The Way Ahead.™

# The Challenge of ChIP Sample Preparations

## Experiences from Affymetrix and Recommendations for the Experimental Design

*Matthias Prucha, PhD*
*Senior Applications Specialist and European Tiling Array Specialist , Affymetrix Europe*

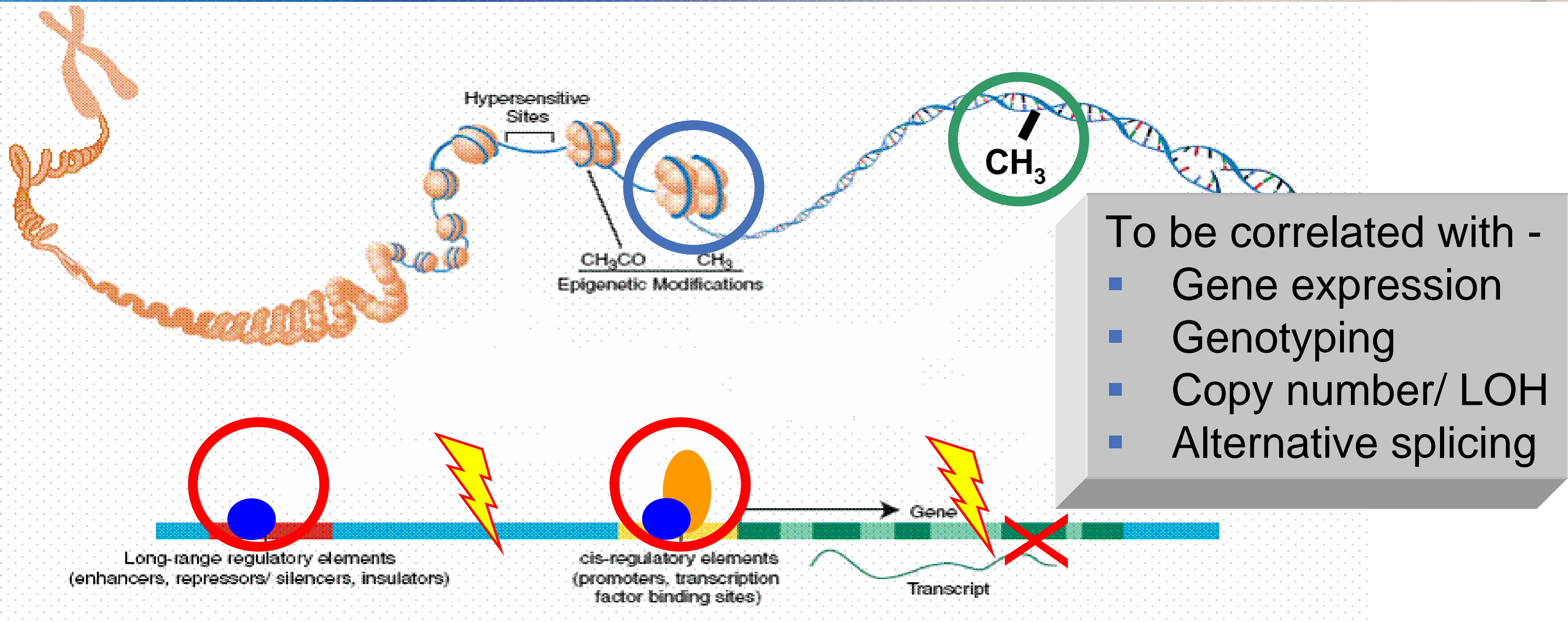
1st European ChIP-on-chip Scientific Forum

# Outline of this Talk

- **The Challenge**
  - **Technically very different ChIP applications**
  - **General ChIP sample preparation workflow**
- ChIP experimental design
- Affymetrix ChIP protocol

# Identification of Altered Gene Regulation

AFFYMETRIX
The Way Ahead.™




To be correlated with -
- Gene expression
- Genotyping
- Copy number/ LOH
- Alternative splicing

Genome-Wide analysis of transcription regulation using ChIP-on-Chip identifies disease-related modifications at the epigenetic and transcription machinery levels:
- **Histone modification**
- **DNA methylation**

} **Epigenetics**

- **Transcription factor- and other regulatory protein binding activity**

1st European ChIP-on-chip Scientific Forum

AFFYMETRIX
The Way Ahead.™

# (Some) ChIP-chip Publications with Affymetrix Tiling Arrays

- **Cawley S et al., Unbiased Mapping of Transcription Factor Binding Sites along Human Chromosomes 21 and 22 Points to Widespread Regulation of Noncoding RNAs. Cell. 2004 Feb 20;116(4):499-509** *
- Bernstein B et al. Genomic Maps and Comparative Analysis of Histone Modification in Human and Mouse. Cell January 28, 2005, Vol 120, 169-181.
- Carroll J, et al. Chromosome-Wide Mapping of Estrogen Receptor Binding Reveals Long-Range Regulation Requiring the Forkhead Protein FoxA1. Cell, July 12, 2005, Vol. 122, 33-43
- Schumacher A, et al. Microarray-based DNA methylation profiling: technology and applications. Nucleic Acids research, 2006 Jan., Vol 34 (528-542)
- Bernstein BE, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell. 2006 Apr 21;125(2):315-26
- Schwartz YB, et al. Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. Nat Genet. 2006 Jun;38(6):700-5. Epub 2006 May 28. PMID: 16732288 *
- Zhang X. et al. Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis. Cell 2006 Sept 22; 126: 1189-1201
- Carroll J. et al. Genome-wide Analysis of estrogen receptor binding sites. Nat. Genetics Online 2006, Oct 1
- Hayashi H., et al. High-resolution mapping of DNA methylation in human genome using oligonucleotide tiling array. Hum Genet 2006 Nov, Online

*: In direct relation to Affymetrix ChIP protocol

# Chromatin Immunoprecipitation Procedure - Optimization Required!

AFFYMETRIX

The Way Ahead.™

DNA

Target protein

+ formaldehyde

DNA

Cell lysis

Sonicate
~ 1kb

Add antibody

Add protein A
beads

Wash/Elute
DNA-Protein
complexes

Reverse
X-links

Isolate DNA

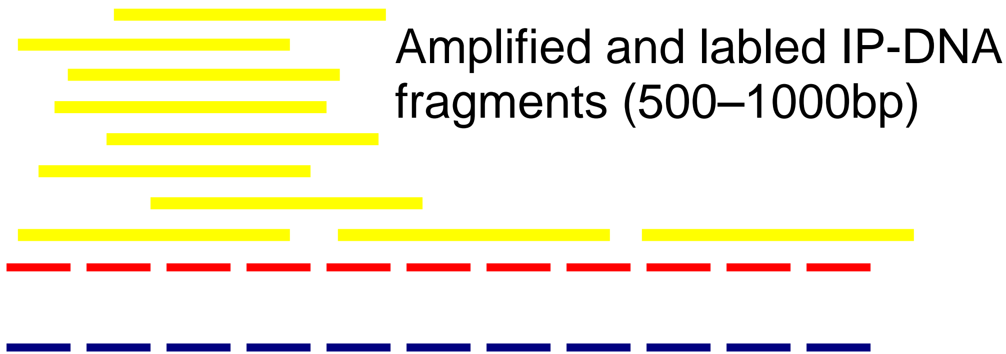
Amplify DNA

Fragmentation,
Labeling, &
Hybridization

Specific to the Affymetrix technology

1st European ChIP-on-chip Scientific Forum

# The Goal: Hybridization of IP-Enriched DNA

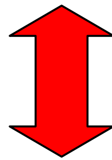


Amplified and labled IP-DNA fragments (500–1000bp)

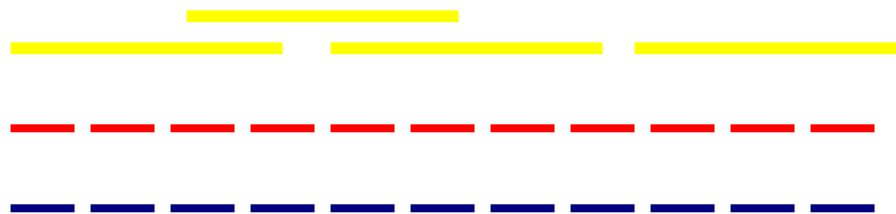
Enriched Sample

PM

MM



Amplified and labled non-specific DNA fragments



Control Sample

PM

MM

AFFYMETRIX

The Way Ahead.™

# Outline of this Talk

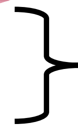
- The Challenge
- **ChIP experimental design**
  - **Complexity effects**
  - **Enrichment analysis**
  - **Amplification methods**
  - **Non-specific control**
- Affymetrix ChIP protocol

# ChIP Challenges are linked!

**Complexity**
- Genome size
- Frequency of IP sites

**Data Analysis**
- Normalisation
- Overlap with another specific IP (PolII e.g.)

**Antibody**
- Affinity
- Specificity
- Stability



**Achievable enrichment**
(check often, it can get lost)

**Enrichment analysis**

**Hybridization**
- Amount per array
- How many cocktails (for array sets)

**Amplification**
- Which PCR
- Complexity reduction
- Linear amplification (T7)

**Non-specific control**
- Input
- Mock IP
- Non-specific Antibody

# Complexity Sources

- Genome size matters
  - Human (3200 MB) vs Arabidopsis (157 MB: ~5% of Human)
- Frequency of binding
  - Very high
    - DNA-methylation (~19% of entire genome found in Arabidopsis, Zhang et al.)
  - ~Mid to high
    - Histone modifications e.g.
  - ~Low to mid
    - Transcription factors and other regulatory DNA-binding proteins show large differences of binding frequency
- Antibody-/ purification specificity
  - Non-specific binding of antibody but also non-specific binding of beads
  - Antibody specificity determines enrichment factor
  - ~95% of the specifically IP'ed DNA is non-specific

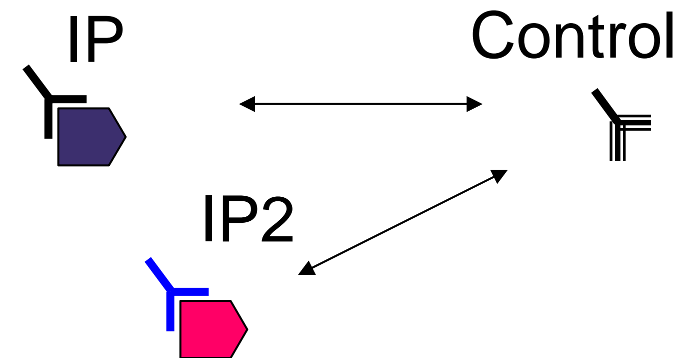
AFFYMETRIX
The Way Ahead.™

# Complexity Effects

- Amount of labeled DNA for array hybridization
  - Thermodynamics (competition for binding) is correlated with complexity
    - Low complexity (low binding frequency): 0.5 – 10µg per array
    - High complexity: >20µg per array might be useful (see other DNA-applications)
- Selection of amplification method
  - PCR-bias effects (short fragments vs long fragments e.g.) seems to be highly correlated with complexity
    - High complexity: DNA-methylation works in Human and Mouse only with PCR after complexity reduction (Schumacher et al.; GenPathway etc.) or with linear amplification (Hayashi et al.)
    - Low complexity: PCR-methods
- Selection of non-specific IP-control
  - Complexity should be similar between specific IP and unspecific control

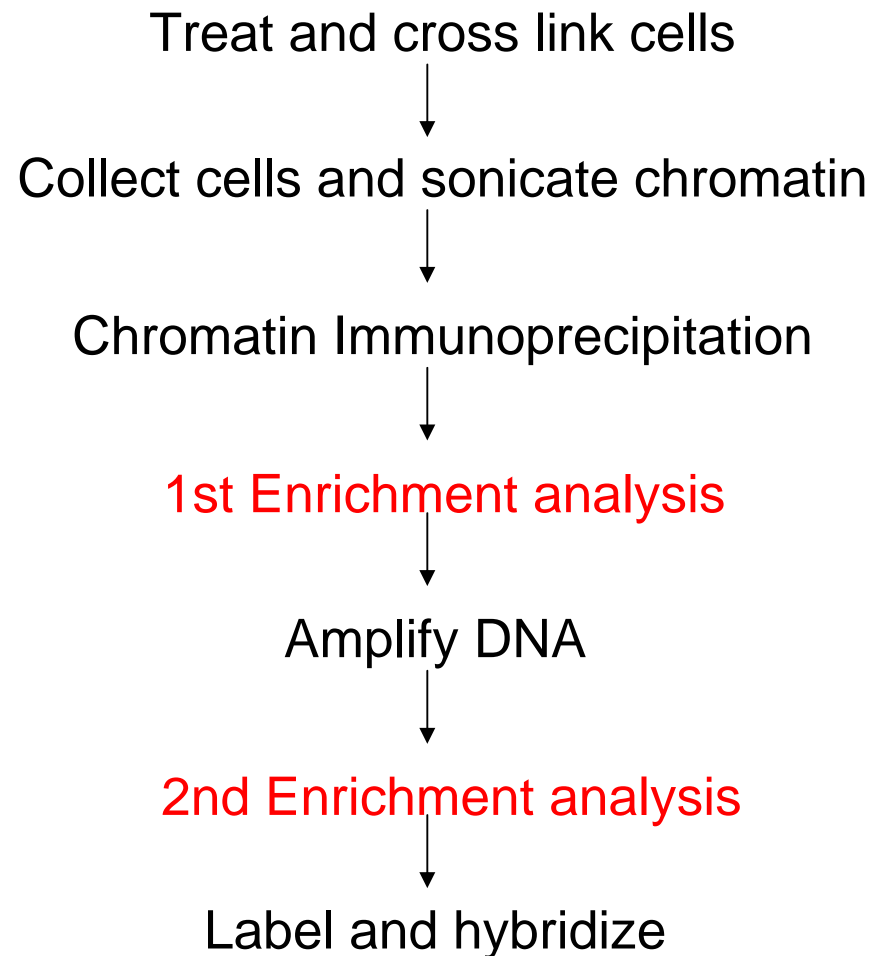
# Antibody Recommendations

- Only use ChIP-qualified (by Vendor) Ab
- Antibodies vary from batch to batch
  - Use large batches for entire experiment
- Antibody specificity might be unknown
  - Use multiple antibodies for same target
  - Use independent IPs for finding overlaps (bioinformatics approach!)
    - RNA-Polymerase II might be good control for TFs
- Amount of Antibody must be optimized for maximum enrichment (titration)



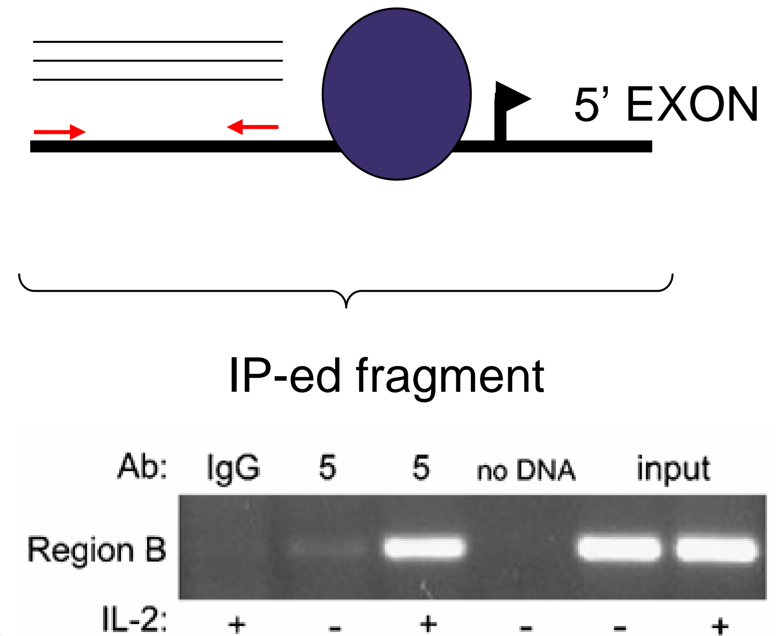
# Simplified Workflow – Validation needed

AFFYMETRIX
The Way Ahead.™



# Validation – Enrichment Analysis

- Enrichment = sensitivity of ChIP
  - >8-fold is ideal
- PCR of known binding sites
  - Use available hypothesis and literature
  - If unknown binding site… discussed later
- RT PCR versus standard PCR
- Tube to tube variation is guaranteed!
- Need to normalize for differences in DNA concentrations
  - Specific IP vs non-specific control

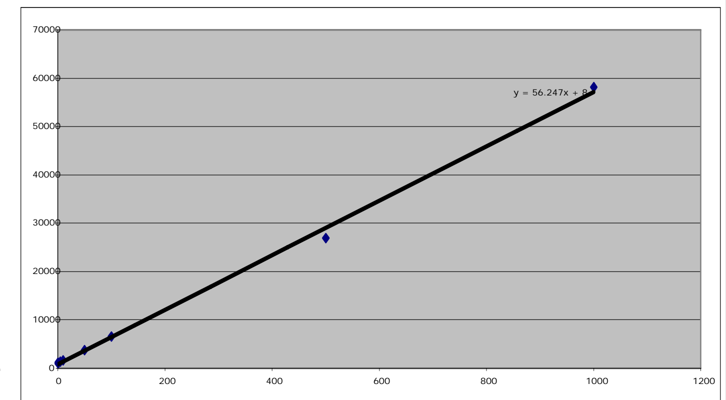


AFFYMETRIX®

The Way Ahead.™

# DNA Normalization

- Quantifying DNA is tricky at very low concentrations and large differences between IP- and non-specific control sample
  - Picogreen from Molecular Probes
    - Measures as little 1pg/ul
    - Requires fluorescent plate reader




- Using non-specific region in IP-sample
  - Normalize using primers to a region where no binding occurs
  - Intergenic region, specialized gene (Rhodopsin e.g.)
  - Makes assumption that most ChIP DNA is non-specific

| **Samples** | Fluorescence | pg/ul |
|---|---|---|
| ER Chip EtoH | 1789 | 16.66071429 |
| ER Chip E2 | 1585 | 13.01785714 |
| Input | 62086 | 1093.392857 |

AFFYMETRIX
The Way Ahead.™

# Unknown Binding Sites

Require a known binding site to monitor through the procedure
- If not:

1. Educated guess. Use expression array data to isolate candidate genes. Make primers to promoter proximal regions
2. Discover binding site: ChIP-cloning approaches
   Low success rate, but only need one positive
3. Blind approach: Parallel ChIP of a known positive control (RNA Pol2 etc). If this is successful, then hybridize unknown ChIP sample with one array (one seventh of the genome, chr 21/22, Encodev2)

# DNA Amplification – PCR bias?

Many amplification options:

- Random Priming (RP) PCR
- Ligation Mediated (LM) PCR (used also as complexity red.)
- Klenow
- ??? DNA – based assay

}
- Compatible with Affy ds DNA kit

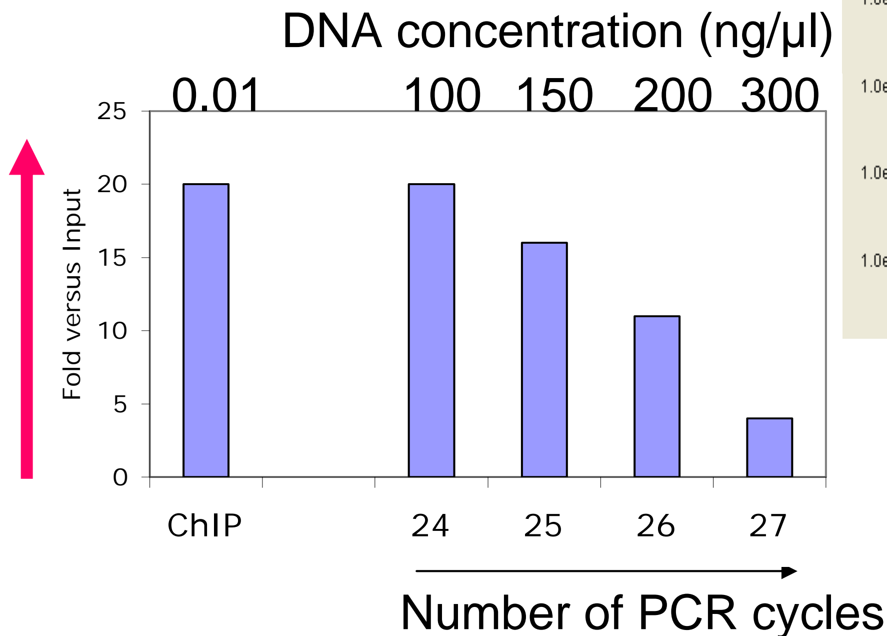
- T7 RP IVT amplification (RNA)
  - Need to be converted to cDNA for hybridization

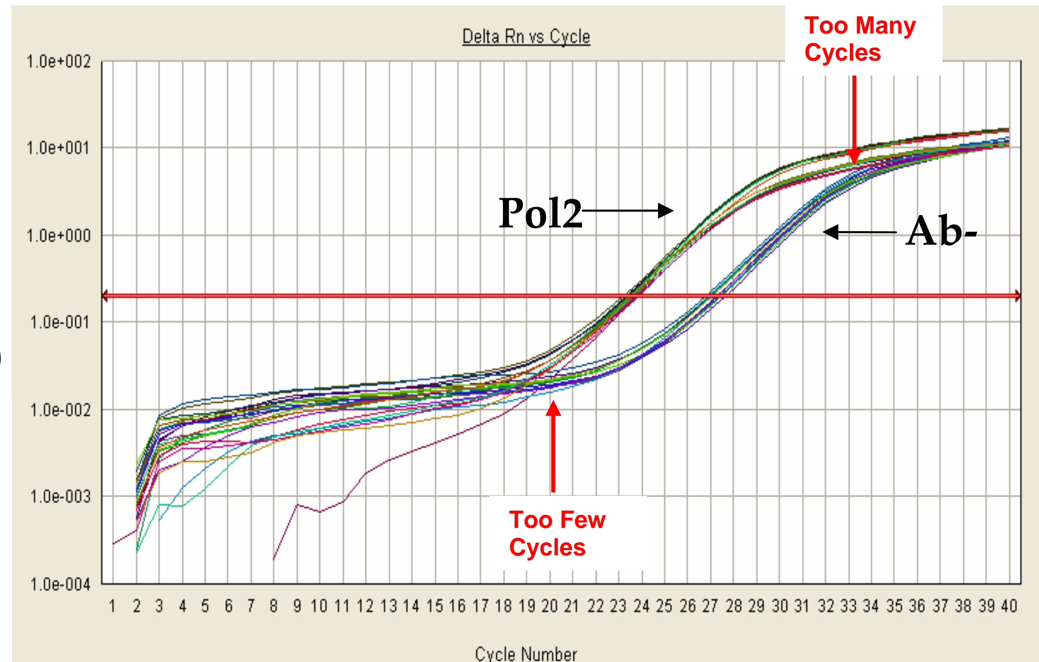
- Compatible with Affy IVT reagents
- Compatible with Affy ds DNA kit

# PCR-Saturation affects Enrichment

AFFYMETRIX
The Way Ahead.™

- Known PCR problems:
- Bias with fragment length
- Saturation with too many cycles




Don't forget to check enrichment after amplification!

AFFYMETRIX
The Way Ahead.™

# IP Control Options

- IP experiment (Experimental sample)
  - 10-15ug Ab to IP+ sample (Affymetrix protocol)
- Control experiment
  - Mock IP-DNA (no Ab)
  - Non-specific Ab, e.g. IgG
  - Input-DNA as processed chromatin (X-linked, fragmented)
  - Input-DNA as unprocessed chromatin (genomic DNA)
- Many control options should be tested in parallel.
  - Often Mock IP or using non-specific Ab works best
    - Similar low complexity as specific IP
    - Same concentration range – less error for DNA-normalization
    - Low (or similar) bias during amplification

**Fixed complexity (by method)**

**Variable complexity**

AFFYMETRIX
The Way Ahead.™

# Outline of this Talk

- The Challenge
- ChIP experimental design
- **Affymetrix ChIP protocol**
  - **Sample material, amount of cells**
  - **Cross linking**
  - **DNA fragmentation (shearing)**
  - **Antibody titration, Enrichment analysis**
  - **DNA-Amplification**
  - **Fragmentation and Labeling**
  - **Hybridization; Wash & Stain**

# Main Message

## This protocol is not an integrated A-Z solution!!!

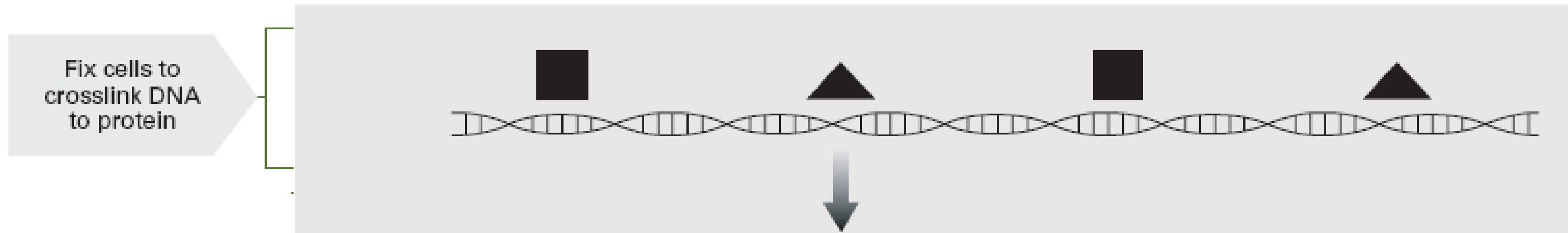
- The Affymetrix ChIP protocol was developed together with external group at Berkeley University (see Schwartz et al, Mark Biggin – group leader)
- Protocol was validated at Affymetrix by transcription factor binding experiments with Human (SP1, see Cawley et al. 2004) and also Mouse samples
- *Should work for low-mid complextity ChIP applications, higher binding frequencies only in smaller genomes*

# Sources of Material

- Suspended cell culture
- Attached cell culture
- Frozen tissues
  - Mouse embryos
  - Drosophila eggs
- Whole blood
- Number of cells
  - Established methods ~ $1x10^8$
  - Not practical for many applications
  - Tolerable ~ $1x10^6$ cells
  - 1000-10,000 cells using carrier chromatin (O'Neill et al, Nature Genet 2006)
  - Low number of cells leads to loss of enrichment

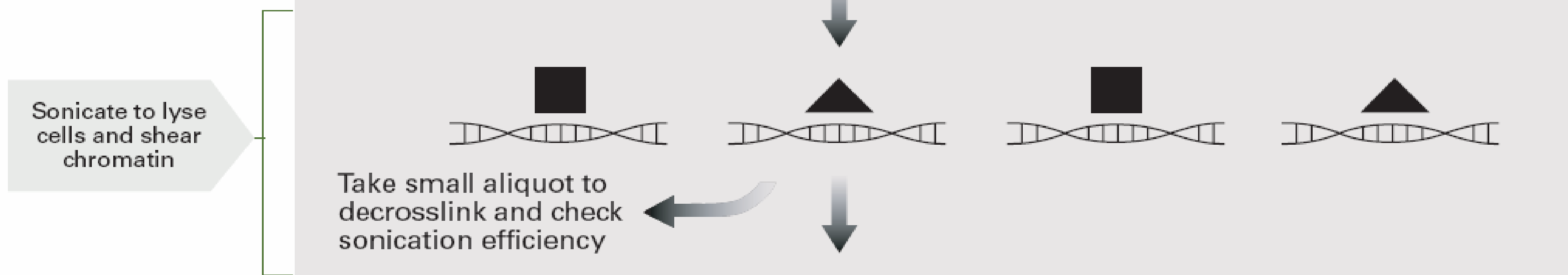
AFFYMETRIX
The Way Ahead.™

# Affy: Cell Preparation / Fixation / Wash



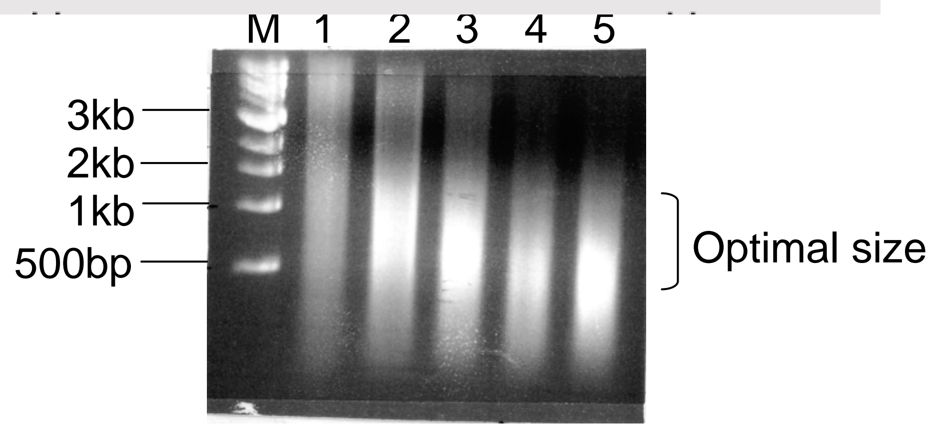

- Cell preparation
  - 5x10$^7$ /reaction
- Fixation
  - Fix in 1% Formaldehyde 10 min
  - Quench w/ 1/20 vol. of 2.5 M Glycine
- Washing
  - Wash with 1x PBS 2x
  - Wash 3x with Lysis buffer ( -80°C Freezing option)

# DNA Shearing (Fragmentation) Optimization




Fragmentation size can also influence PCR (bias effects)

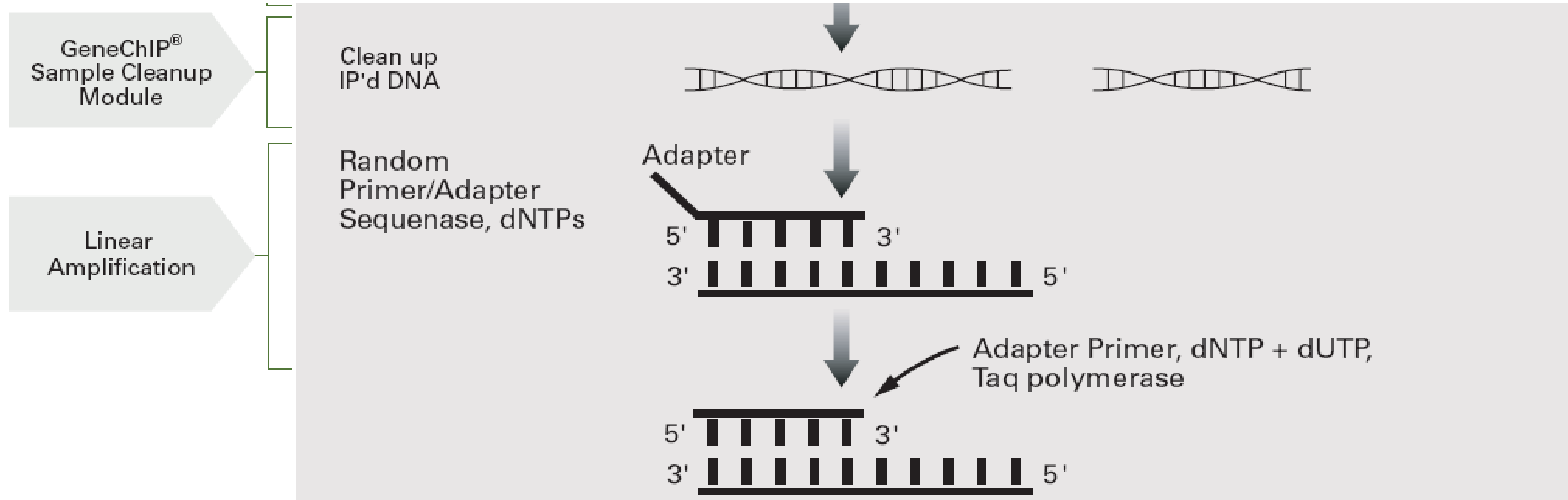
1: 2 x 5 sec
2: 2 x 10 sec
3: 3 x 10 sec
4: 3 x 20 sec
5: 3 x 30 sec




- Variables
  - Number and duration of pulses
  - Sonication needle specs, volume, tube shape and material
  - Highly cell type dependent
    - Salt concentrations (Over-fragm. found, lower Ca++ and Mg++ helped)
    - Additional enzyme (MNase) sometimes needed (Jurkat cells e.g.)

AFFYMETRIX® The Way Ahead.™

# DNA Amplification Step 1 – Linear Amplification

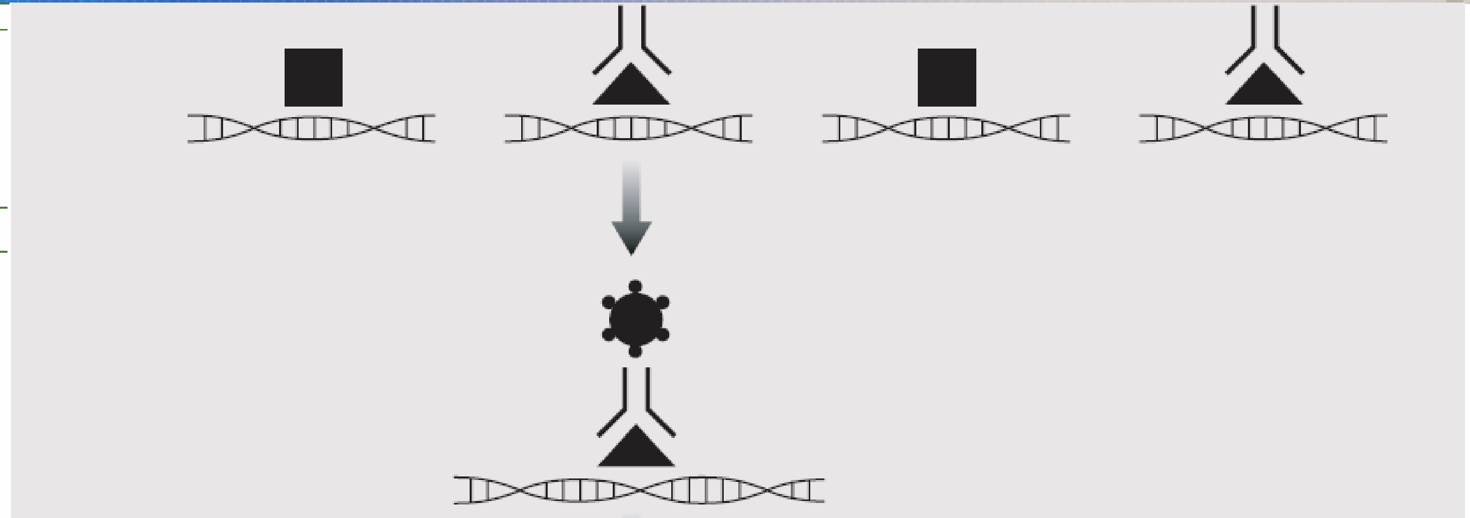



- Tested for low amount of IP-DNA (1-4 ng)
- Substrate inhibition possible
- Temperature gradient set up can be difficult with some Thermocyclers
- Size exclusion columns (Amersham/GEH, two rounds) needed to generate enough yield (remove primers, enrich larger fragments)
  - These columns are storage sensitive

AFFYMETRIX

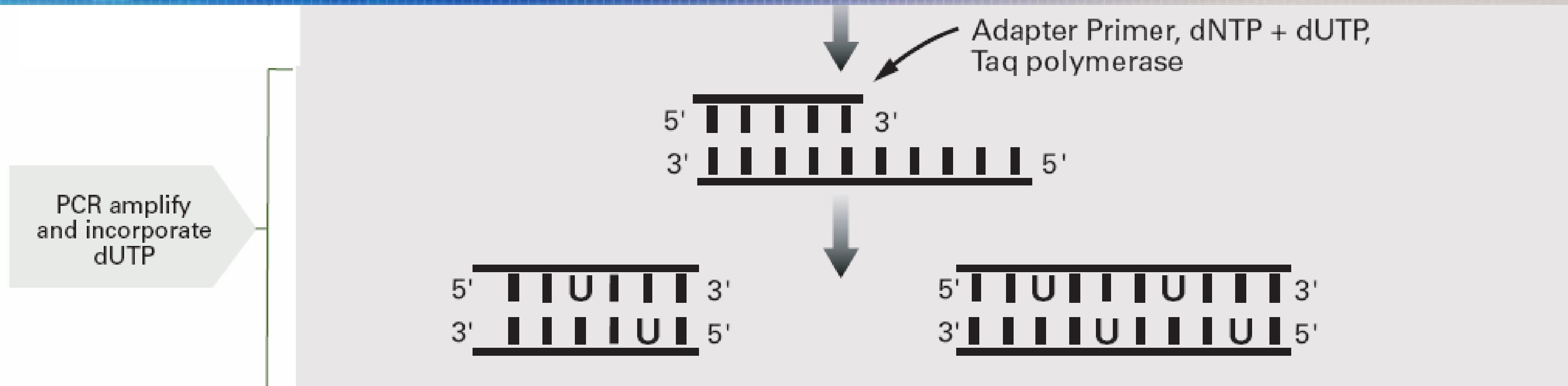
The Way Ahead.™

# Pre-Clearing of Chromatin with Beads




- Part of Affymetrix protocol
- Designed to reduce non-specific interactions of beads with chromatin
  - Some scientists consider it unnecessary
  - Magnetic beads are a good alternative

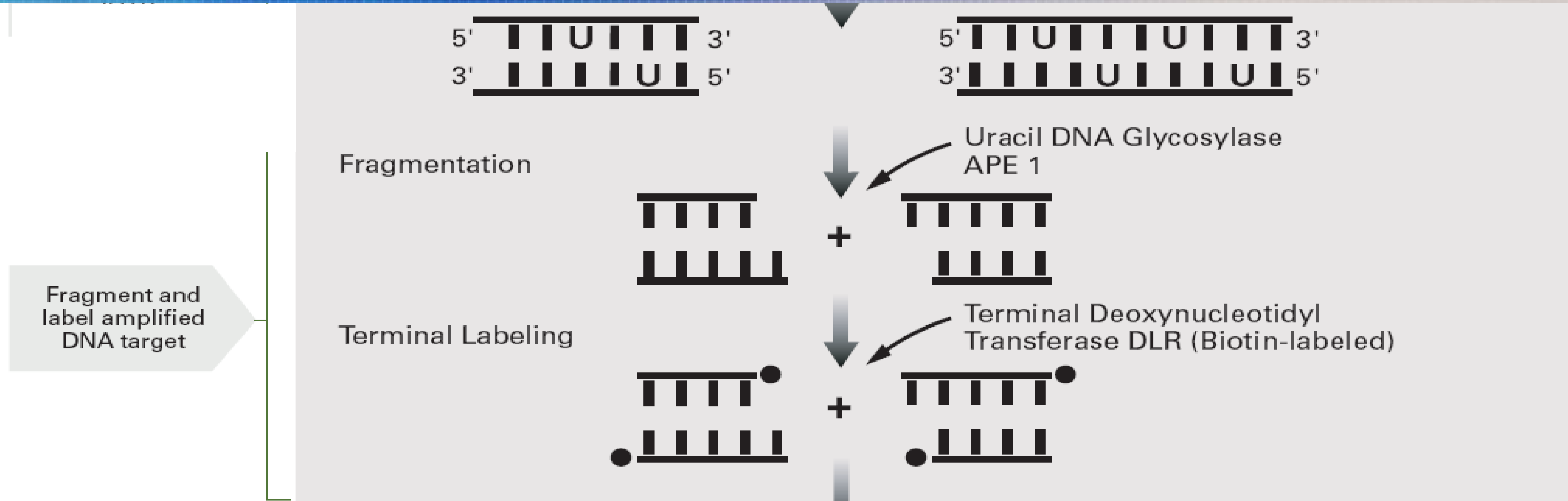
# DNA Amplification
# Step 2: PCR + dUTP Incorporation




- Problems discussed earlier (complexity-bias; saturation)
- dUTP incorporation prepares for fragmentation reaction
- dUTP can inhibit modern Taq polymerases with side activities (proof reading, often not fully documented)

AFFYMETRIX
The Way Ahead.™

# Fragmentation & Labeling

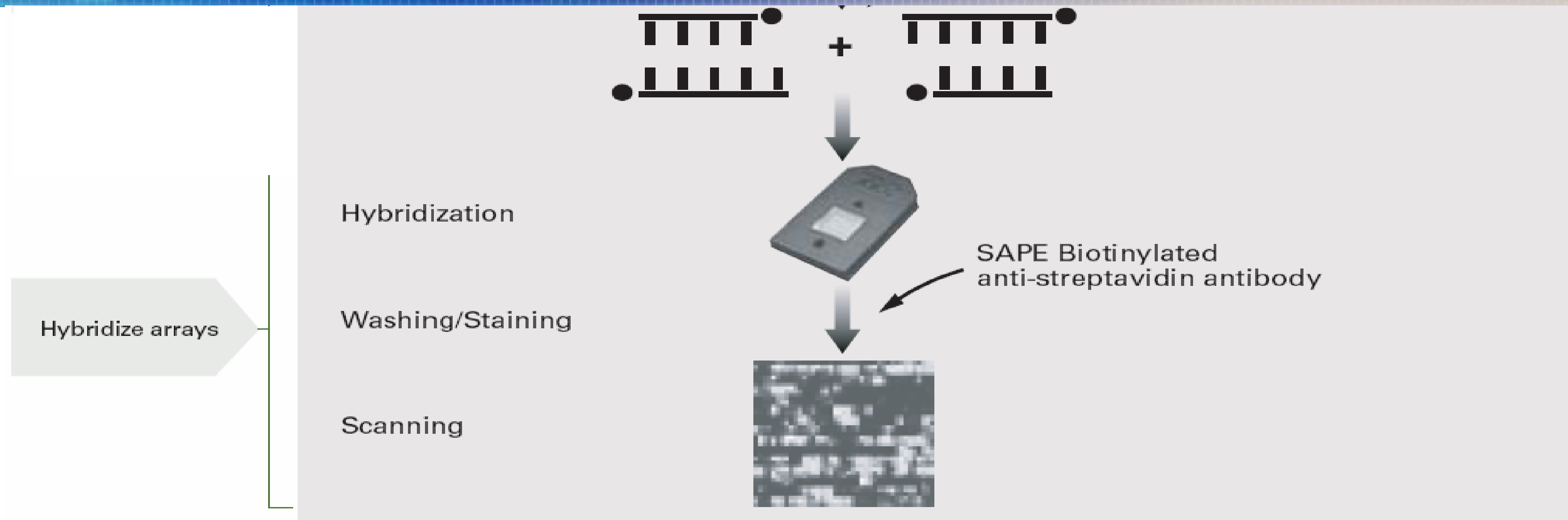



- Fragmentation reaction
  - DsDNA (7-9 ug)
  - Adjust enzymes for DNA amount
  - 37°C 1hour, 93°C 2min
- Labeling reaction
  - Fragmented dsDNA + DLR
  - 37°C 1hour, 70°C 2min

If amplified IP DNA is already available without dUTP used:

- DNA fragmentation & labeling reagents (for 100K SNP-mapping e.g.) can be used as alternative

# Hybridization; Wash & Stain




- Amount of fragmented and labeled dsDNA might vary according to complexity and reached enrichment (0.5 – 20 µg)
- Affymetrix Hyb, Wash&Stain kit should be used (w/ additional background blocker)
- 7G GCS-3000 needed ; Specific Fluidics protocols provided

# ChIP Assay Optimization (Summary)

- Does intended ChIP application fits to the Affymetrix protocol (complexity…) or do you need to adapt amplification method?
- Sonication conditions of fixed cells
- Antibody titration
  - Antibody affinity and specificity can vary, so the amount of antibody may to be titrated to achieve optimal sample enrichment
- Enrichment analysis needs to be established
  - Repeat after amplification again
- PCR amplification of IP-DNA
  - The optimal number of PCR cycles may require optimization to avoid saturation and ensure that the IP enrichment is maintained
- Adjust amount of amplified dsDNA for fragmentation, labeling and hybridization