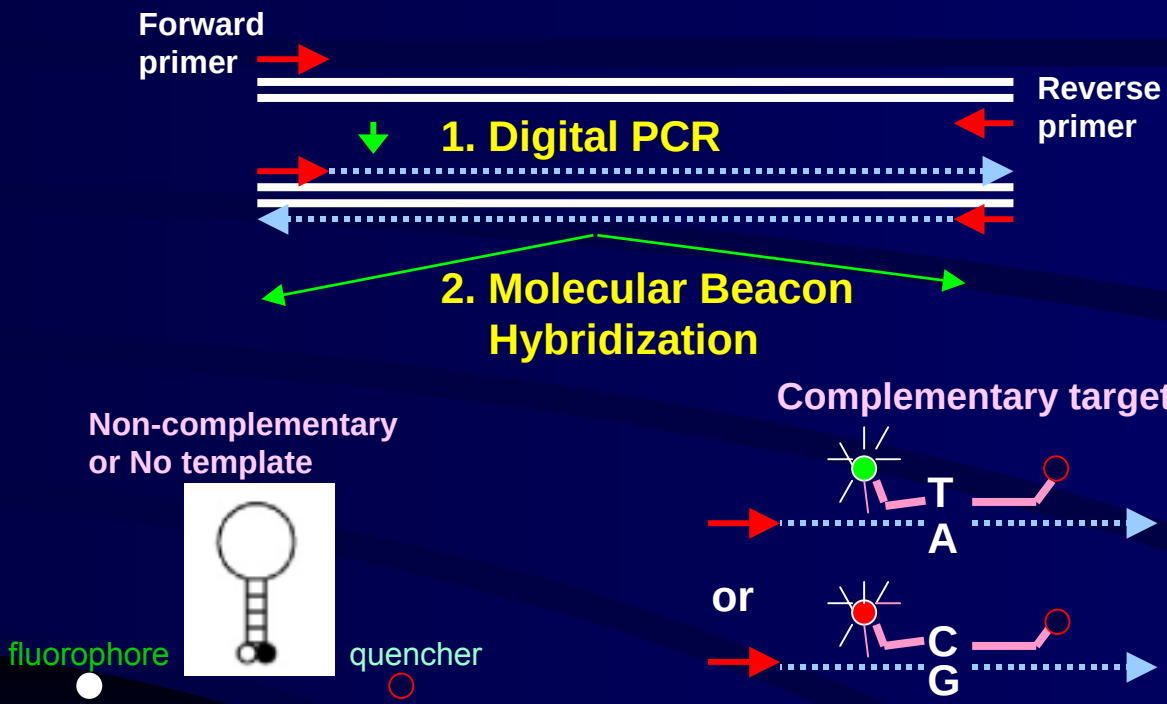
↓

PCR / hybrid with MB

↓

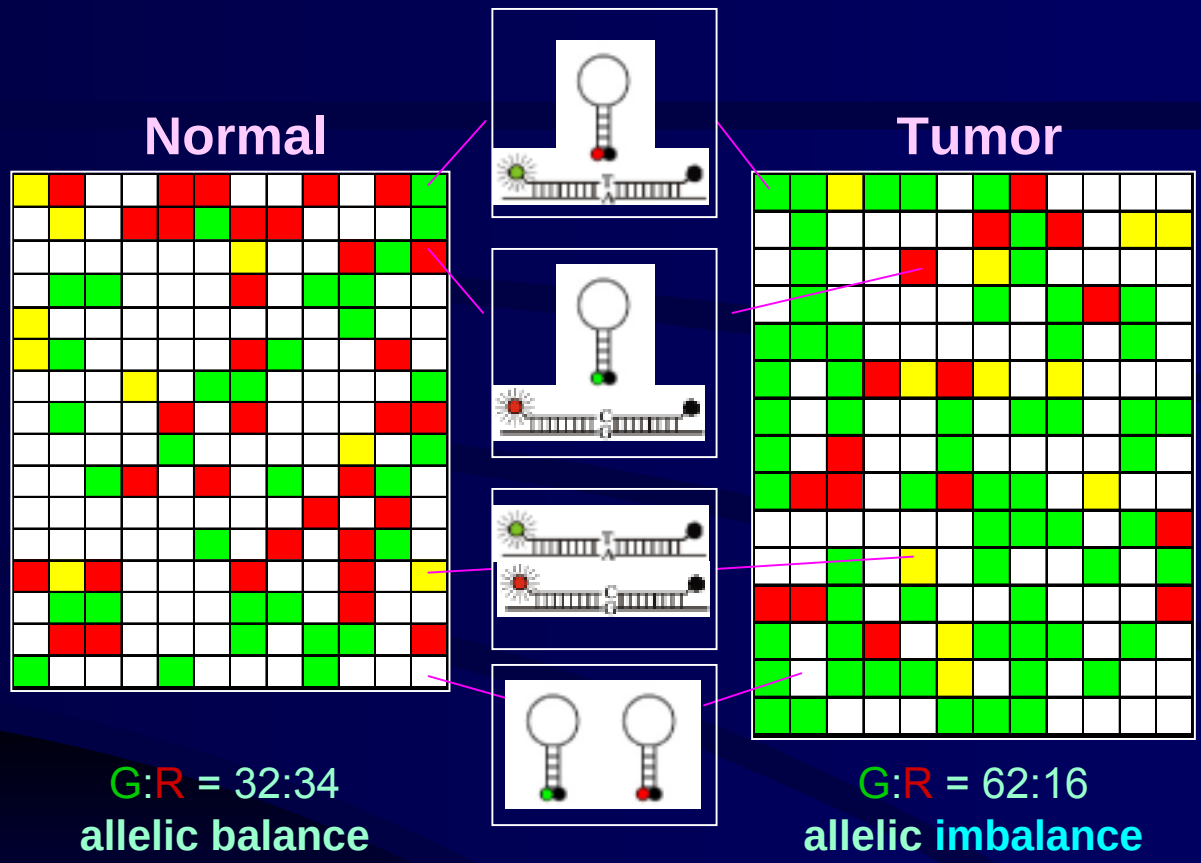
Digital PCR pattern *all-or-none*





Equal template size (<100 bp), no degradation

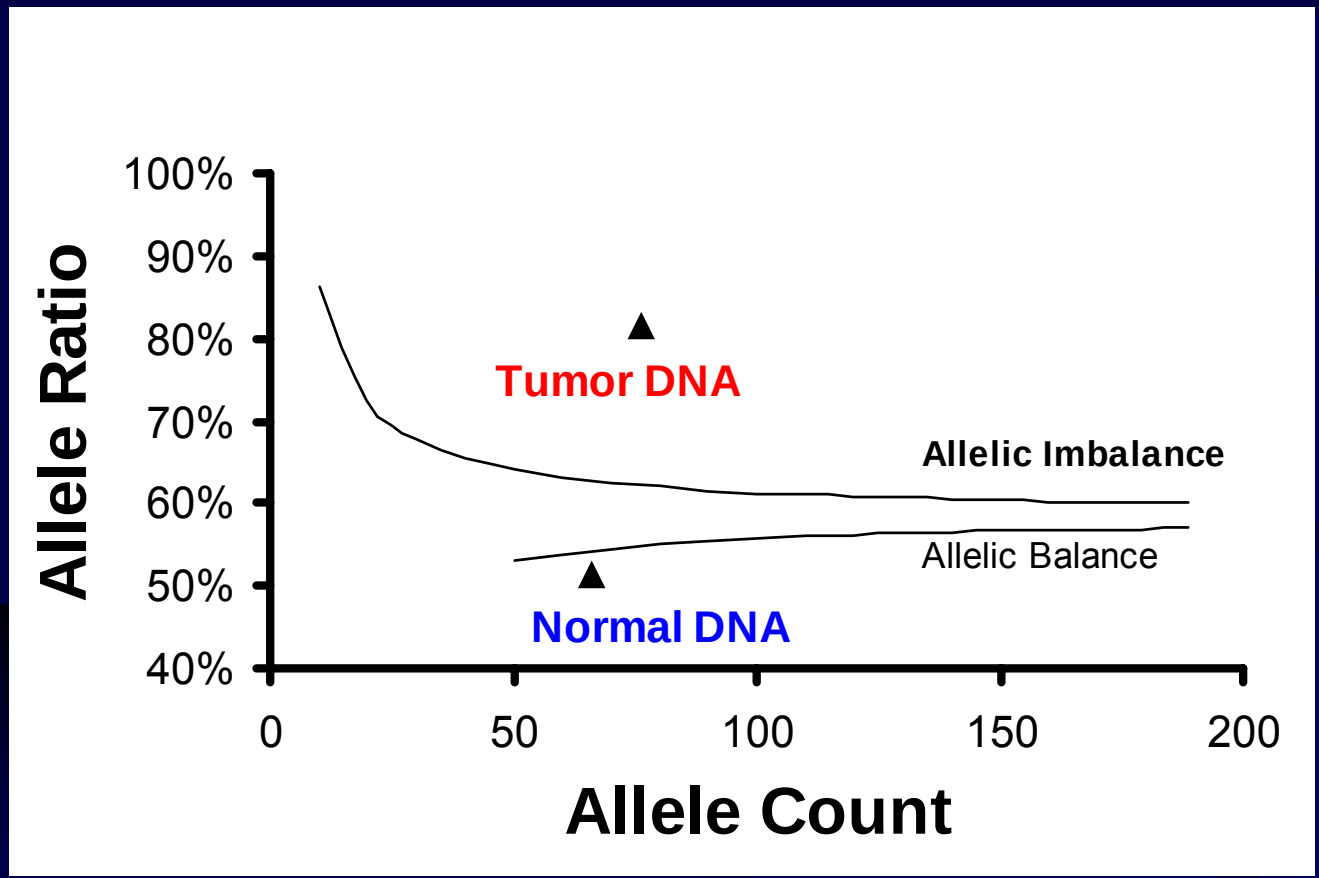
Assess genetic abnormality by counting alleles



Sequencial probability ratio test

(SPRT analysis)-

determines allelic imbalance



The advantages of Digital SNP analysis in assessing allelic imbalance

1. As compared with microsatellite markers the PCR products derived from the two SNP alleles at every locus *are of the same size* and therefore their analysis is *not biased by the preferential DNA degradation of larger alleles*.

2. The Digital SNP approach, which amplifies single allele templates in the PCR reaction, can precisely determine the number of alleles examined in each experiment.

Accordingly, SNP genotyping is "*digital*" involving the detection or absence of a specific allele *rather than "analogue*" as is microsatellite genotyping that measures the length of microsatellites.

3. A statistical method, *SPRT*, can be employed to conclude *whether allelic imbalance is present in the background DNA*.

Indeed, it has been shown that allelic imbalance can be demonstrated in *much higher percentage* in colorectal carcinomas using Digital SNP analysis than the traditional methods using microsatellite markers.

[Science. *244:* 207-11, 1989. & Nat Biotechnol. *19:* 78-81., 2001.]

Assessment of Plasma DNA Levels, Allelic Imbalance, and CA 125 as Diagnostic Tests for Cancer

Hsueh-Wei Chang, Shing M. Lee, Steven N. Goodman, Gad Singer, Sarah K. R. Cho, Lori J. Sokoll, Fredrick J. Montz, Richard Roden, Zhen Zhang, Daniel W. Chan, Robert J. Kurman, Ie-Ming Shih

J National Cancer Institute 94 (22), 1697-1703. (2002)

Allelic status in plasma DNA of ovarian cancer patients

Patient No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Stage	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3
Chromosome 1p / SNP 8118																											
Chromosome 5q / SNP 1756																											
Chromosome 7p / SNP 273																											
Chromosome 8p / SNP 1085																											
Chromosome 12q / SNP 852																											
Chromosome 15q / SNP 1861																											
Chromosome 17p / SNP p53																											
Chromosome 18p / SNP 331																											
Plasma DNA (ng/ml)	18	266	65	199	38	3	1309	264	15	8	17	186	23	42	61	92	210	26	6	10	94	746	36	28	131	8	110
Fraction of tumor DNA	0.3	N/A	N/A	0.5	0.3	0.4	0.3	0.4	0.3	0.3	0.5	0.3	0.4	N/A	N/A	0.7	0.7	0.4	N/A	0.5	0.3	0.7	0.3	0.3	0.3	0.3	0.9
CA125 (U/ml)	X	167	X	18	25	X	580	11	19	15	46	28	115	8	X	1260	76	36	X	220	875	X	9	30	270	16	3185

Patient No.	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
Stage	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	4	4
Chromosome 1p / SNP 8118																											
Chromosome 5q / SNP 1756																											
Chromosome 7p / SNP 273																											
Chromosome 8p / SNP 1085																											
Chromosome 12q / SNP 852																											
Chromosome 15q / SNP 1861																											
Chromosome 17p / SNP p53																											
Chromosome 18p / SNP 331																											
Plasma DNA (ng/ml)	2	146	38	3	67	149	5	338	491	103	92	147	65	285	20	538	39	25	40	14	78	35	895	2481	18	504	1190
Fraction of tumor DNA	0.4	0.4	0.4	N/A	0.5	0.7	0.4	0.6	0.4	0.6	0.6	0.7	0.7	0.7	0.4	0.9	0.3	0.5	0.3	0.3	0.4	0.6	0.7	0.9	0.4	0.7	0.7
CA125 (U/ml)	8	84	X	29	600	75	15	250	377	66	979	13	979	209	19	231	643	319	788	X	X	580	365	66	38	2397	304

Allelic balance

Allelic imbalance

Undetermined

Homozygous (non-informative)

Vol. 8, 2580–2585, August 2002 Clinical Cancer Research

Detection of Allelic Imbalance in Ascitic Supernatant by Digital Single Nucleotide Polymorphism Analysis[1]

Hsueh-Wei Chang, Syed Z. Ali, Sarah K. R. Cho, Robert J. Kurman, and Ie-Ming Shih[2]

Departments of Pathology [H-W. C., S. Z. A., S. K. R. C., R. J. K., I-M. S.] and Gynecology and Obstetrics [R. J. K.], The Johns Hopkins University School of Medicine, Baltimore, Maryland 21231

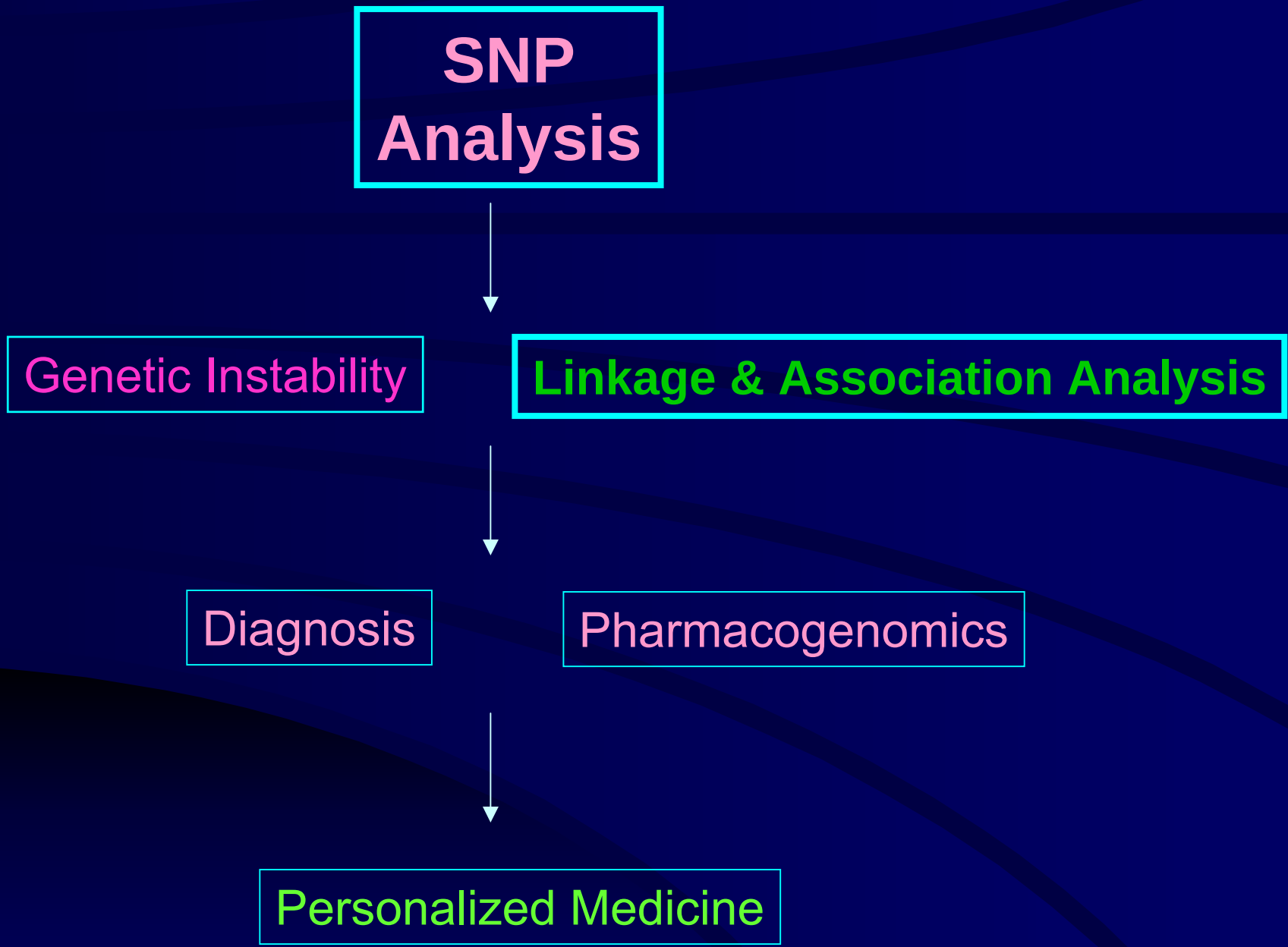
Diverse Tumorigenic Pathways in Ovarian Serous Carcinoma

Am J Pathol 2002, 160:1223–1228

Gad Singer*, Robert J. Kurman,*†
Hsueh-Wei Chang,* Sarah K.R. Cho,* and
Ie-Ming Shih*

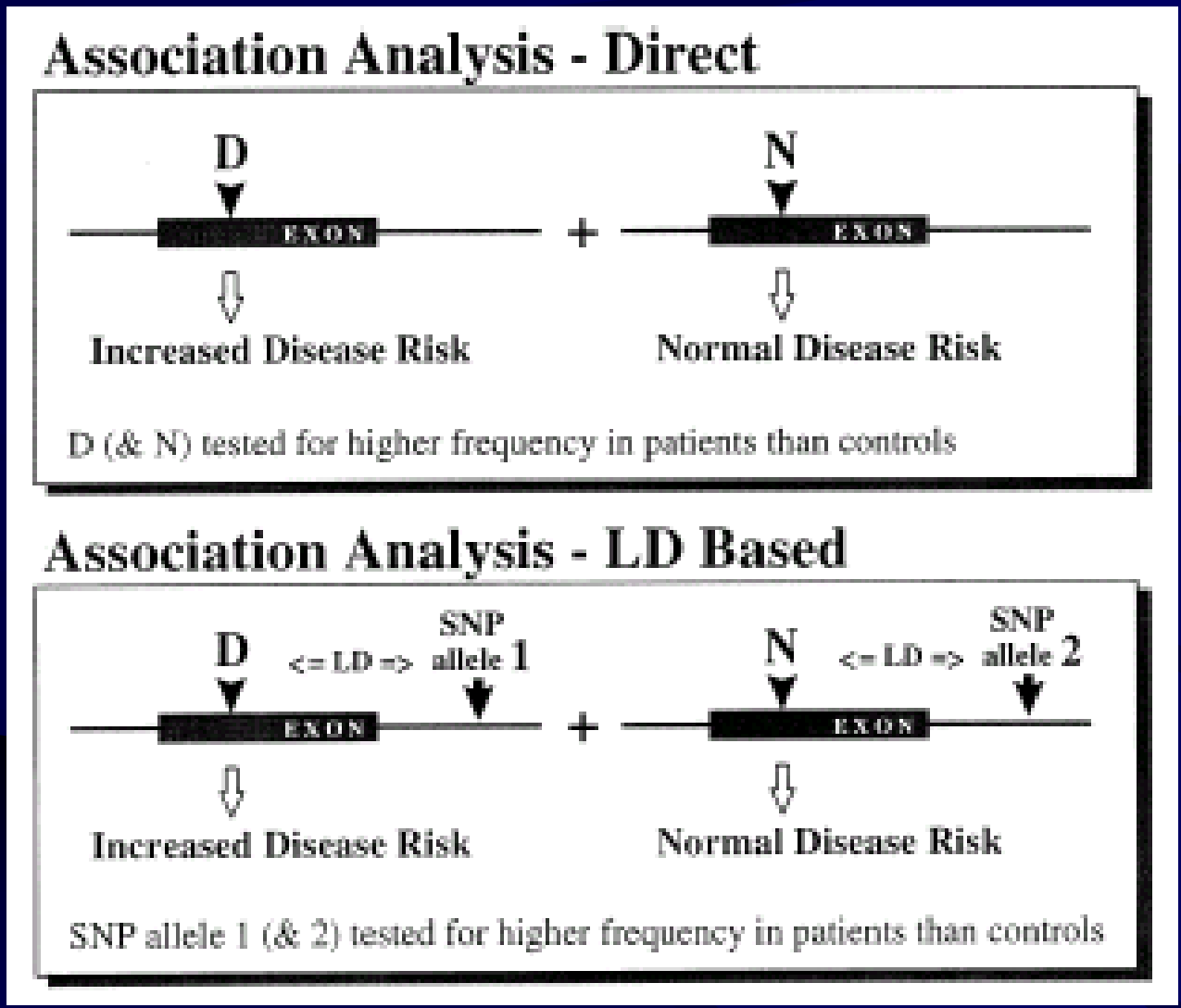
Conclusion I

Digital SNP analysis can count the allelic imbalance in samples from plasma, ascites, paraffin …etc. in the presence of background (normal) DNA.


SNP Analysis
Genetic Instability
Linkage & Association Analysis
Diagnosis
Pharmacogenomics
Personalized Medicine

Strategies for association analysis

Gene 234 (1999) 177–186



LD (Linkage disequilibrium)

Some Diseases Involve Many Genes

- There are a number of classic "genetic diseases" caused by mutations of a single gene
  - Huntington's, Cystic Fibrosis, PKU, etc.
- There are also many diseases that are the result of the interactions of many genes:
  - asthma, heart disease, cancer
- Each of these genes may be considered to be a **risk factor** for the disease.
- Groups of SNP markers may be associated with a disease without determining mechanism

Direct Medical Applications

- Diagnosis
  - Type of cancer
  - Aggressive or benign?
- Monitor treatment outcome
  - Is a treatment having the desired effect on the target tissue?

People React Differently to Drugs

- Side effects
- Effectiveness
- There are genes that control these reactions
- SNP markers can be used to identify these genes

from "One Drug Fits All" to personalized therapy

SNP or DNA-based marker discovery

Expression of P-glycoprotein and Genotype (MDR1 3435)

P-gp Genotype	P-gp Expression
T T	Reduced
C T	Normal
C C	Increased

Pharmacogenomics and Asthma

Drug Category	Effector Variability
β-agonists	Receptor at codons 16, 27
Corticosteroids	IL-4 polymorphisms
Leukotriene inhibitors	ALOX5 core promoter expression

Hypertension: The α-Adducin Gene Variant and the Rate of MI or CVA

Genotype	Odds Ratio of Clinical Events
Wild Type (n=653)	1.09 (0.78-1.52)
Variant Allele (n=385)	0.49 (0.32-0.77)

Psaty, JAMA, 2002

Warfarin Toxicity and CYP2C9 Variability: Clinical Correlates

Warfarin maintenance dose (mg)

Wild type: *1*1	**5.63**
Variant: *1*2	**4.88**
Variant: *1*3	**3.32**

Higashi, JAMA, 2002

Functional SNPs in the lymphotoxin-α gene that are associated with susceptibility to myocardial infarction

Kouichi Ozaki[1], Yozo Ohnishi[1], Aritoshi Iida[2], Akihiko Sekine[2], Ryo Yamada[3], Tatsuhiko Tsunoda[4], Hiroshi Sato[5], Hideyuki Sato[5], Masatsugu Hori[5], Yusuke Nakamura[2,6] & Toshihiro Tanaka[1]

Nat Genet. 2002 Dec;32(4):650-4.

European Heart Journal (2001) **22**, 845–848
doi:10.1053/euhj.2000.2400, available online at http://www.idealibrary.com on IDEAL®

The human G-protein β3 subunit C825T polymorphism is associated with coronary artery vasoconstriction

A. Meirhaeghe[1], C. Bauters[1,2], N. Helbecque[1], M. Hamon[3], E. McFadden[2], J.-M. Lablanche[2], M. Bertrand[2] and P. Amouyel[1,4]

Comparison of genotype and allele frequencies of the C825T polymorphism of GNB3 between hypertensive and normotensive groups

	Genotypes (frequency)		
Group	*CC*	*CT*	*TT*
Hypertensive	27 (0.25)	71 (0.65)	12 (0.11)
Normotensive	101 (0.53)	82 (0.43)	6 (0.03)

Promoter SNP

J Mol Med (2001) 79:732–737
DOI 10.1007/s001090100265

ORIGINAL ARTICLE

Laura Viitanen · Jussi Pihlajamäki · Raija Miettinen
Päivi Kärkkäinen · Ilkka Vauhkonen · Pirjo Halonen
Anu Kareinen · Seppo Lehto · Markku Laakso

Apolipoprotein E gene promoter (–219G/T) polymorphism is associated with premature coronary heart disease

[CANCER RESEARCH 62, 4992–4995, September 1, 2002]

A Novel Polymorphism in Human Cytosine DNA-Methyltransferase-3B Promoter Is Associated with an Increased Risk of Lung Cancer[1]

Hongbing Shen,[2] Luo Wang,[2] Margaret R. Spitz, Waun K. Hong, Li Mao, and Qingyi Wei[3]

Departments of Epidemiology [H. S., M. R. S., Q. W.] and Thoracic and Head and Neck Medical Oncology [L. W., W. K. H., L. M.], The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030

[CANCER RESEARCH 62, 152–155, January 1, 2002]

The *XRCC1* Arg399Gln Polymorphism, Sunburn, and Non-melanoma Skin Cancer: Evidence of Gene-Environment Interaction[1]

Heather H. Nelson,[2] Karl T. Kelsey, Leila A. Mott, and Margaret R. Karagas

Department of Cancer Cell Biology, Harvard School of Public Health, Boston, Massachusetts 02115 [H. H. N., K. T. K.], and Section of Biostatistics and Epidemiology, Department of Community and Family Medicine, Dartmouth Medical School, Lebanon, New Hampshire 03756 [L. A. M., M. R. K.]

Genotype Analysis Using Human Hair Shaft

Cancer Epidemiology, Biomarkers & Prevention *11: 925–929, 2002*

Hsueh-Wei Chang, Ching-Yu Yen, Shyun-Yui Liu, Gad Singer, and Ie-Ming Shih*

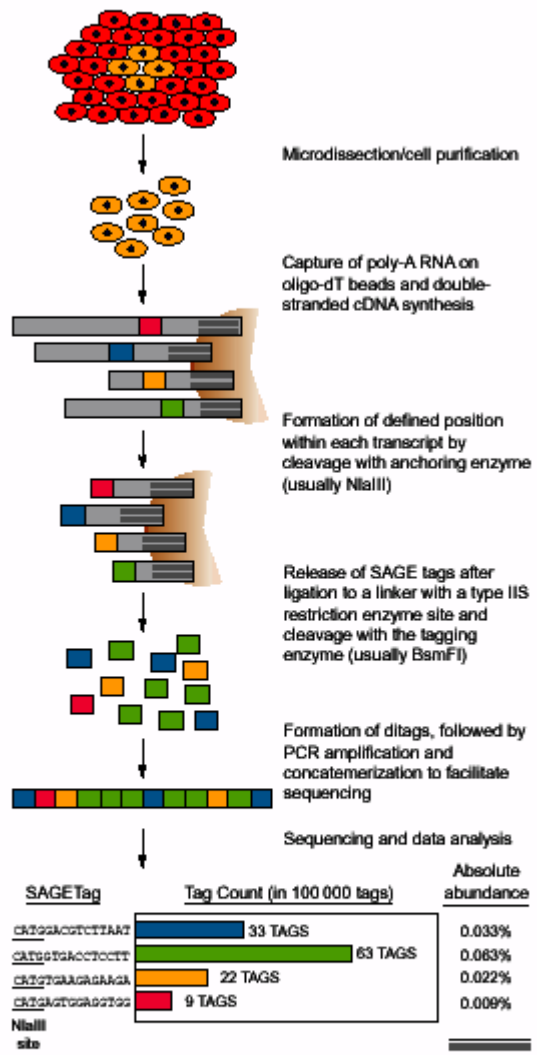
mRNA and protein marker discovery

Gene Expression (phenotype) Profiling

- Sequence bulk cDNAs from different tissues
  - SAGE (sequence short tags from cDNAs)
  - Microarrays

Data mining from SAGE (Serial Analysis of Gene Expression)

FIGURE 1. Schematic of the SAGE method as applied to small cell populations



Cells of interest are microdissected or purified by affinity selection. Polyâ RNA is directly captured from cell lysates using oligo-dT coated beads[44] and converted to cDNA. A frequently cutting anchoring enzyme, usually NlaIII, is used to cleave cDNA molecules, leaving the 3′ end of the cDNA attached to the beads. Linkers are ligated to the immobilized cDNA fragments. These contain a site for a type IIs restriction enzyme (usually BsmFI), and the tagging enzyme is then used to release a short (15 bp) tag from the cDNA. These tags are ligated tail to tail to form ditags, which can be amplified by PCR, and then concatemerized and cloned. Sequencing of concatemer clones reveals the identity and abundance of each tag. Absolute abundances are calculated by dividing the observed abundance of any tag by the total number of tags analysed.

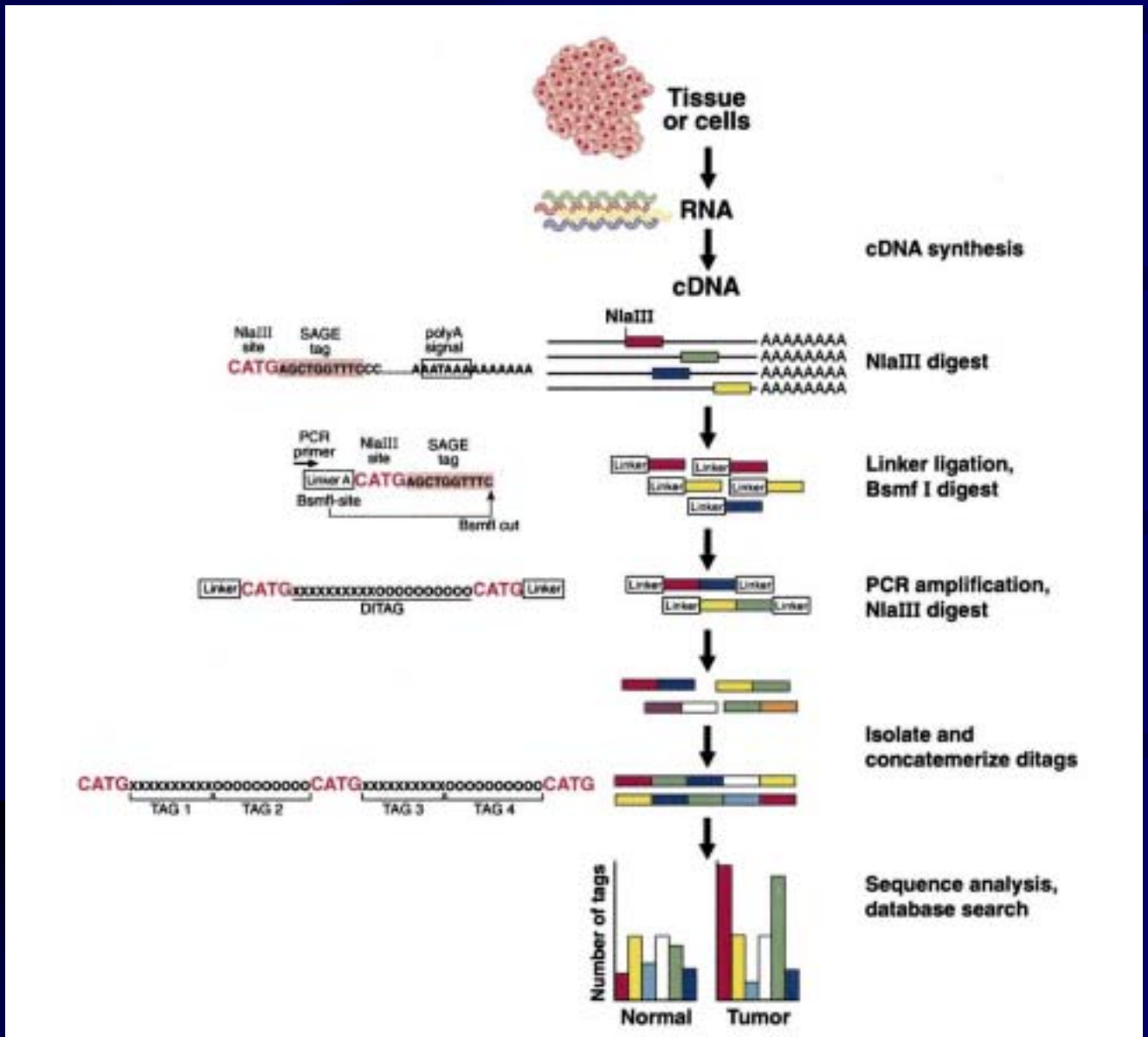


Fig 1. Brief outline of SAGE library generation and analysis. mRNA purified from cells are converted to double-stranded cDNA, and SAGE tags are derived, concatemerized, and subcloned by sequential enzymatic steps. A specific example is depicted on the left-hand side.

Serial Analysis of Gene Expression
Tag to Gene Mapping
SAGEmap

NCBI

SAGE: Measuring Gene Expression

Information
Home
About SAGE
About website & data
About mappings
About tools
Data submissions
Extraction stats
Resources
Virtual Northern
xProfiler
xProfiler example
Tag to gene mapper
Gene to tag mapper
Library Search
Downloads
Library info
Mappings
Sequences
Tag & ditag lists

Serial Analysis of Gene Expression, or SAGE, is an experimental technique designed to gain a quantitative measure of gene expression.
The SAGE technique itself includes several steps utilizing molecular biological, DNA sequencing and bioinformatics techniques.
These steps have been used to produce 9 or 10 base "tags", which are then, in some manner, assigned gene descriptions.
For experimental reasons, these tags are immediately adjacent to the 3' end of the 3'-most NlaIII restriction site in cDNA sequences.

Online Data Analysis

Information on SAGE libraries and access to the data produced from these libraries can be reviewed and analyzed with several online tools via the xProfiler Virtual Northern

Differential expression analysis queries

Colon cancer vs. normal colon

http://www.ncbi.nlm.nih.gov/SAGE

NCBI

Serial Analysis of Gene Expression Tag to Gene Mapping SAGEmap

CGAP UniGene OMIM PubMed Entrez BLAST

SAGEmap xProfiler

Information
Home
About SAGE
About website & data
About mappings
About tools
Data submissions
Extraction stats

Resources
Virtual Northern
xProfiler
xProfiler example
Tag to gene mapper
Gene to tag mapper
Library Search

Downloads
Library info
Mappings
Sequences
Tag & ditag lists

FTP site

PubMed search

xProfiler:

1. Type in names for Groups A & B (optional)
2. Select libraries to put into Groups A & B below
3. Alter fold difference factor (default 2 fold)
4. Alter coefficient of variance cutoffs (default disabled)
5. Press Calculate

Group A name:
Group B name:

Factor: 2.0 X difference

0 % Coefficient of variance cutoffs 0 %

Limited by tags:

<TD WIDTH="60" VALIGN="t

A B

WHOLE BLOOD

SAGE Duke leukocyte (48169 tags)
Bulk tissue, blood, normal human adult leukocyte total RNA

BOWEL

SAGE HCT116 (60322 tags)
Cell line, colon, cell line derived from colorectal carcinoma, ATCC: CCL-247, Mutation in the Ras gene, codon 13, Wild type p53, RER+

SAGE Caco 2 (61601 tags)
Cell line, colon, colorectal carcinoma cell line (RER-), ATCC: HTB-37, 72 year old male

SAGE SW837 (60986 tags)
Cell line, colon, cancer cell line, Mismatch proficient(RER-) with a mutant p53(248arg -> trp) and a mutant APC

SAGE RKO (52064 tags)
Cell line, colon, cancer cell line, Wild type p53, RER+

SAGE NC1 (50115 tags)
Bulk tissue, normal colonic epithelium

SAGE NC2 (49552 tags)

Quick browser:

Click on an organ in image to show a new list of SAGE libraries

Use the full-featured Library Browser to choose from a different list of libraries

NCBI

Serial Analysis of Gene Expression
Tag to Gene Mapping
SAGEmap

CGAP UniGene OMIM PubMed Entrez BLAST

SAGE Data Analysis

For this query, there are **57381** unique SAGE tags of which the 100 most likely different by greater than 2 fold are shown. For each of these tags, the probability that there is greater than a **2 fold difference in expression levels** between Groups A and B is given.

To download the entire list, see the bottom of this page.

Group A: Colon cancer (total tags: 106641)

SAGE_Tu102 : Colon, primary tumor (*total tags: 57636*)
SAGE_Tu98 : Colon, primary tumor (*total tags: 49005*)

Group B: Normal colon (total tags: 99667)

SAGE_NC1 : Normal colonic epithelium (*total tags: 50115*)
SAGE_NC2 : Normal colonic epithelium (*total tags: 49552*)

Color = **RED** if expression of tag in **Group A > GroupB**

Color = **GREEN** if expression of tag in **Group B > Group A**

#	*SAGE tag*	*UniGene Id*	*Gene description*	*A:B*	*Grp A (CoV)*	*Grp B (CoV)*	*A:B >2x*
1	CATAGGTTTA	Hs.1650	solute carrier family 26, member 3	AB	7 (154%)	105 (36%)	100%
		Hs.323430	EST				
2	CCCATCGTCC	N/A	**WARNING**: Tag matches mitochondrial DNA	AB	1236 (131%)	2672 (8%)	100%
3	ATCGTGGCGG	Hs.5372	claudin 4	AB	33 (66%)	285 (49%)	100%
4	CTGGCCCTCG	Hs.166184	intersectin 2	AB	30 (21%)	238 (79%)	100%
		Hs.350470	trefoil factor 1 (breast cancer, estrogen-inducible sequence expressed in)				
		Hs.7720	dynein, cytoplasmic, heavy polypeptide 1				
5	ATACTCCACT	Hs.778	guanylate cyclase activator 2A (guanylin)	AB	14 (93%)	142 (21%)	100%
6	TAAATTGCAA	Hs.56205	insulin induced gene 1	AB	9 (50%)	162 (38%)	100%
		Hs.84905	cytokeratin 20				
7	GACATCAAGT	Hs.182265	keratin 19	AB	65 (80%)	285 (54%)	100%
8	GTCATCACCA	Hs.107382	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 37	AB	0 (0%)	57 (31%)	100%
		Hs.257045	hypothetical protein dJ726C3.2				
		Hs.32966	guanylate cyclase activator 2B				

NCBI Serial Analysis of Gene Expression Tag to Gene Mapping SAGEmap

CGAP UniGene OMIM PubMed Entrez BLAST

SAGE Tag to Gene Mapping

SAGEtag (10 bases): ATCGTGGCGG | Homo sapiens | NlaIII | Submit

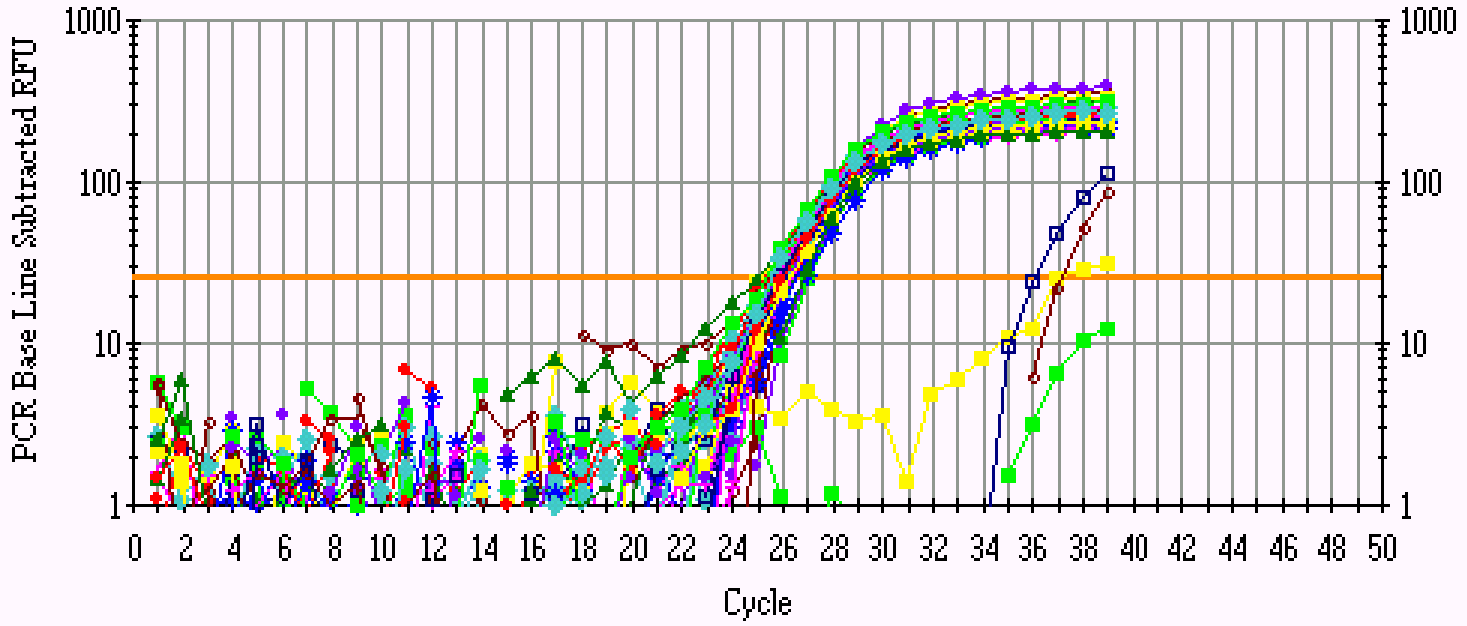
Reliable UniGene clusters matched to this tag:

Hs.5372 : **claudin 4**

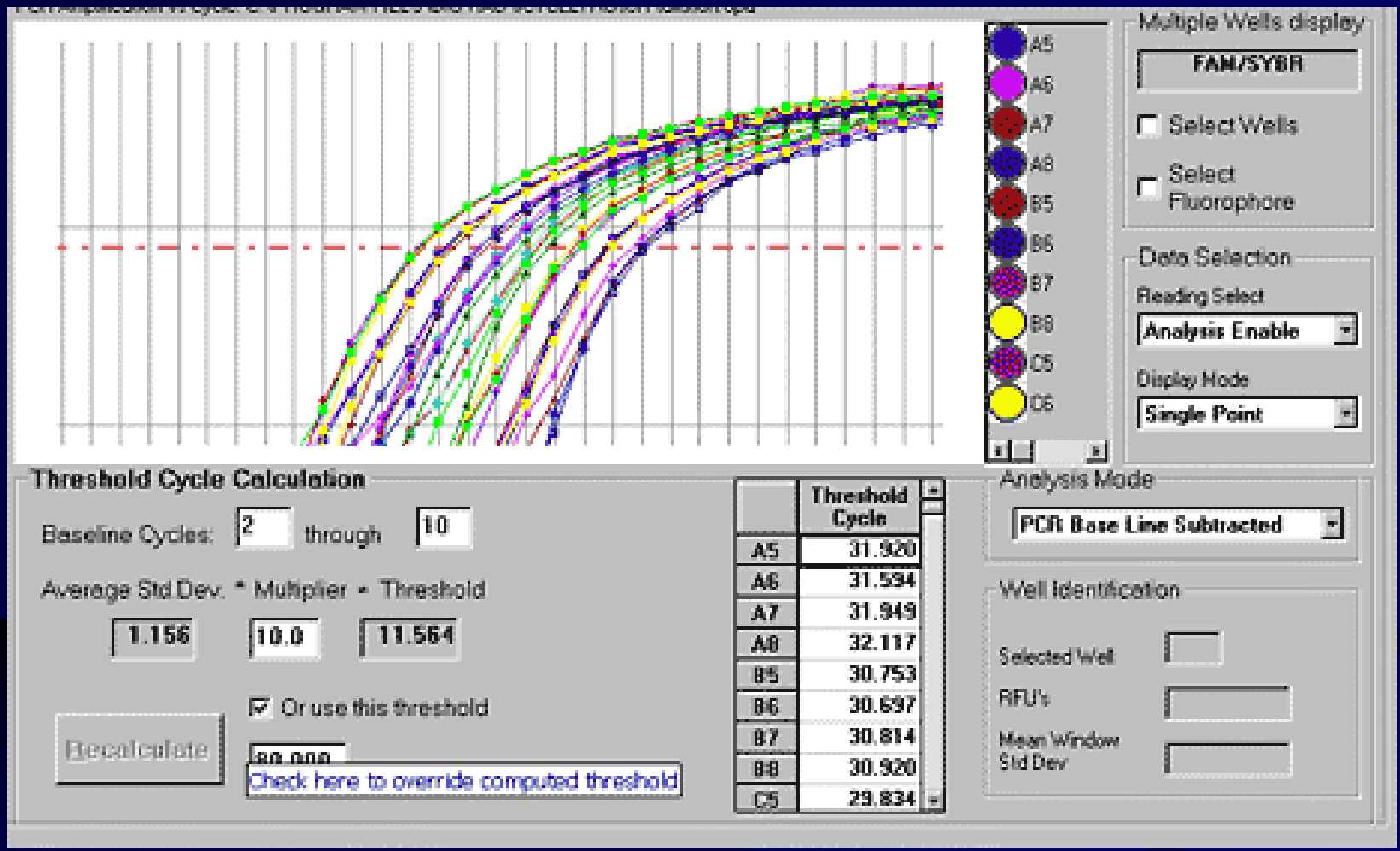
SAGE library data for this tag:

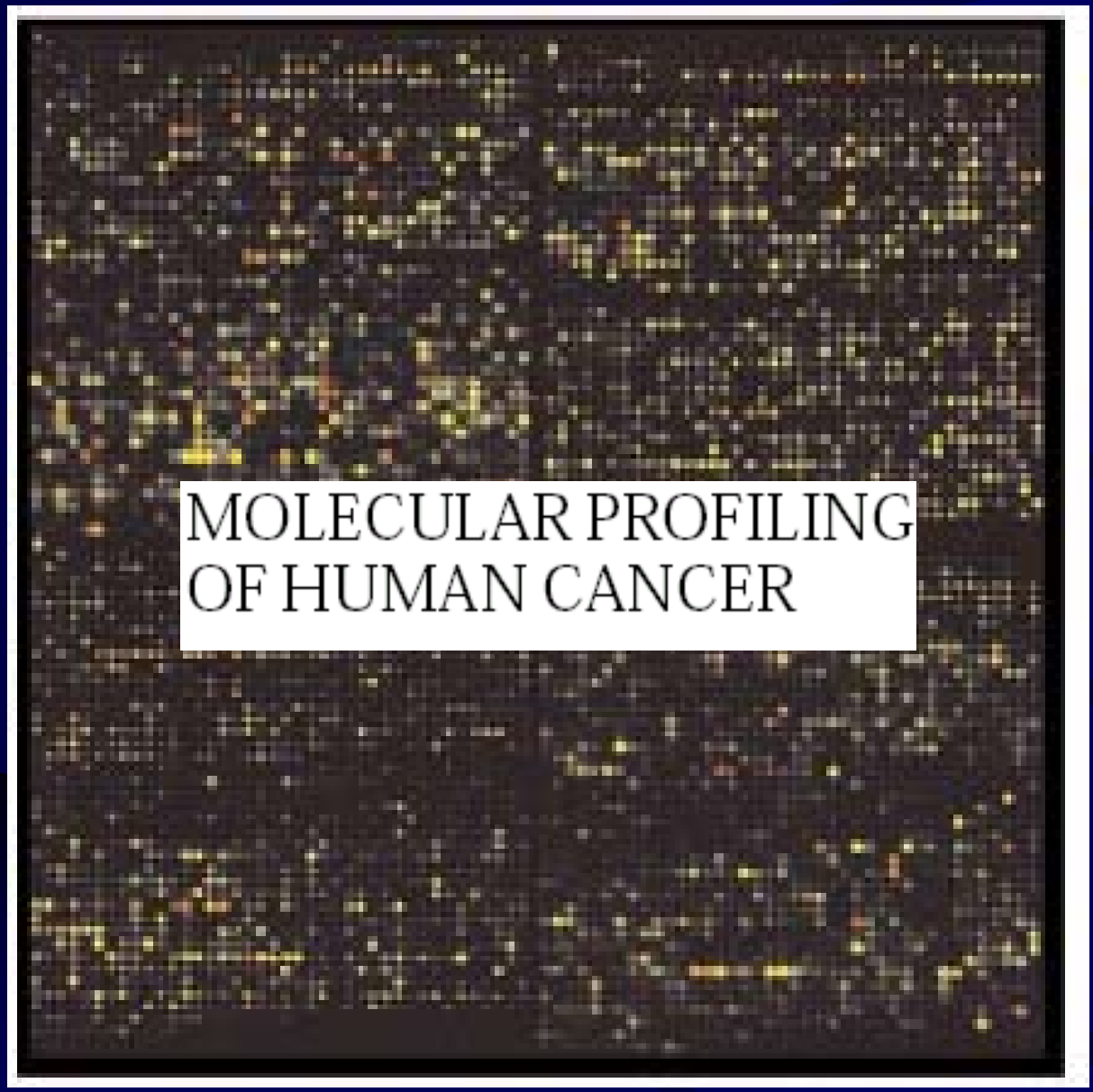
Library name	*Tags per million*	*Tag counts*	*Total tags*
SAGE HCT116	198	12	60322
SAGE Caco 2	340	21	61601
SAGE Chen LNCaP	144	9	62267
SAGE Chen LNCaP no-DHT	123	8	64631
SAGE Chen Normal Pr	75	5	66193
SAGE Chen Tumor Pr	87	6	68384
SAGE CAPAN1	290	11	37926
SAGE CAPAN2	706	27	38240
SAGE HS766T	539	17	31506
SAGE Panc1	281	7	24879
SAGE Duke H54 lacZ	59	4	67101
SAGE SW837	3115	190	60986
SAGE CPDR LNCaP-C	144	6	41590
SAGE CPDR LNCaP-T	45	2	44122
SAGE PR317 prostate tumor	76	5	65109
SAGE NHA(5th)	19	1	52196
SAGE NC1	3851	193	50115
SAGE NC2	1856	92	49552
SAGE Panc 91-16113	648	22	33941
SAGE Panc 96-6252	615	22	35745
SAGE OVCA432-2	699	2	2861
SAGE OV1063-3	282	11	38938
SAGE Tu102	173	10	57636
SAGE Tu98	469	23	49005
SAGE SciencePark MCF7 control 3h	338	2	5903
SAGE SciencePark MCF7 Control 0h	16	1	61079

Internal control



Expression of same gene in different samples




MOLECULAR PROFILING
OF HUMAN CANCER

Summary

Personalized medicine simply means the prescription of specific therapeutics best suited for an individual based on pharmacogenetic and pharmacogenomic information.

Several technologies are used including single nucleotide polymorphism genotyping. haplotyping, gene expression studies by biochip/microarrays, SAGE(serial analysis of gene expression) and proteomics.

Molecular diagnostics will play an important role in the development of personalized medicine, in which therapy and diagnosis will be integrated.

Personalized therapy is financially feasible, as it will reduce the costs of drug development by shortening the drug development cycle.

Personalized medicine is anticipated to be an acceptable part of medical practice by the year **2010**.

The PATH TO personalized medicine - Archive - Microsoft Internet Explorer

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)

上一頁 下一頁 停止 重新整理 首頁 搜尋 我的最愛 記錄 郵件 列印 編輯 討論

網址(D) C:\My Documents\wei\cg\bioinformatics\The PATH TO personalized medicine - Archive.htm 移至 連結

Bio-IT World

Technology for the Life Sciences SEARCH

THE NEW NEW PHARMACOGENOMICS

THE PATH TO PERSONALIZED MEDICINE

[PART TWO OF A TWO-PART SERIES]

The tactics have changed, sometimes dramatically, but hints of

完成 Internet

開始 Micr... Acro... 檔... The ... AM 03:30

The threat posed by personalised medicine

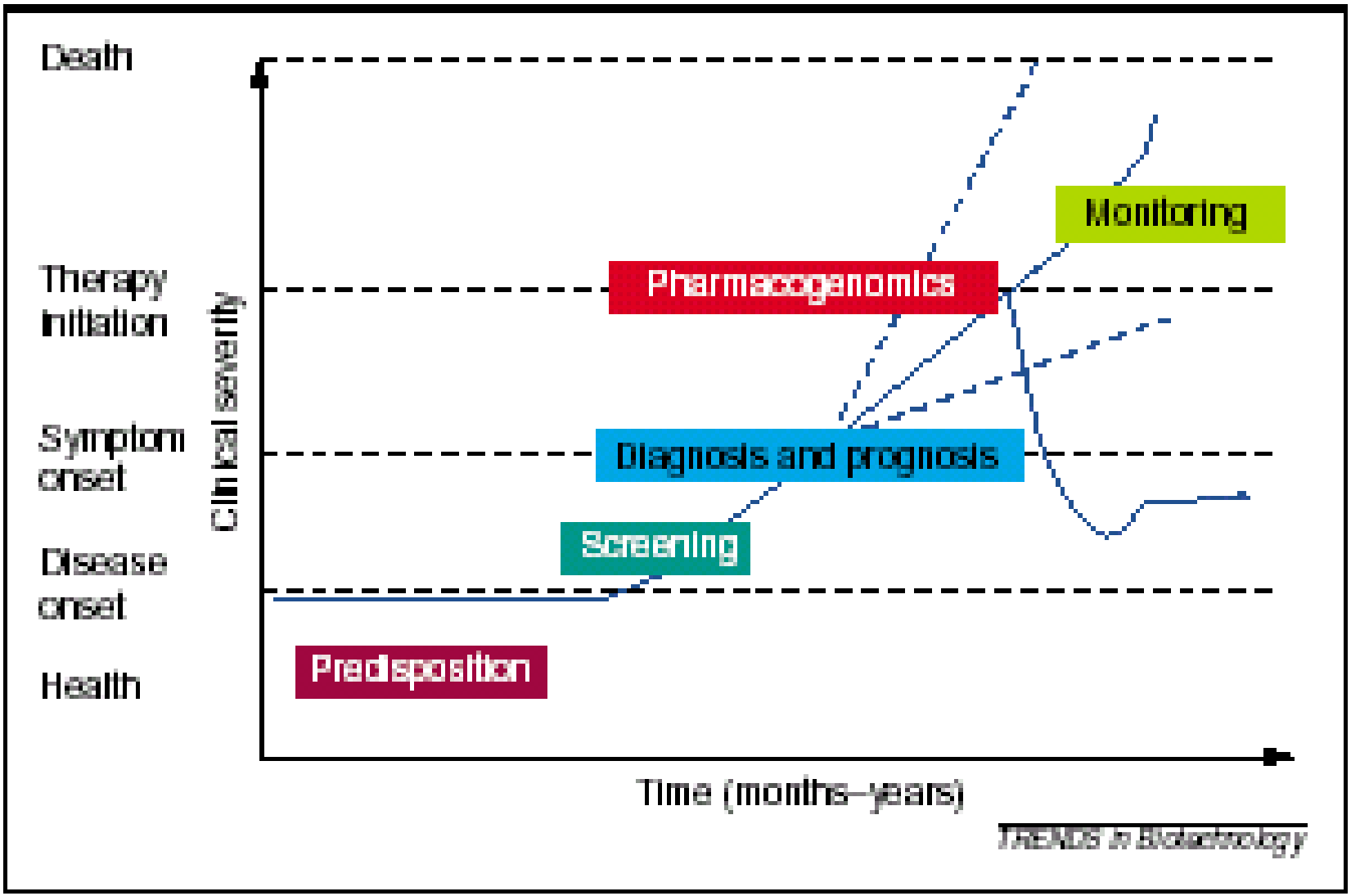
Despite the promise of personalised medicine for the pharmaceutical industry as well as patients, many companies are understandably wary of the impending changes in the way medicines will be discovered, developed and prescribed.

These changes could be severely disruptive to the business of the big drug companies, which rely on blockbuster products to drive their sales and earnings growths.

Think of the damage to sales if a $1 billion product is found to work in only the 20% of patients with the appropriate genetic profile, for instance.

Research, intervention and personalized medicine opportunities at stages of a hypothetical disease.

TRENDS in Biotechnology Vol.19 No.12 December 2001



- **Genetic Polymorphism:**

  A difference in DNA sequence among individuals, groups, or populations.

- **Genetic Mutation:**

  A change in the nucleotide sequence of a DNA molecule. Genetic mutations are a kind of genetic polymorphism.

  Single nucleotide polymorphisms (SNPs) are an abundant form of genome variation, distinguished from rare variations by a requirement for the least abundant allele to have a frequency of 1% or more.

The genotype-phenotype relation

- **Phenotype**: The observable properties of an individual as they have developed under the combined influences of the individual's genotype and the effects of environmental factors.

- **Genotype**: An exact description of the genetic constitution of an individual, with respect to a single trait or a larger set of traits (sequence of complete set of genes, both dominant and recessive - “SNP scoring”).

- The genotype-phenotype relation forms the basis and motivation for SNP research. If SNPs account for diversity in genotypes, then SNPs also can be mapped to account for diversity in phenotypes.

Genotype and Haplotype

- In the most basic sense, a haplotype is a "haploid genotype".

- **Haplotype:** particular pattern of sequential SNPs (or alleles) found on a single chromosome. These SNPs tend to be inherited together over time.

Gene Expression (phenotype) Profiling

- Sequence bulk cDNAs from different tissues
  - NCBI CGAP website allows "digital differential display"
- SAGE (sequence short tags from cDNAs)
- Microarrays

Impact on Bioinformatics

Genomics produces high-throughput, high-quality data, and bioinformatics provides the analysis and interpretation of these massive data sets.

Personalized medicine, the use of marker-assisted diagnosis and targeted therapies derived from an *individual's molecular profile*, will impact the way drugs are developed and medicine is practiced.

Summary
[personalized medicine & bioinformatics-hwchang]

Personalized medicine simply means the prescription of specific therapeutics best suited for an individual based on pharmacogenetic and pharmacogenomic information.

Several technologies are used including single nucleotide polymorphism genotyping. haplotyping, gene expression studies by biochip/microarrays, SAGE(serial analysis of gene expression) and proteomics.

Molecular diagnostics will play an important role in the development of personalized medicine, in which therapy and diagnosis will be integrated.

Personalized therapy is financially feasible, as it will reduce the costs of drug development by shortening the drug development cycle.

Personalized medicine is anticipated to be an acceptable part of medical practice by the year **2010**.

Using genomic and bioinformatic approaches to study cancer and diseases genetics including:

1. Digital SNP Analysis

Digital SNP analysis can individually counted the paternal or maternal alleles within any clinical DNA samples, thus allowing a quantitative measure of AI in the presence of normal DNA. The significance of this technology is to use SNP molecular diagnostic test in identifying asymptomatic patients with early and clinically curable neoplastic diseases. DNA from plasma, ascites, paraffin, and urine samples can perform well using this technology.

2. Data mining from Serial Analysis of Gene Expression (SAGE)

Data mining for SAGE is able to screen the tumor markers and/or tumor suppressors via the aide of real time RT-PCR. Different tumors are selected for this study, including ovarian caners, oral cancers now. Several tumor candidate genes were under cloning and waiting for protein expression.

3. SNP Genotyping and Association Study

DNA from any source, e.g., hair shafts, blood, or buccal, is suitable for analysis. Under properly sampling for certain diseases, the risk factor for specific SNP associated with diseases is estimated.