An update on technologies relevant to carrier screening

- Distinctions: What exactly are you trying to do?
  - Monogenic disorders
  - Copy number variations
  - Polygenic

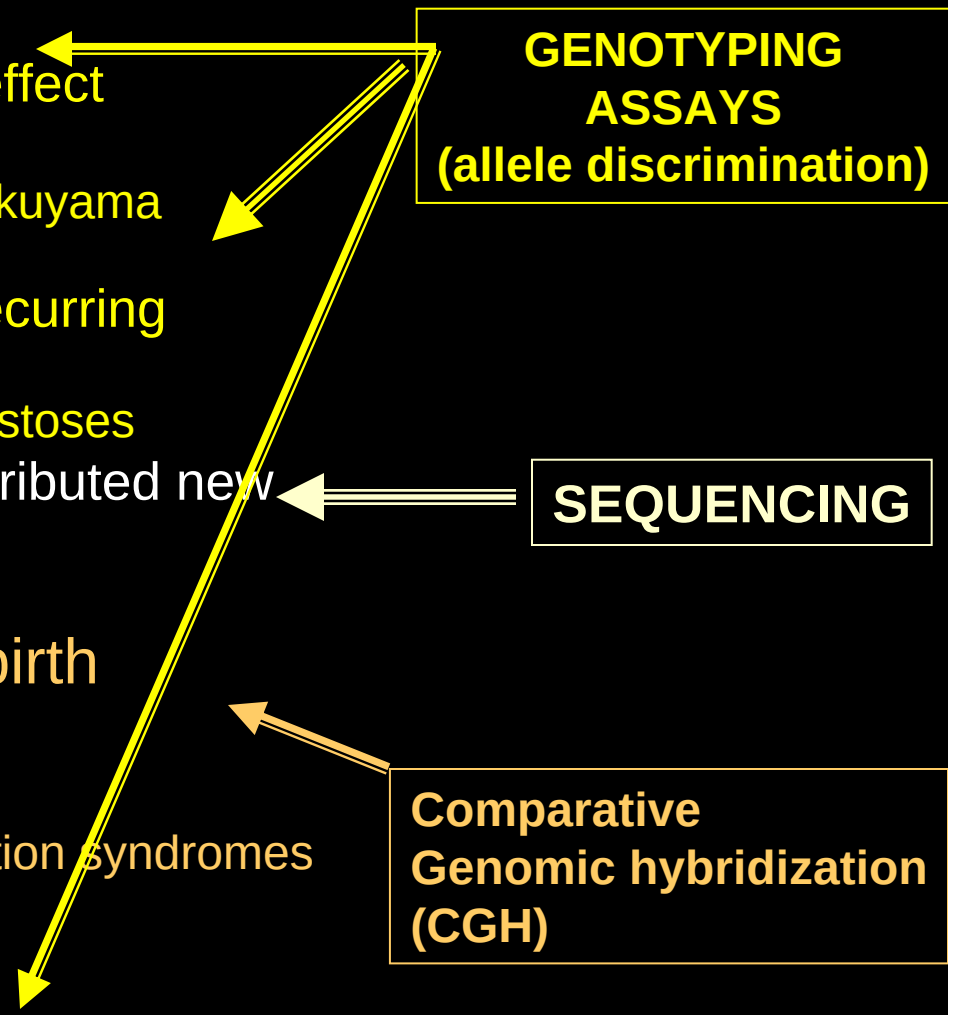
Eric Hoffman, Children's National Medical Center, Washington DC

Intense gratitude for research support to:
NICHD (MRDDRC, NCMRR, Wellstone Center)
NIAMS, NINDS, Department of Defense

Carrier screening: Newborn vs. Adult

- Dr. Alexander: “Whatever is done in neonatal screening can be applied to carriers”
- NOW: Newborn screening for patients
  - diagnosis
  - mass spec, biochemistry
- FUTURE? Newborn screening for carriers
  - ethical considerations
  - if we broach ethical issues, likely genetic (not mass spec or biochem)
- Technology likely same for newborn vs. adult.
  - Dr. Calonge. Technology same, and available.

- Distinctions
  - **Monogenic** disorders
    - Common mutations, founder effect
      - CF, Sickle cell, Goucher
      - Limb girdle dystrophy 2B, Fukuyama muscular dystrophy
    - High mutation rate, *de novo* recurring mutations
      - Achondroplasia, cranial synostoses
    - High mutation rate, widely distributed new mutations
      - DMD, NF, TS
  - **Copy number variations** (birth defects, etc)
    - Common recurring
      - Prader Willi, Angelman, deletion syndromes
    - Personal CNVs
      - MR
  - **Polygenic** (SNP associations)
    - 7 strongly validated loci T2DM

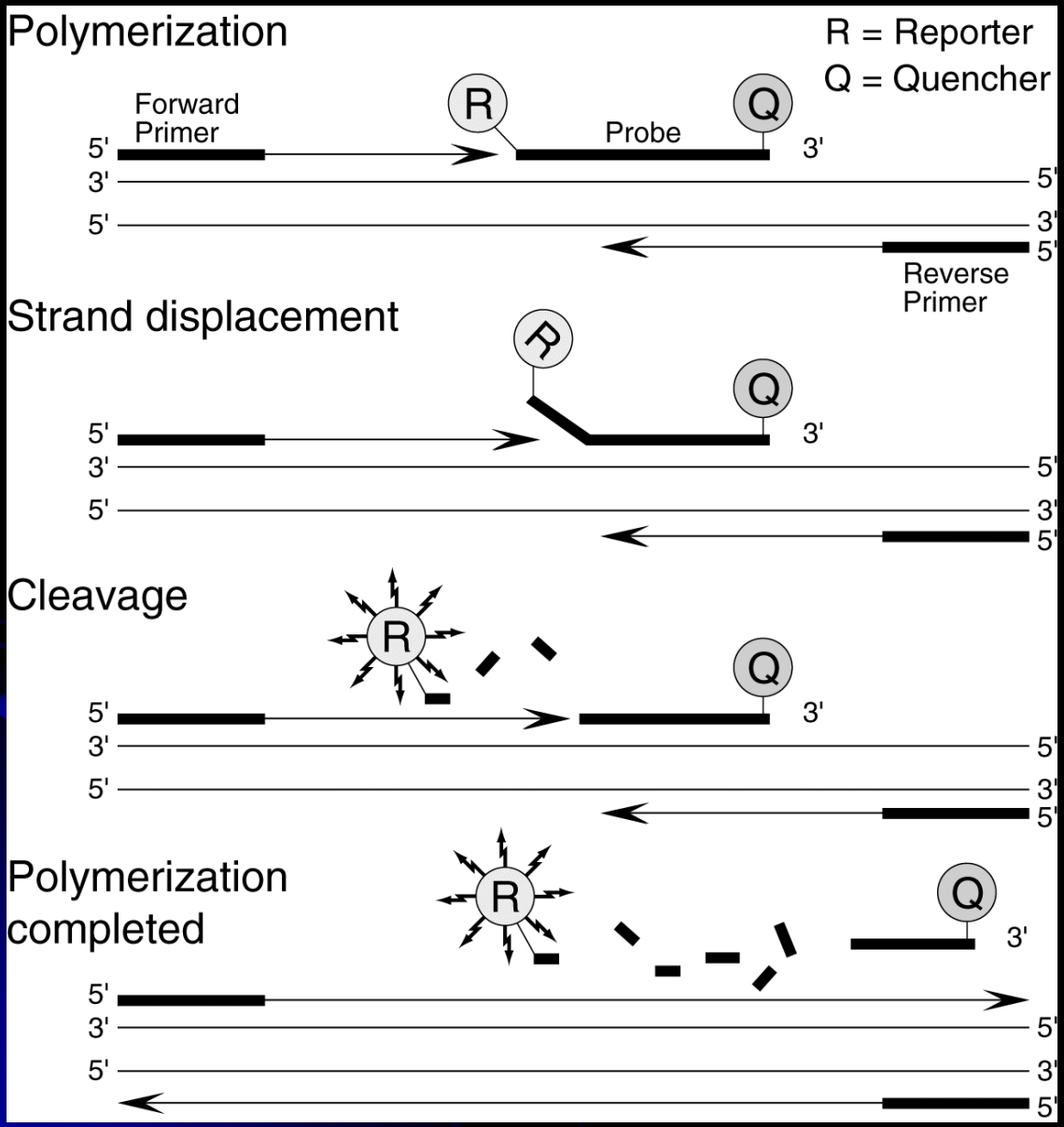


An update on technologies relevant to carrier screening

GENOTYPING

- Distinctions
  - Monogenic disorders
    - Common mutations, founder effect
      - CF, Sickle cell, Goucher
      - Limb girdle dystrophy 2B, Fukuyama muscular dystrophy
    - High mutation rate, *de novo* recurring mutations
      - Achondroplasia, cranial synostoses
    - High mutation rate, widely distributed new mutations
      - DMD, NF, TS
  - Copy number variations (birth defects, etc)
  - Polygenic (SNP associations)

Genotyping assays



TaqMan assays

Quantitative PCR

- Red fluor (product 1)
  - Internal control
- Yellow fluor (product 2)
  - Viral DNA of interest

Alternative: Roche Light cycler
Roche recently bought
Nimblegen, 454

Genotyping

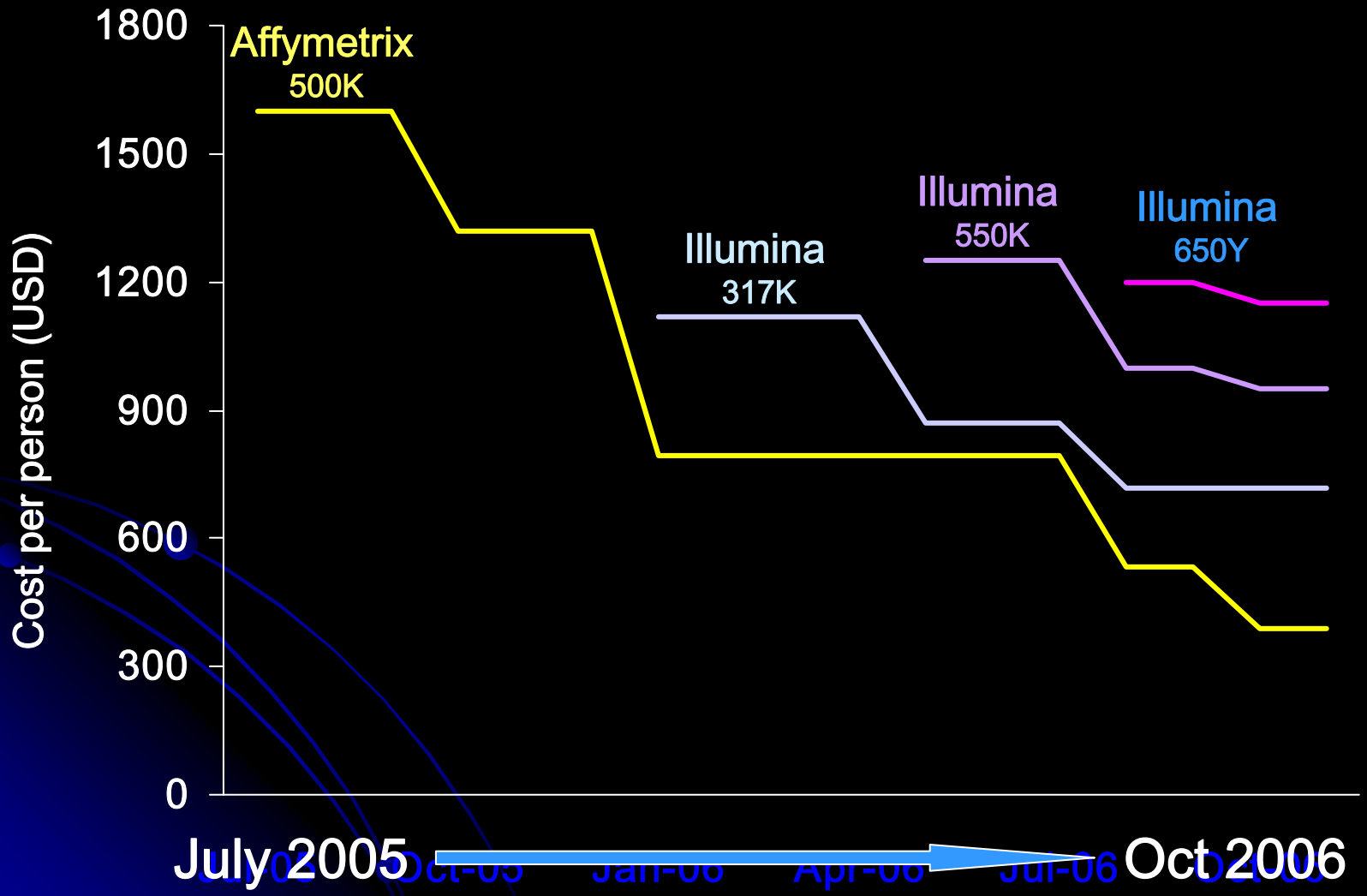
Allele discrimination assays

- Red fluor (allele 1)
- Yellow fluor (allele 2)

Other genotyping assays

- SNP chips
  - Not really.
  - Do not target known mutations.
  - Call rates not as high as TaqMan
  - More expensive
  - Overkill
- Custom Mol Dx microarrays?
  - Affymetrix and others working on ‘POC’ Mol Dx machines
  - More focused on expression arrays, cancer

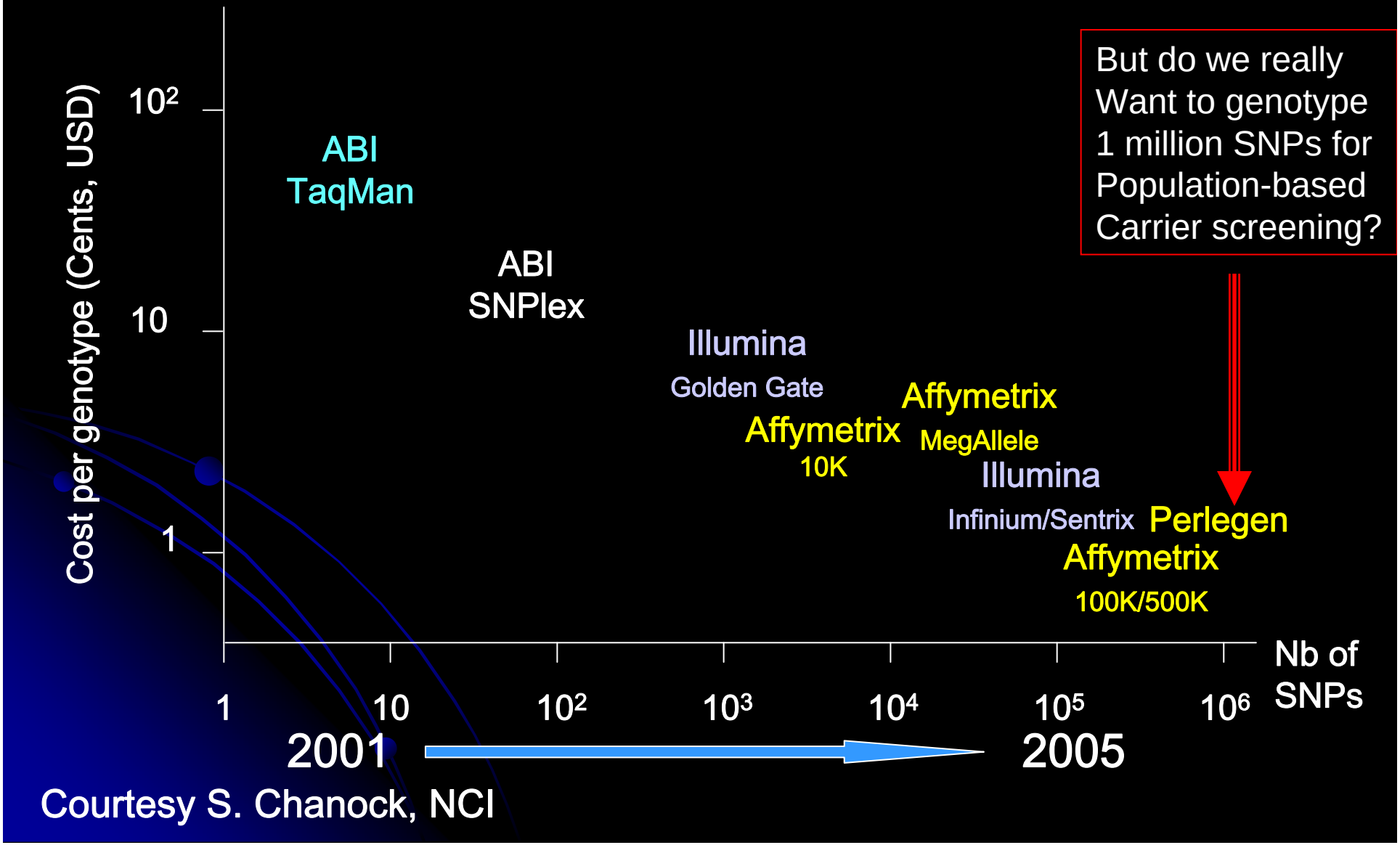
Continued Progress in Genotyping Technology



Courtesy S. Gabriel, Broad/MIT

Progress in Genotyping Technology

Intoxication by numbers



Courtesy S. Chanock, NCI

Highly parallel TaqMan for population-based carrier screens

- Panels of mutations for single disease, or number of diseases
- BioTrove

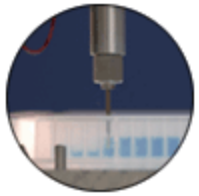
BioTrove

BREAKING NEWS

- BioTrove Names New Division Vice President-General Managers
- Children's National Medical Center to Investigate Diabetes Genetic Markers with OpenArray™ System
- BioTrove Names Edward "Buzz" Sztukowski Senior Vice President and Chief Business Officer
- BioTrove, Inc. CEO Albert Luderer to Present at Piper Jaffray Health Care Conference
- Applied Biosystems and BioTrove, Inc. to Collaborate on Integrated Platform for High-Throughput Genotyping
- BioTrove, Inc. Appoints Jeffrey C. Leathe as Chief Financial Officer
- Agilent Technologies and BioTrove Sign Collaborative Marketing Agreement for Ultra-

BioTrove offers researchers the latest in high throughput technologies:

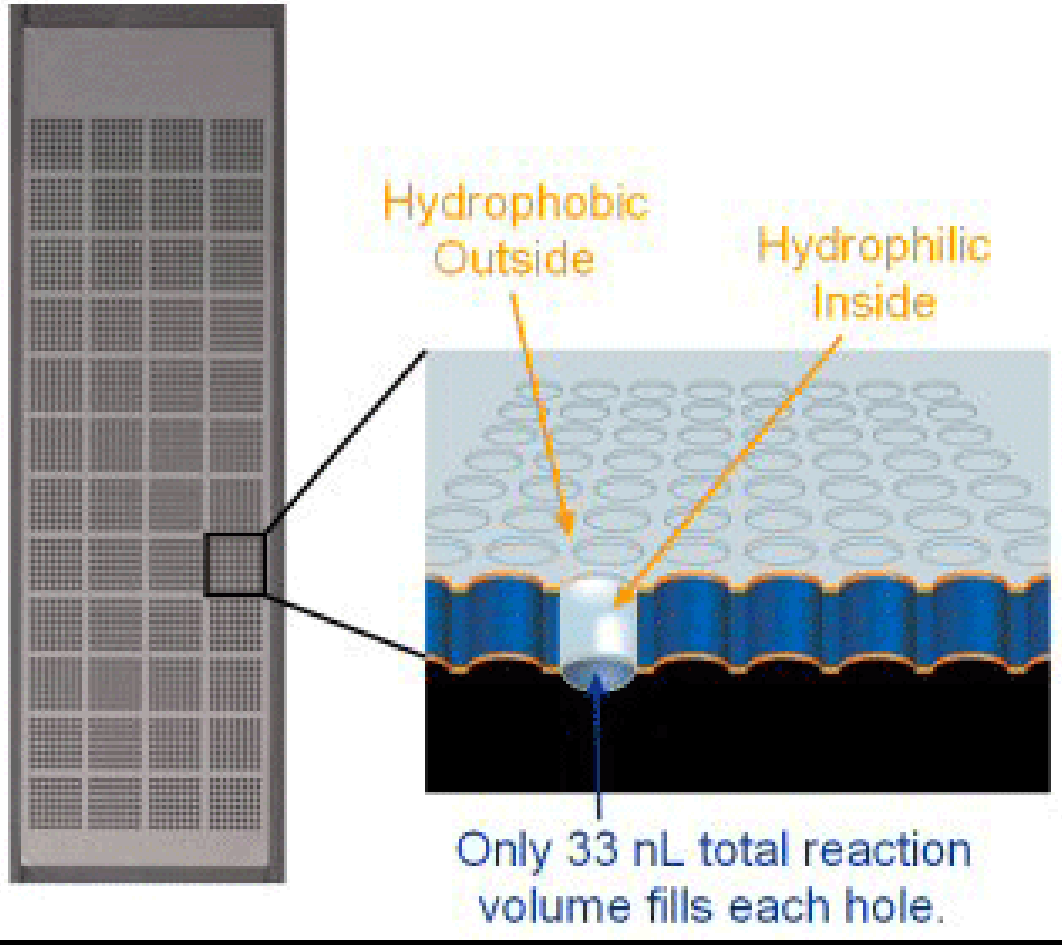
OpenArray™ plates hold over 3000 nanoliter scale PCR reactions in a flexible format.

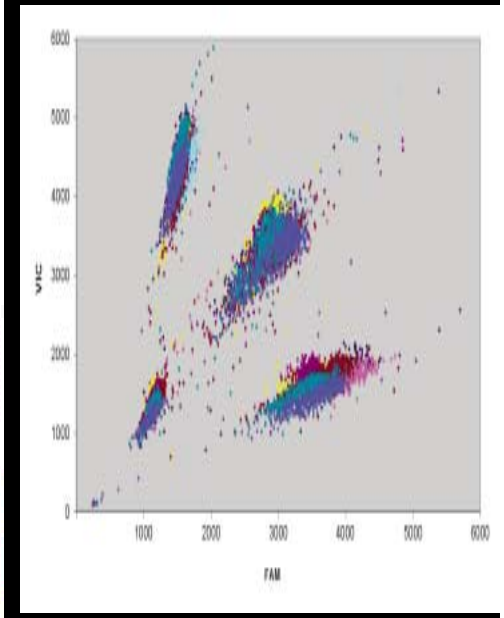


RapidFire™ technology performs pharmaceutical compound screening with mass spectrometry in < 8 seconds.

64 holes are in each 8 x8 subarray.

3072 holes are in each OpenArray™ plate.





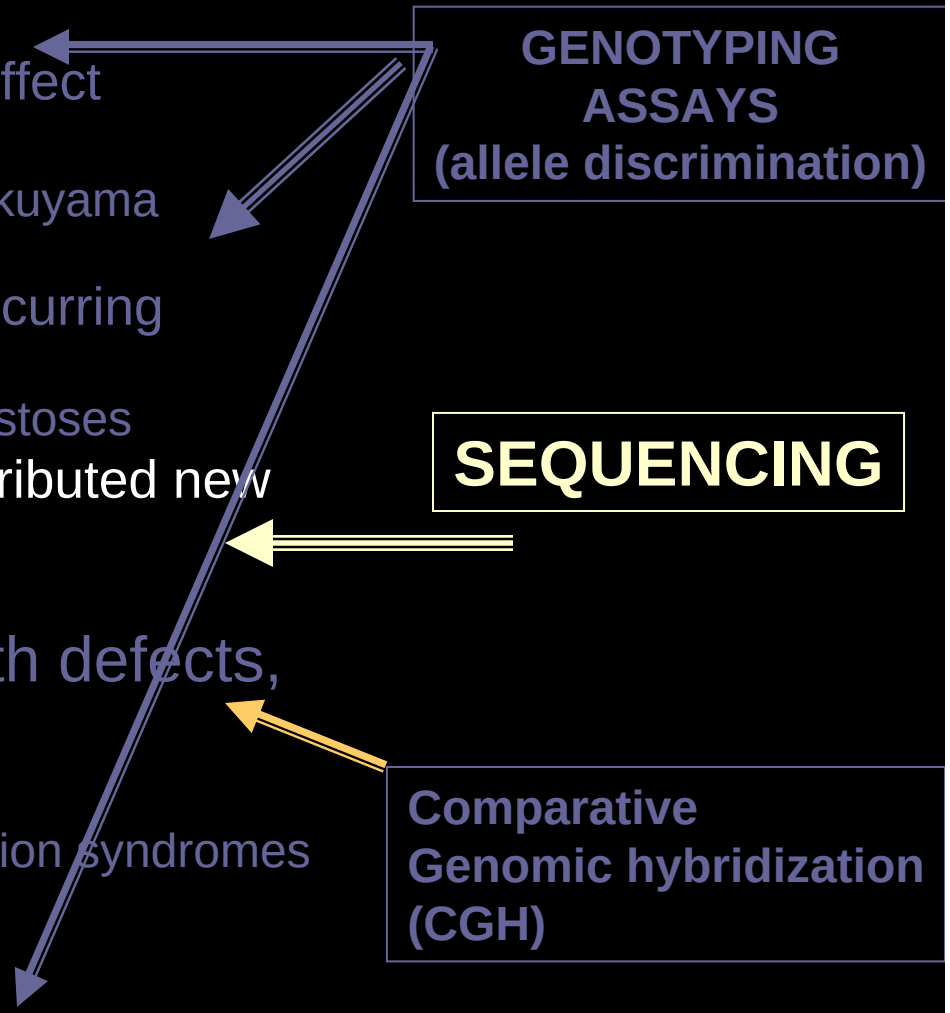
24,000 genotypes Overlaid from 8 arrays

- TaqMan assays are pre-loaded into wells, complete flexibility in format
- Sickle cell: put single mutation in all wells, genotype 3,072 people/plate
- CF: put 23 mutations as subarray, genotype 150 subjects/plate

Other genotyping platforms

- Illuminex
  - multiplex PCR, sequence tagging of products, bead pull down, read out
- Many others…….
- Bottom line:
  - Highly parallel genotyping
    - Cheap
    - Accurate
    - Moving into lab medicine
    - Becoming automated, nanoscale

- Distinctions
  - Monogenic disorders
    - Common mutations, founder effect
      - CF, Sickle cell, Goucher
      - Limb girdle dystrophy 2B, Fukuyama muscular dystrophy
    - High mutation rate, *de novo* recurring mutations
      - Achondroplasia, cranial synostoses
    - High mutation rate, widely distributed new mutations
      - DMD, NF, TS
  - Copy number variations (birth defects, etc)
    - Common recurring
      - Prader Willi, Angelman, deletion syndromes
    - Personal CNVs
      - MR
  - Polygenic (SNP associations)
    - 7 strongly validated loci T2DM



- Distinctions
  - Monogenic disorders
    - **High mutation rate, widely distributed new mutations**
      - **Scope of problem:**
      - **DMD**
        - **Recessive, identify female carriers**
        - **High mutation rate: even if found ALL carriers would only reduce disease frequency by 50%**
        - ***Unless you want to screen each egg or pre-implantation embryo***
        - **2.7 million bp for entire gene (add some for promoter); 11 kb coding sequence (79 exons)**
      - **NF**
        - **Dominant. NOT identifying carriers, identifying patients**
        - **But many clinically mild, could find them pre-symptomatic**
        - ***But this discussion not point of workshop***
      - **TS**
        - **Dominant. NOT identifying carriers, identifying patients**
        - **Ditto**

SEQUENCING

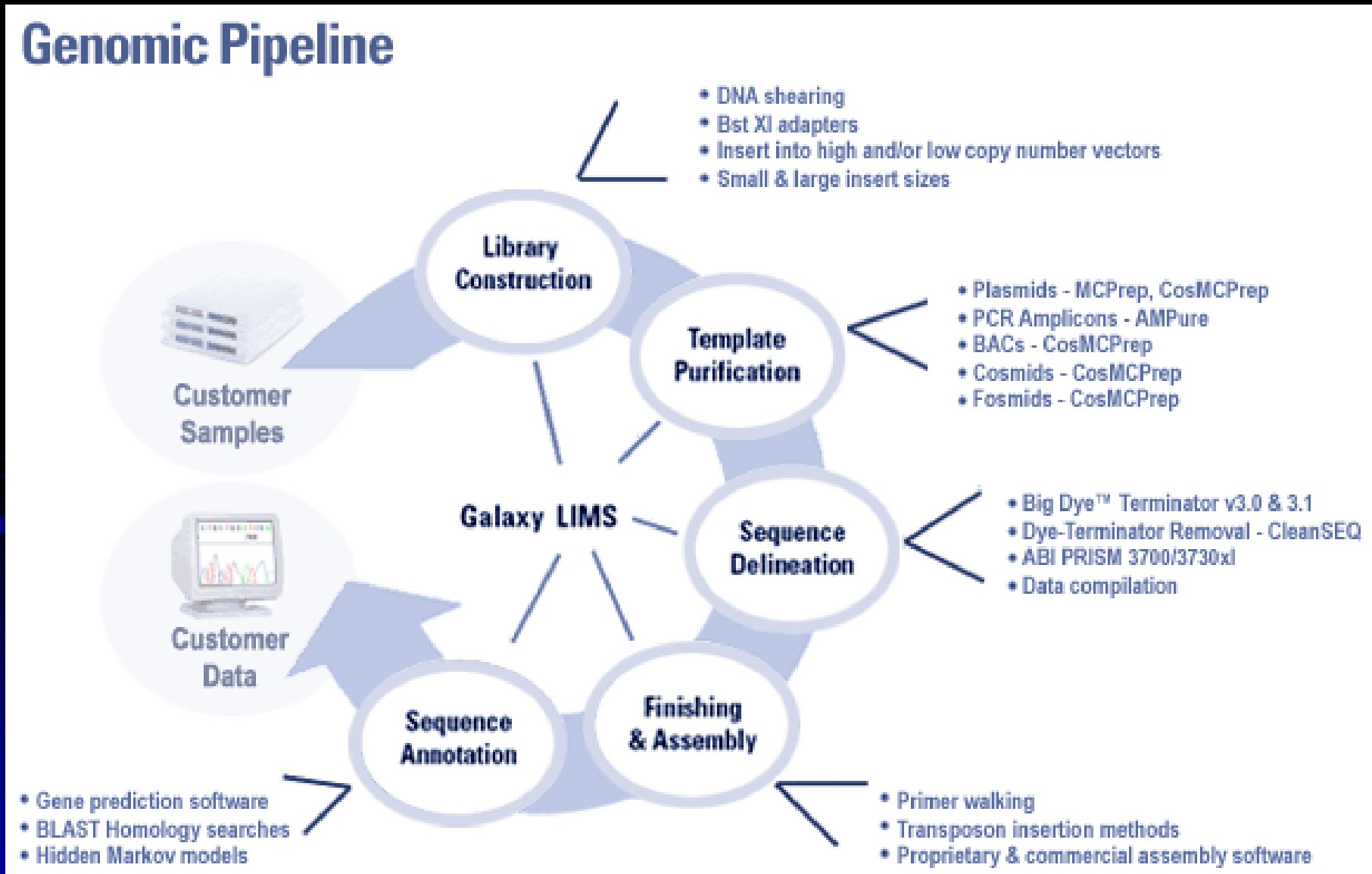
Do you want to sequence everything, or just disease genes?

- **High throughput whole genome sequencing**
- **Next gen sequencing: $1,000 genome**

- Targeted sequencing (e.g. dystrophin gene)

- *Great advances in whole genome sequencing do not necessarily advance targeted sequencing*

Next gen sequencing pipeline
Shot gun methods; Whole genome



Designs for Discount Genes

Four companies want to shake up the world of genetics, making it possible to read out a person's DNA blueprint for under $10,000. Here's where they are now:

Company	Machine	Cost of machine	Current cost to sequence a person's genome	Time required to sequence a person's genome
Applera	Applied Biosystems 3730 xl	About $300,000	About $10 million	Several years
Roche Holdings	454 Genome Sequencer	About $500,000	$2 million–$3 million	1–2 months
Applera	Applied Biosystems Solid System	$500,000–$600,000	About $300,000	About 2 months
Illumina	Genome Analyzer	$400,000	About $300,000. Will be $100,000 by 2008	2–3 months
Helicos BioSciences*	Heliscope	About $2 million	About $100,000	6–7 weeks

*Available later this year

Sources: the companies; J. Craig Venter Institute

- Wall Street Journal, Oct. 4, 2007

Not quite ready for POC in all path labs.
But even if it was ready, would we want or need it for population-based screening?

Do you want to sequence everything, or just disease genes?

- High throughput whole genome sequencing
- Next gen sequencing: $1,000 genome

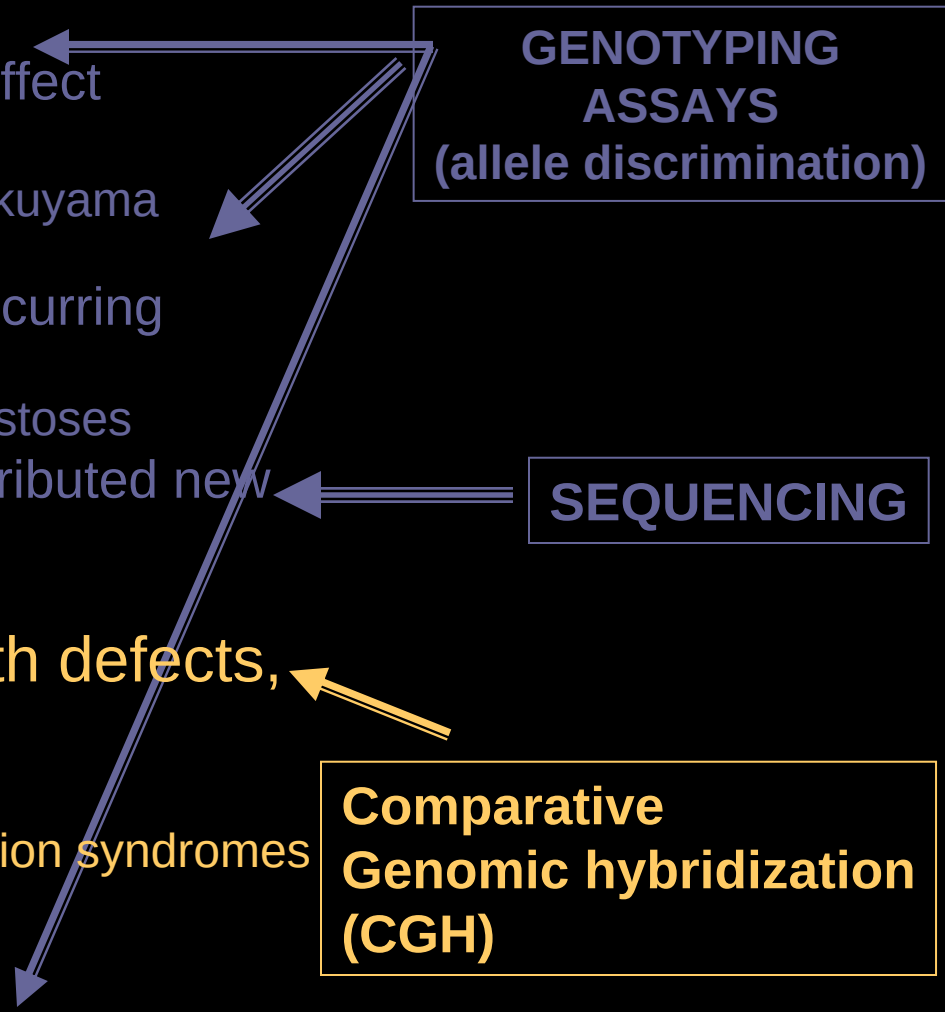
- **Targeted sequencing (e.g. dystrophin gene)**

- *Great advances in whole genome sequencing do not necessarily advance targeted sequencing*

Re-sequencing arrays

- **Affymetrix**
  - Custom seq arrays
    - 300,000 bp of ‘real estate’
    - Sequencing by hybridization
    - Applications:
      - 10 cardiac genes Harvard Partners
      - Mitochondrial re-sequencing (bit of trouble with GC rich regions)
      - NCI oncogene re-sequencing
  - Major issue
    - Multiple PCRs
      - Things don’t like to be multiplexed
      - Dystrophin gene alone: 100 PCRs
        - Can work on multiplexing some
        - But still many independent PCRs, then mixing

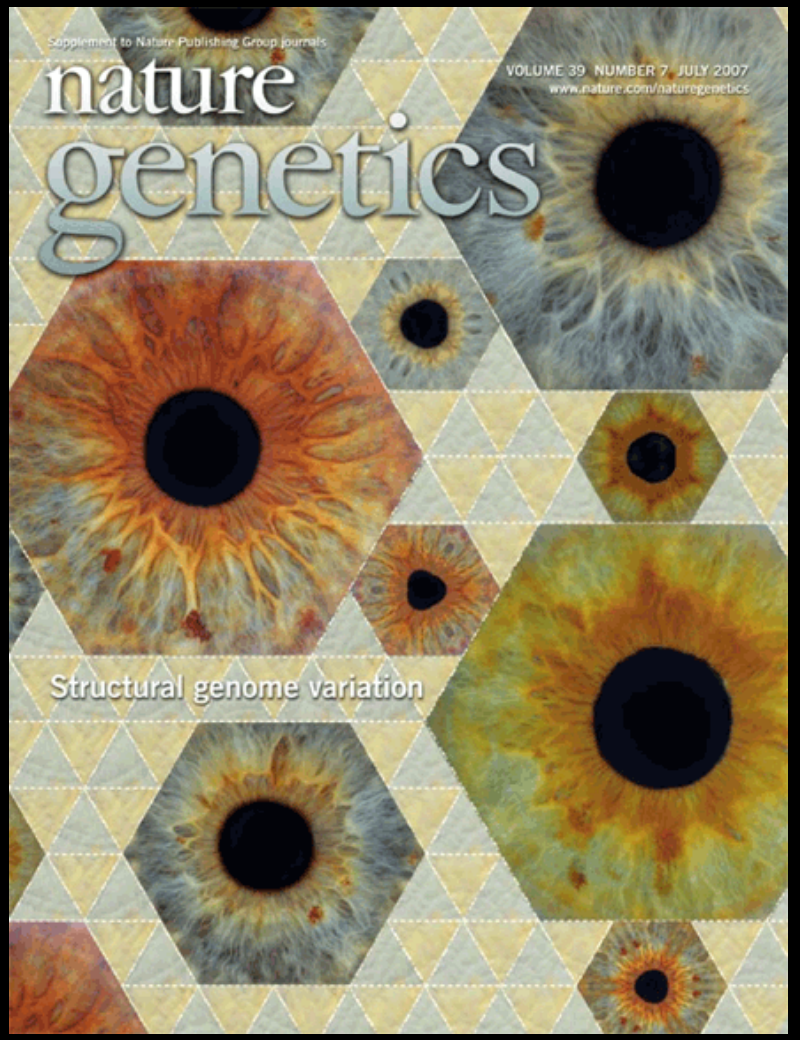
- Distinctions
  - Monogenic disorders
    - Common mutations, founder effect
      - CF, Sickle cell, Goucher
      - Limb girdle dystrophy 2B, Fukuyama muscular dystrophy
    - High mutation rate, *de novo* recurring mutations
      - Achondroplasia, cranial synostoses
    - High mutation rate, widely distributed new mutations
      - DMD, NF, TS
  - Copy number variations (birth defects, etc)
    - Common recurring
      - Prader Willi, Angelman, deletion syndromes
    - Personal CNVs
      - MR
  - Polygenic (SNP associations)
    - 7 strongly validated loci T2DM



Summary of 12 Surveys of Structural Variation

CNVs between individuals amount to 4 Mb (1/800 bp) of genetic difference, and less conservative estimates put this figure in the range of 5–24 Mb.

CNVs account for more nucleotide variation on average than SNPs: 2.5MB, 1/1,200 bp



Scherer et al. Nature Genetics 39, S7-S15 (2007)

Examples of Disease Associated CNVs:
But almost all not a 'carrier' situation (patient diagnosis)

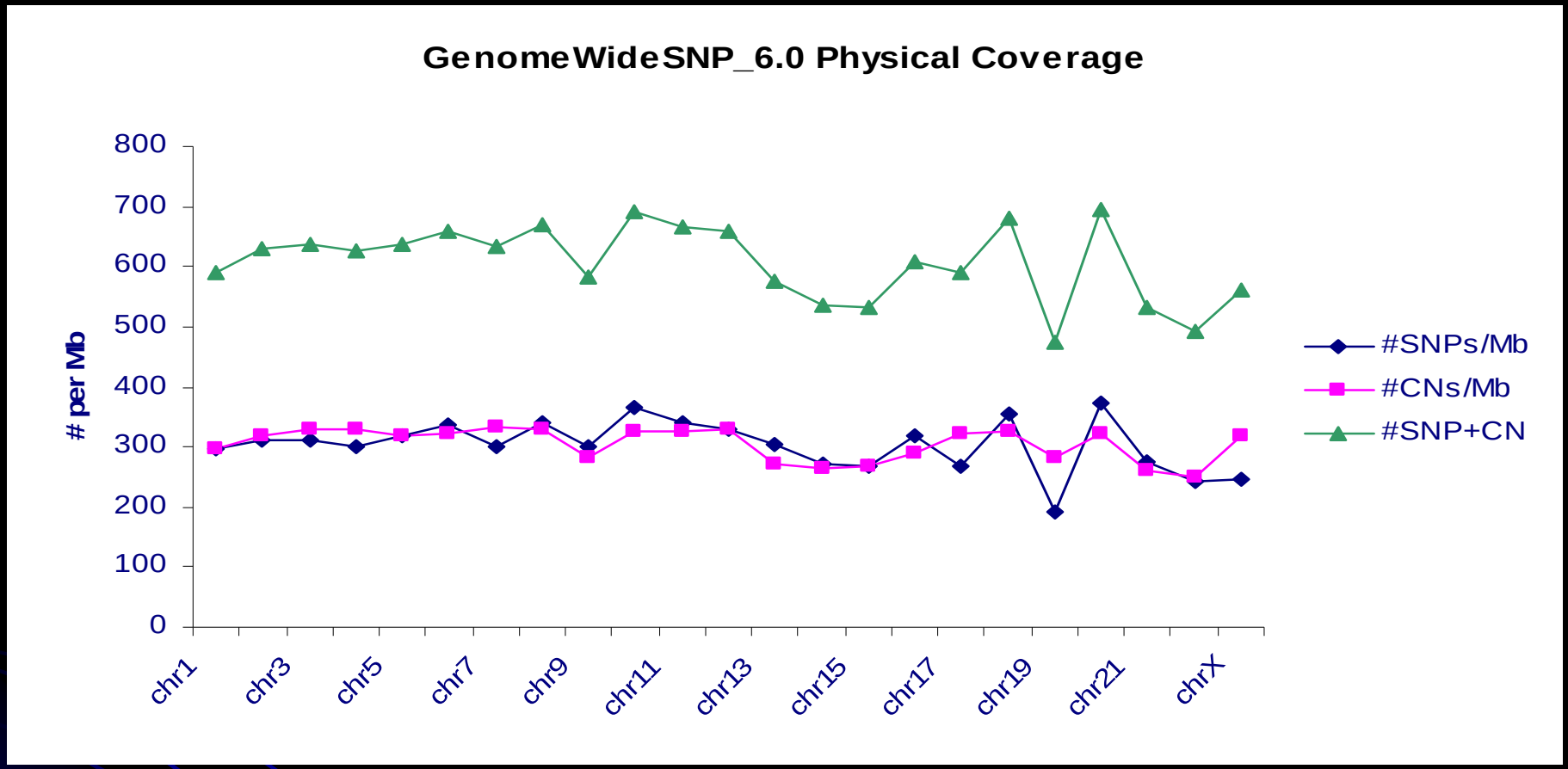
Disease	Gene	Phenotype
Charcot-Marie-Tooth type 1A	*PMP22*	Demyelination, peripheral neuropathy
X-linked hypopituitarism	*SOX3*	In males, short stature, mild mental retardation
Autosomal dominant leukodystrophy	*LMNB1*	Demyelination, white brain matter abnormalities
Parkinson's	*SNCA*	Neuron degeneration, rigidity, tremor
Alzheimer's	*APP*	Amyloid beta precursor protein buildup
Altered drug metabolism	*CYP2D6*	Increased side effects, increased or decreased efficacy
HIV/AIDS	*CCL3L1*	Increased susceptibility to infection and disease
Lupus	*FCGR3B*	Increased susceptibility to kidney failure
Smith-Magenis syndrome	*RAI1*	Mental retardation
Pelizaeus-Merzbacher	*PLP1*	Demyelination, paralysis of legs, involuntary jerking of head
Spinal muscular atrophy	*SMN1*	Spinal deterioration, milder disease w/later onset
Rett-like syndrome	*MECP2*	Mental retardation, spasticity, language/speech problems

Platform Overview

	Affymetrix SNP 6.0	Illumina 1M
List Price	$375	~$650
Total # Genetic Markers	~1.8 million	~1.0 million
Number of SNPs	~906K	~1050K
Number non-polymorphic CNV probes	~946K	~22K
Use of Whole Genome Amplified Samples	**YES**	Up to 4% decrease in Call Rate**
Company Demonstrated* Call Rates	**99.8%** (270 HapMap samples)	99.66% (125 DNA samples)
Open Informatics Site	**YES**	NO
Open access to algorithms	**YES**	NO
3rd Party Software Compatible	YES	YES
Scanner throughput per day	**40 samples**	24 samples
Open Automation platform	**YES**	NO
Median Marker Spacing	680 bases	1,700 bases

***Data from Specification Sheets on company websites.**
****From Illumina website http://www.illumina.com/pagesnrn.ilmn?ID=82**

SNP and CNV markers across multiple chromosomes

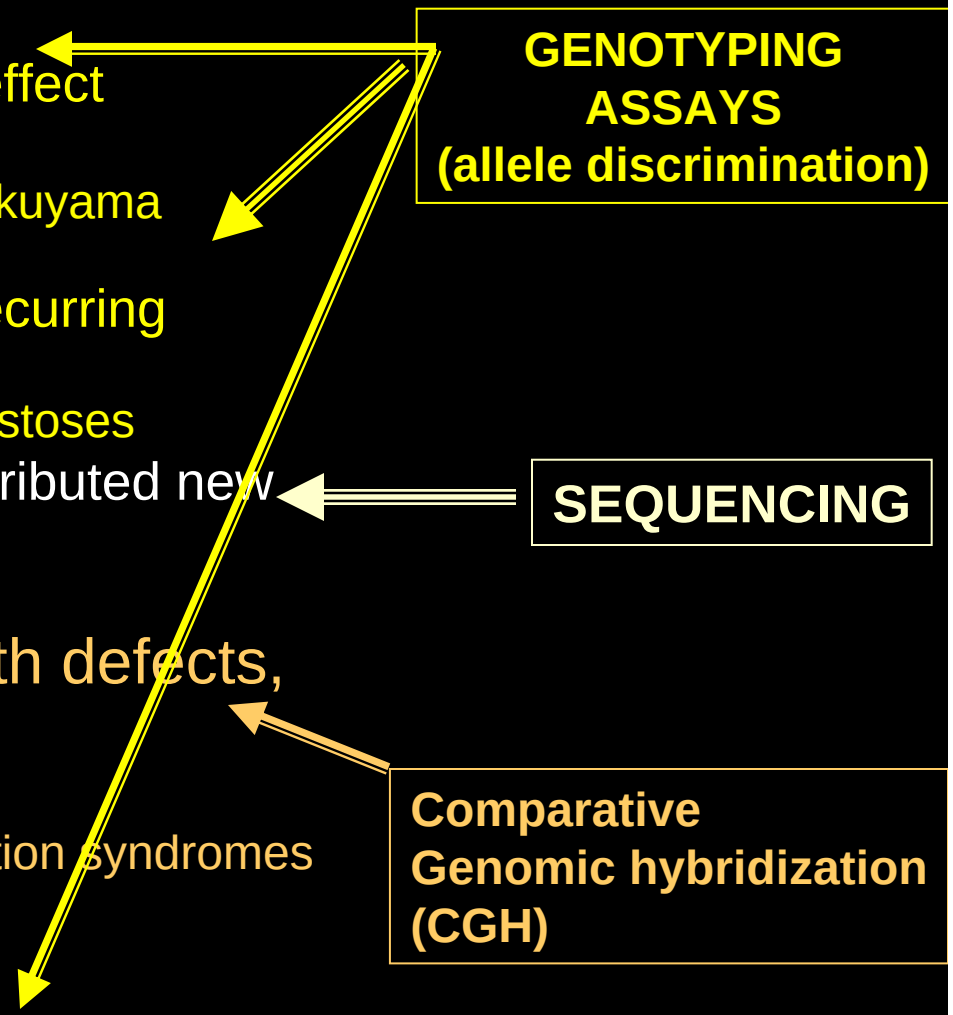


Median SNP + CNV inter-marker distance = 696 base pairs

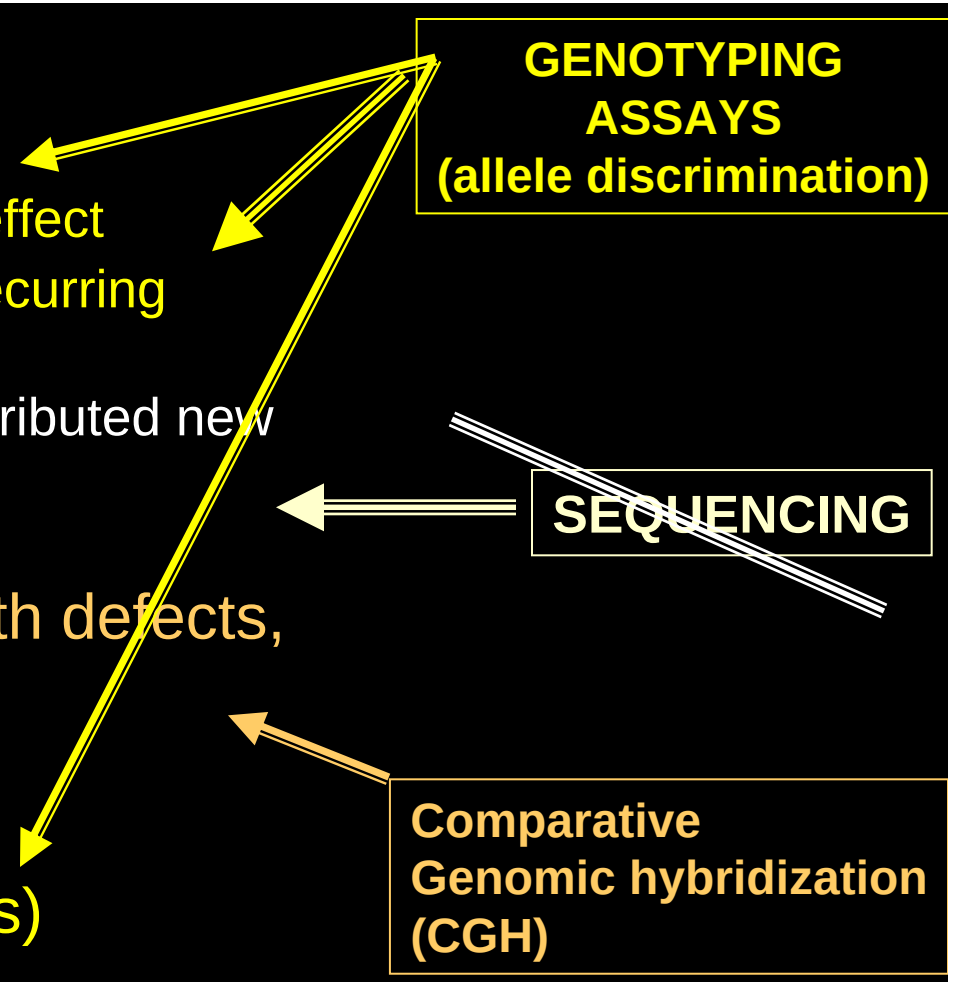
SNP chips great, but

- Will not sequence disease genes;
- Current versions do NOT contain common disease gene mutations;
- Likely movement into hybrid CGH – disease genotyping chips

- Distinctions
  - Monogenic disorders
    - Common mutations, founder effect
      - CF, Sickle cell, Goucher
      - Limb girdle dystrophy 2B, Fukuyama muscular dystrophy
    - High mutation rate, *de novo* recurring mutations
      - Achondroplasia, cranial synostoses
    - High mutation rate, widely distributed new mutations
      - DMD, NF, TS
  - Copy number variations (birth defects, etc)
    - Common recurring
      - Prader Willi, Angelman, deletion syndromes
    - Personal CNVs
      - MR
  - Polygenic (SNP associations)
    - 7 strongly validated loci T2DM



- Distinctions
  - Monogenic disorders
    - Common mutations, founder effect
    - High mutation rate, *de novo* recurring mutations
    - High mutation rate, widely distributed new mutations
      - DMD, NF, TS
  - Copy number variations (birth defects, etc)
    - Common recurring
    - Personal CNVs
  - Polygenic (SNP associations)
    - 7 strongly validated loci T2DM



Prediction: Hybrid chip-based assays (TaqMan, hyb assays)
- All common mutations in recessive disease
- Quantitative assays (CGH-type) for common copy number variations