diagenode

Innovating Epigenetics Solutions

WELCOME TO DIAGENODE

ChIP Workshop

Jessica Apulei & Juri Kazakevych

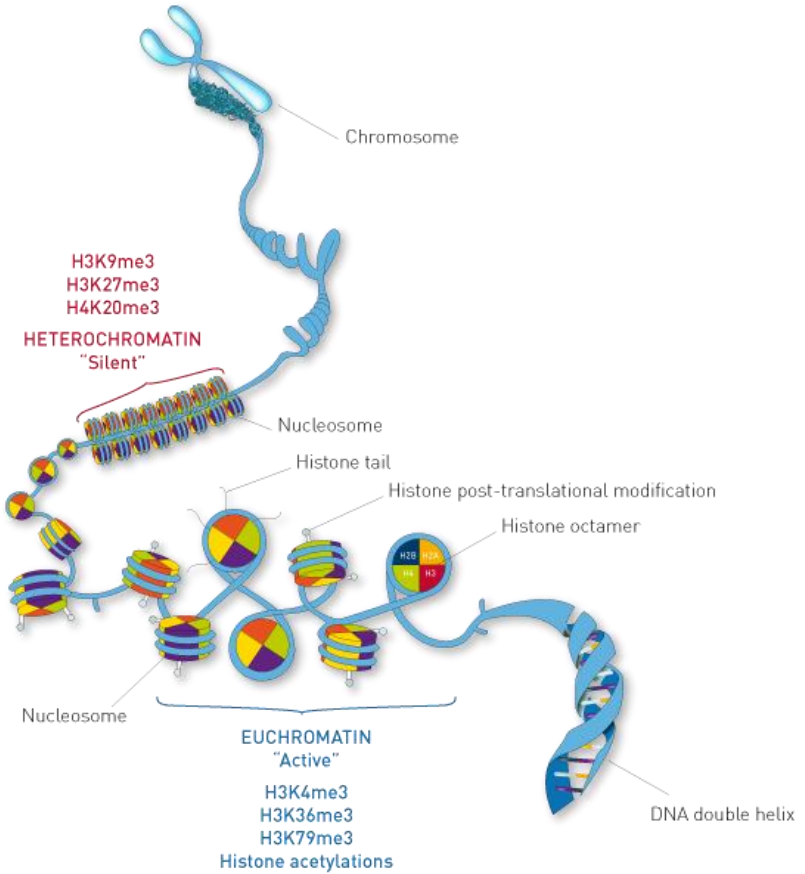
8-9 December 2020

OBJECTIVES

- ChIP overview
  - Chip-qPCR vs. ChIP-Seq
- Chromatin preparation
  - Fixation, Cell lysis and Chromatin shearing
- Setting up IP
  - Antibodies, Replicates, inputs, controls
- ChIP-qPCR
- ChIP-seq: library prep & sequencing
- ChIP-seq: analysis
- Overview of Alternate methods

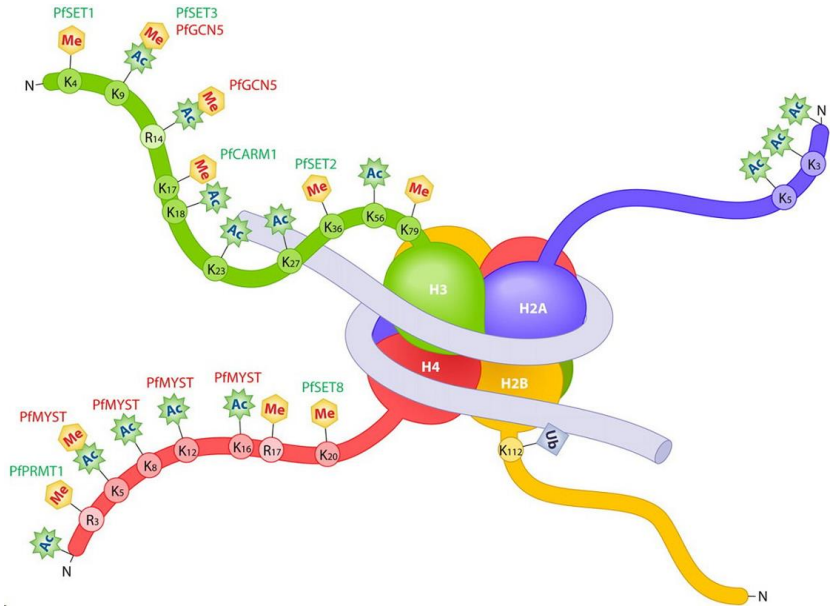
What is Chromatin?

A complex of DNA and proteins found in eukaryotic cells



What is Chromatin?

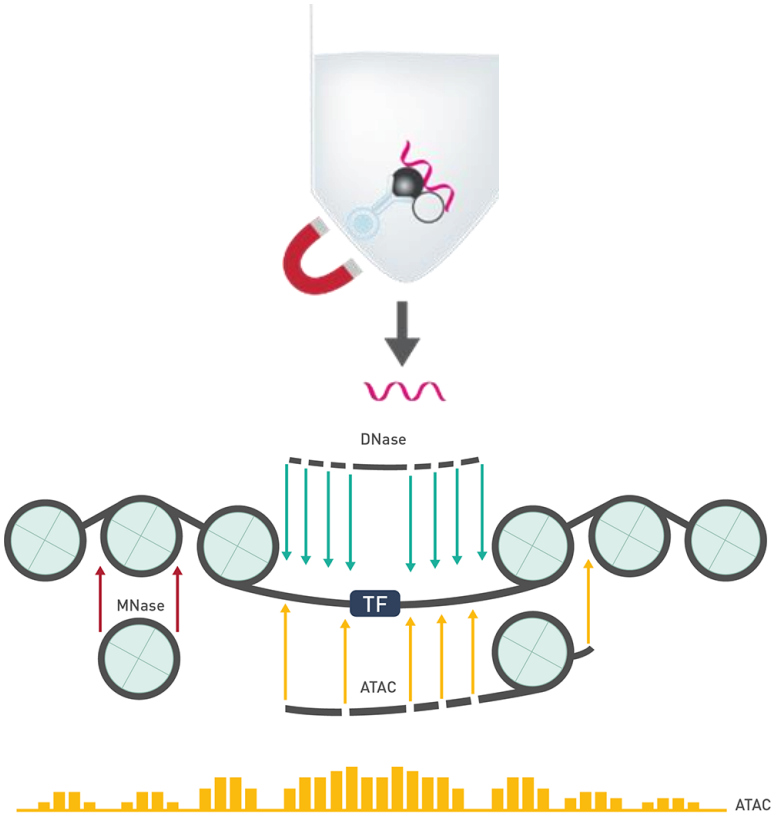
- Identifying genome-wide DNA binding sites for histones, transcription factors and other proteins
- Defines transcription factor binding sites
- Reveals gene regulatory networks in combination with RNA sequencing and methylation analysis



CHROMATIN ANALYSIS

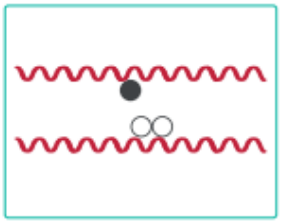
- Interaction between proteins and DNA (immuno- assays)
  - ChIP-qPCR
  - ChIP-Seq
  - Cut&Run and Cut&TAG

- Methods to study chromatin accessibility (non-immuno assays)
  - ATAC-seq



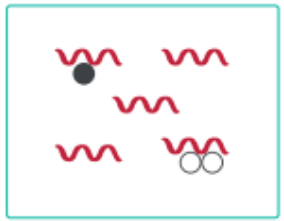
Workflow: Chromatin ImmunoPrecipitation (ChIP):

Step 1



Cross link to fix proteins to DNA

Step 2

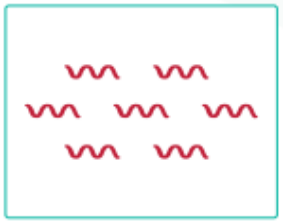


Shear chromatin

Step 3

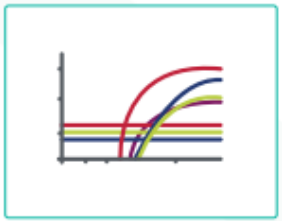
Immunoprecipitate with antibody and magnetic beads

Step 4



Reverse crosslinks and purify

Step 5



Analyze by qPCR

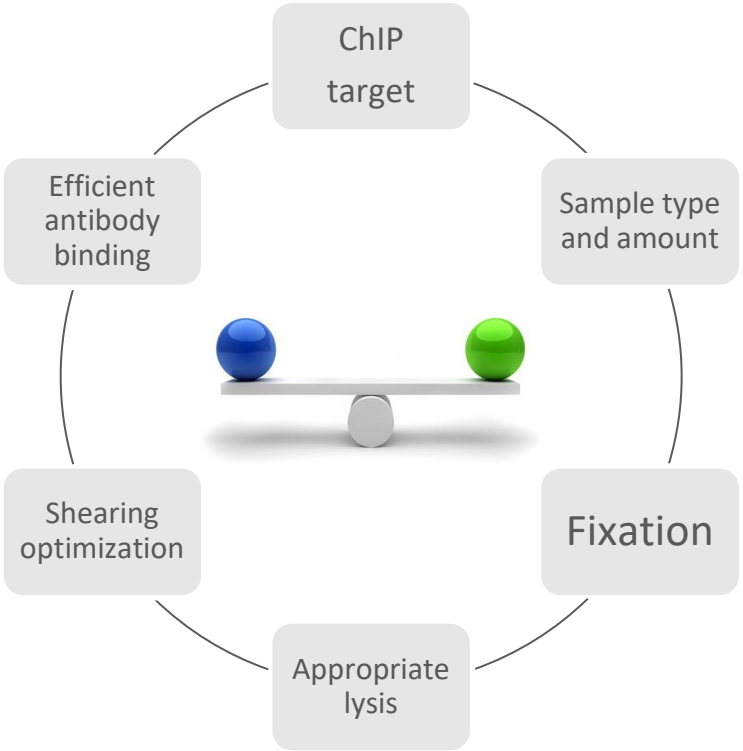
Step 6

Prepare the libraries for NGS

ChIP-qPCR or ChIP-Seq?

ChIP-qPCR	ChIP-Seq
Single-locus data QC step for ChIP-seq	Genome-wide data suitable for exploratory analysis
Low-cost Fast	High sequencing costs Longer protocol High sequencing turnaround time

Summary – Tips to Prepare Good Chromatin



Guide for successful chromatin preparation using the Bioruptor® Pico

https://www.diagenode.com/files/protocols/bioruptor-pico-chromatin-preparation-guide.pdf

Starting material: Cells and Tissues

- **Cells**
  ChIP: 1 million/IP for histones, 4 million/IP for TF (less depending on histone/TF)
  low-input ChIP: 10k/IP for histones

- **Tissues**
  Amount – 20-30 mg/IP
  Dounce homogenization for soft tissues (e.g. liver or brain)
  Bead beater like TissueLyser for hard fibrous frozen tissues (e.g. muscles)

- **FFPE tissue**
  Challenging due to extensive crosslinking
  Heptane instead of xylene for de-paraffinization -> easier, non-toxic workflow

Fixation

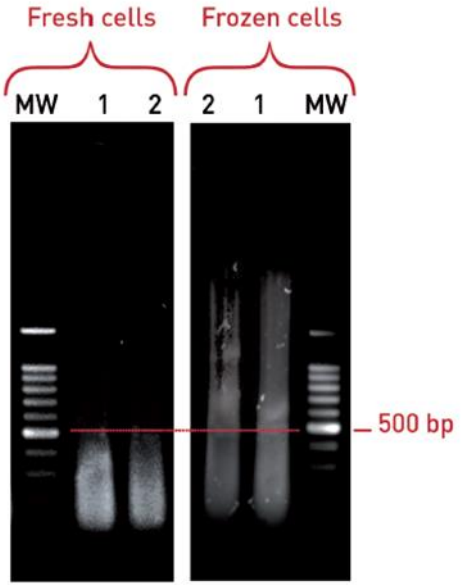
- Covalent stabilization of protein-DNA interactions; Reversible
  - Directly in medium for weak or rare protein-DNA interaction
  - For histone marks, cells can be resuspended by trypsinization before fixation

- Common fixative: Formaldehyde
  - Fresh
  - Methanol-free not mandatory

Target / Fixator	Formaldehyde	ChIP Cross-link Gold C01019021
Histones	Yes (8-10 min)	No need
Transcriptional factors directly bound to DNA	Yes (10-20 min)	No need
Indirect higher order and/or dynamic interactions	Yes (10-15 min)	Yes (30-45 min)

Cell Lysis

- Two step lysis – standard protocol, difficult cells
  - Remove soluble cytosolic proteins first
  - Improves sonication efficiency
  - Reduces background
- One step lysis – for low cell numbers
  - Lyse cells directly with an SDS-containing buffer
- Tips/Tricks/Critical steps:
  - Incubate on ice to start lysis and get narrower fragments size
  - Centrifuge to remove soluble membranes and cytosol
  - Avoid freezing chromatin if possible



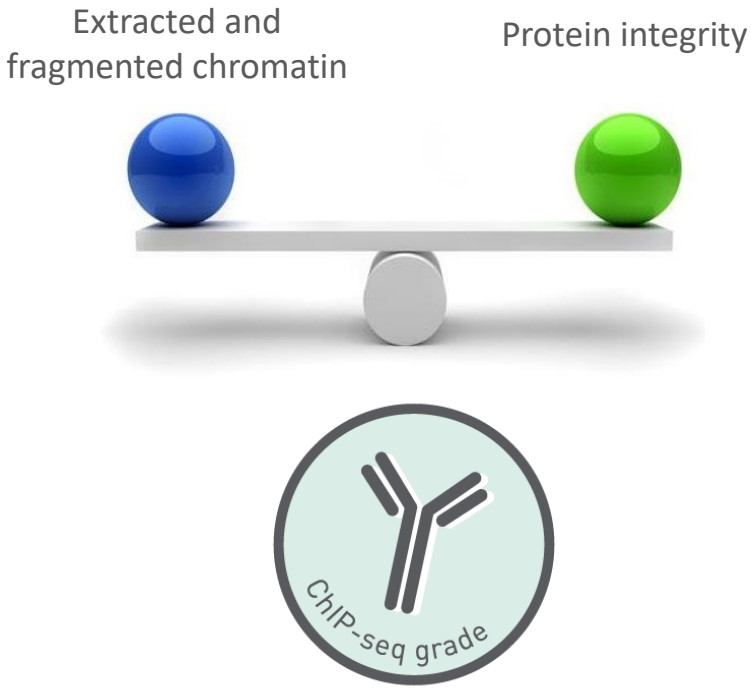
Stopping Points

- **Cells**
  - Fix cells, lyse, isolate & shear chromatin -> freeze
  - Fix cells, lyse, isolate chromatin -> freeze
  - Fix cells -> freeze

- **Tissues**
  - Fix tissue, lyse, isolate & shear chromatin -> freeze
  - Freeze prior fixation

Secrets of ChIP Success

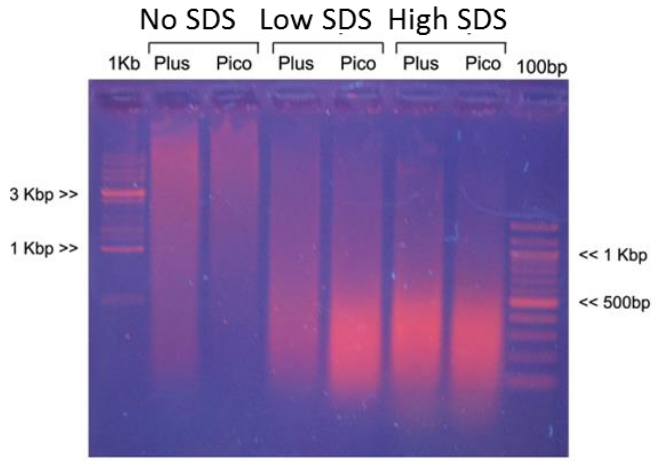
- Prepare “good” chromatin
  - Suitable fragment size and available epitopes
- Use a good antibody at the right concentration
- Optimize for highest specific signal and the lowest background

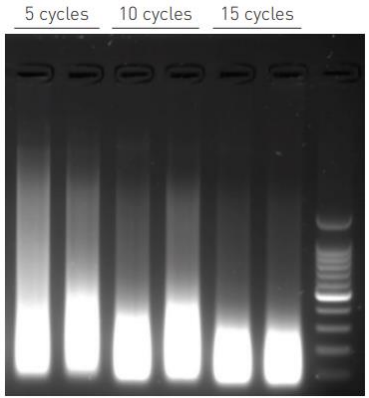


Chromatin Shearing

- 100-800 bp fragments, peak 200-500bp
- Use a good sonicator
  - Gentle - not to dislodge protein
  - Uniform and reproducible energy
  - Temperature control at 4°C
  - Multiplex and easy to use

- Shearing buffer with detergents, preferably SDS
  - Increase sonication efficiency and chromatin yield
  - Improve epitope availability
  - Balance shearing and downstream IP
- Sample concentration
- Select the shortest time resulting in efficient shearing

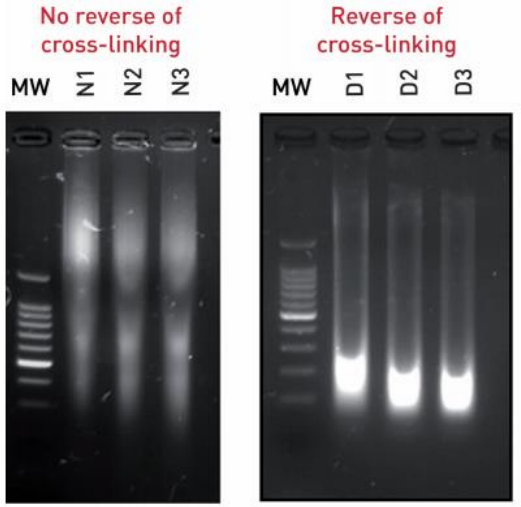




Bioruptor Pico

Analyzing Fragment Size

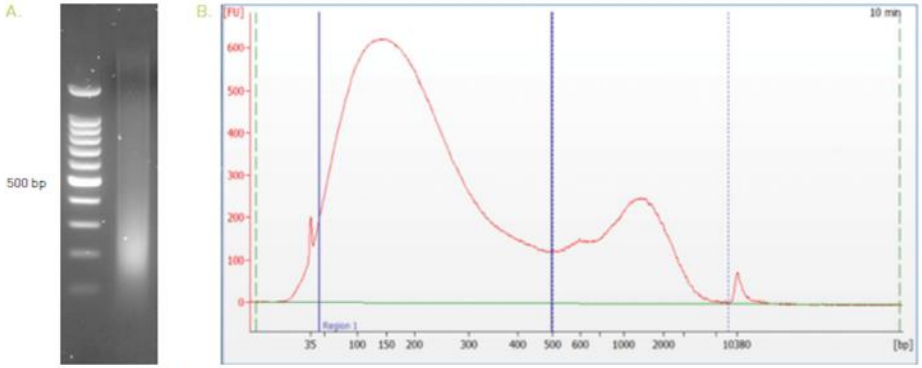
- **De-crosslink**
  - Residual crosslinking retards migration
- **RNase treatment**
  - reduces background
- **DNA purification**
  - IPure beads + DiaMag magnetic rack
  - **Low inputs:** DiaPure columns (eluted in 6 µl)
- **Electrophoretic analysis**
  - 1.2 - 1.8% agarose gel
  - 300 ng or 60k cells per lane
  - **Low inputs:** FragmentAnalyzer, 2k cells



Analyzing fragment size

- Use agarose gel or fragment analyzer

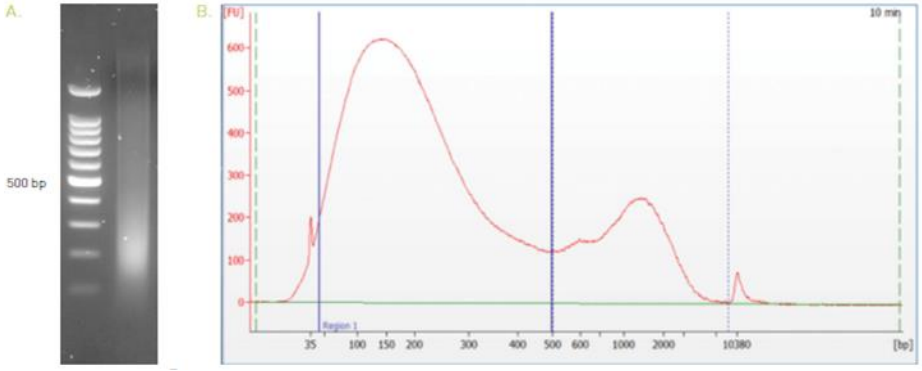
- Bioanalyzer or Tapestation:
  - Over-representation of HMW fragments
  - Log-based -> visual misinterpretation of fragment distribution
  - More sensitive to overloading, incomplete reverse crosslinking and residual contaminants

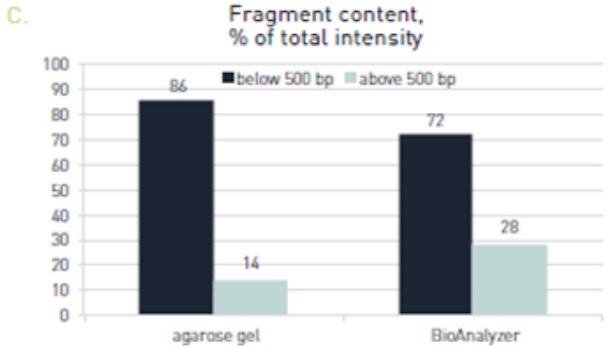


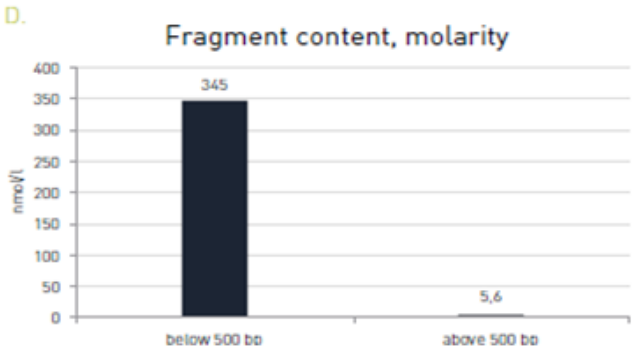
Analyzing fragment size

- Use agarose gel or fragment analyzer

- Bioanalyzer or Tapestation:
  - Over-representation of HMW fragments
  - Log-based -> visual misinterpretation of fragment distribution
  - More sensitive to overloading, incomplete reverse crosslinking and residual contaminants



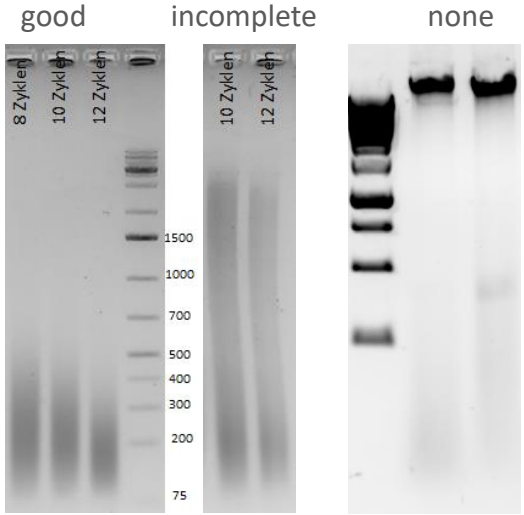




Troubleshooting Chromatin Shearing

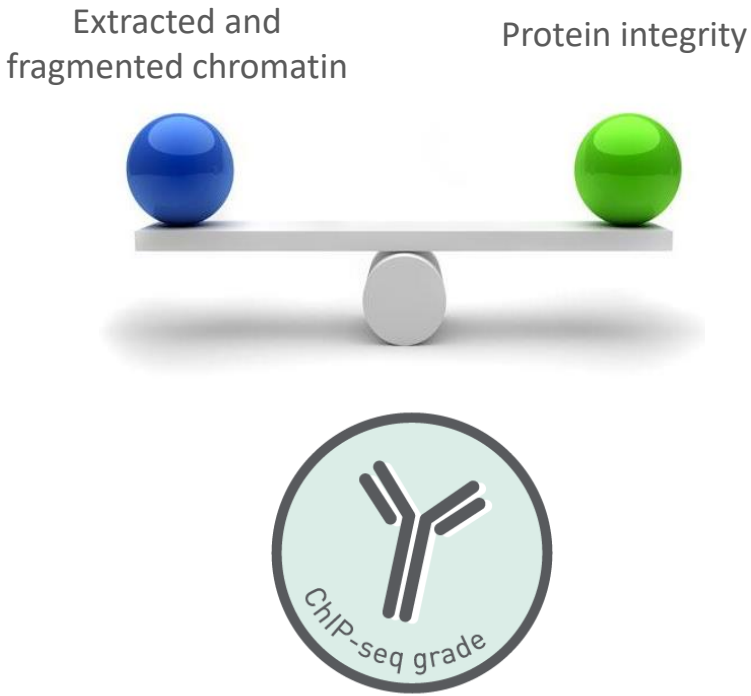
- **No shearing at all**
  - Incomplete lysis – check buffer composition
  - Check instrument efficiency - QC test on sonicator

- **Incomplete shearing**
  - Over-fixation: check fixative and duration
  - Too high cell density
  - Changes in sample require adjustment of shearing protocol
    - Fresh vs. Frozen chromatin
    - Different sample types
  - Wrong consumables (tubes)
  - Sample out of sonication focus
    - droplets on walls/lid of tube
    - Wrong sample volume
  - Wrong temperature (should be 4°C for chromatin)
  - None of the above? -> Check instrument efficiency - QC test on sonicator

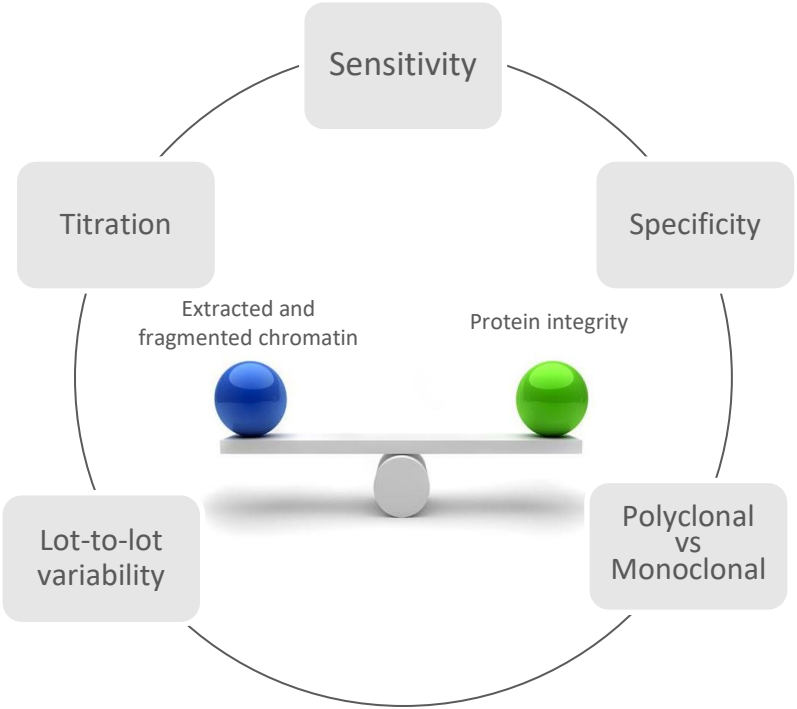


Secrets of ChIP Success

- Prepare “good” chromatin
  - Suitable fragment size and available epitopes
- Use a good antibody at the right concentration
- Optimize for highest specific signal and the lowest background



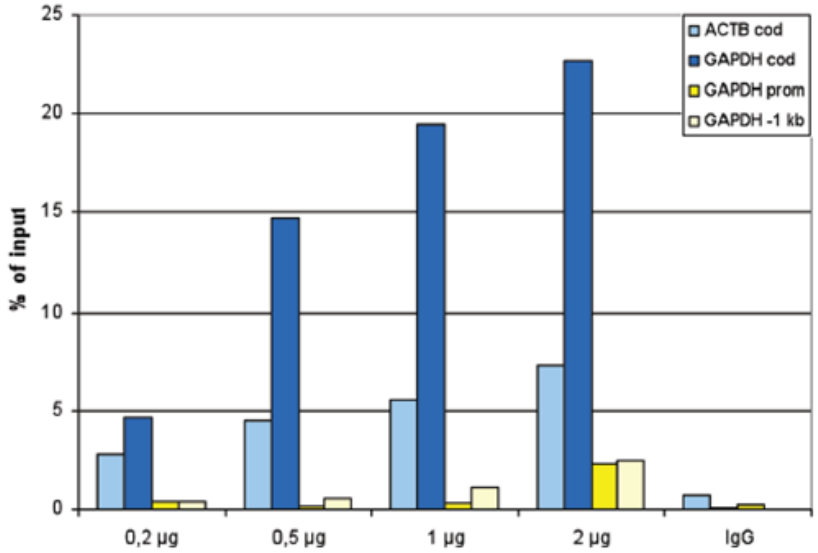
Antibodies for ChIP

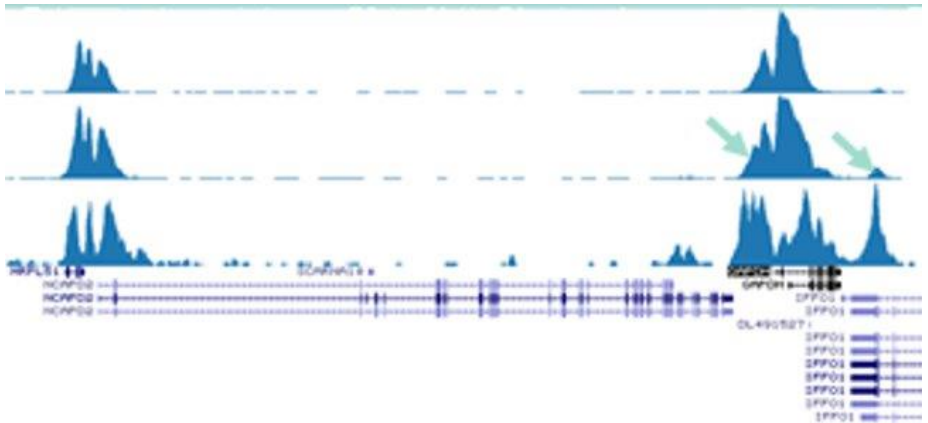


https://www.diagenode.com/en/categories/chip-seq-grade-antibodies

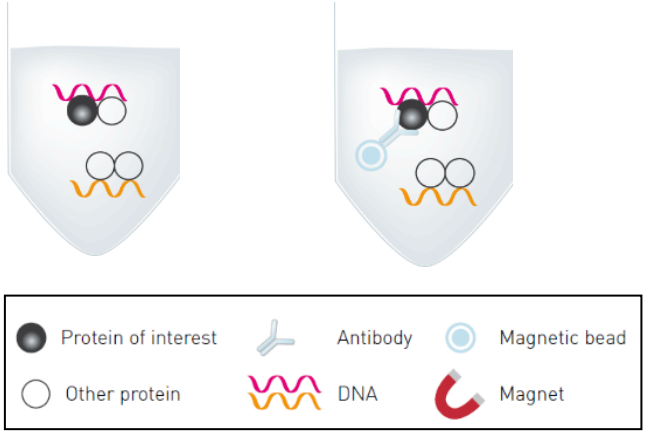
Antibodies for ChIP

Polyclonal H3K36me3 antibody titration (Diagenode C15410192)
1µg IgG as negative IP control
Chromatin from 100.000 cells





Setting up IP



Components:

- Sheared chromatin
- **ChIP grade antibodies**
  -> optimized quantity
- ProteinA/G magnetic beads
- ChIP buffer
- Protease inhibitor cocktail

Antibodies for ChIP – What beads?

Agarose beads

Required: centrifuge

Sensitive to handling

High background

Risk of carry-over

Magnetic beads

Required: magnetic rack

Robust

Low background

Easy separation

Limit antibody amounts to bead capacity!

Protein G or A beads

- Both bind to IgG antibodies and are structurally similar
- Slightly different affinities for IgG subclasses across different species.
- Use appropriate depending on the IgG subtype you are using:

Protein A

Rabbit
Pig
Dog
Cat

Protein G

Mouse
Rat
Human

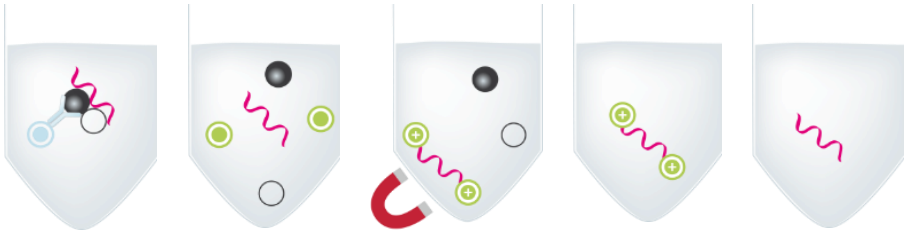
Setting up IP: Input Sample

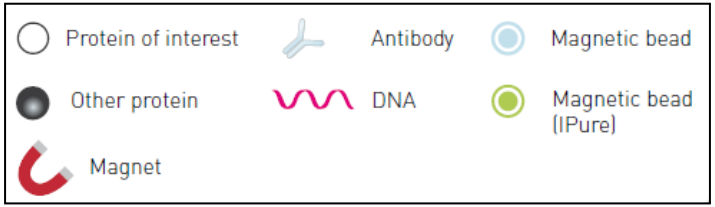
- Fraction of sheared chromatin is kept aside as INPUT
  - Processed in parallel with IP-samples from reversed crosslinking
  - Include one input for each chromatin sample
- Key reference for ChIP-qPCR and ChIP-seq analysis
- ChIP-qPCR: used to calculate the recovery (% of input)
- ChIP-seq: mandatory for bio-informatics analysis
  - Normalization for mappability of a region, avoid duplication bias etc.
  - Input pooling can be considered for ChIP-seq on very similar samples

Setting Up IP: Additional Controls

- Positive control (H3K4me3, CTCF):
  - Confirm overall efficiency of ChIP workflow
  - ChIP optimization for new target
- Negative Control (IgG)
  - Measure of non specific IP background
  - Include one negative IgG control in each series of ChIP reactions
- Not necessary to sequence these but good control for qPCR
- Biological Replicates
  - ChIP-qPCR ≥3
  - ChIP-seq ≥2

ChIP Protocol – Elution, de-crosslinking and DNA isolation



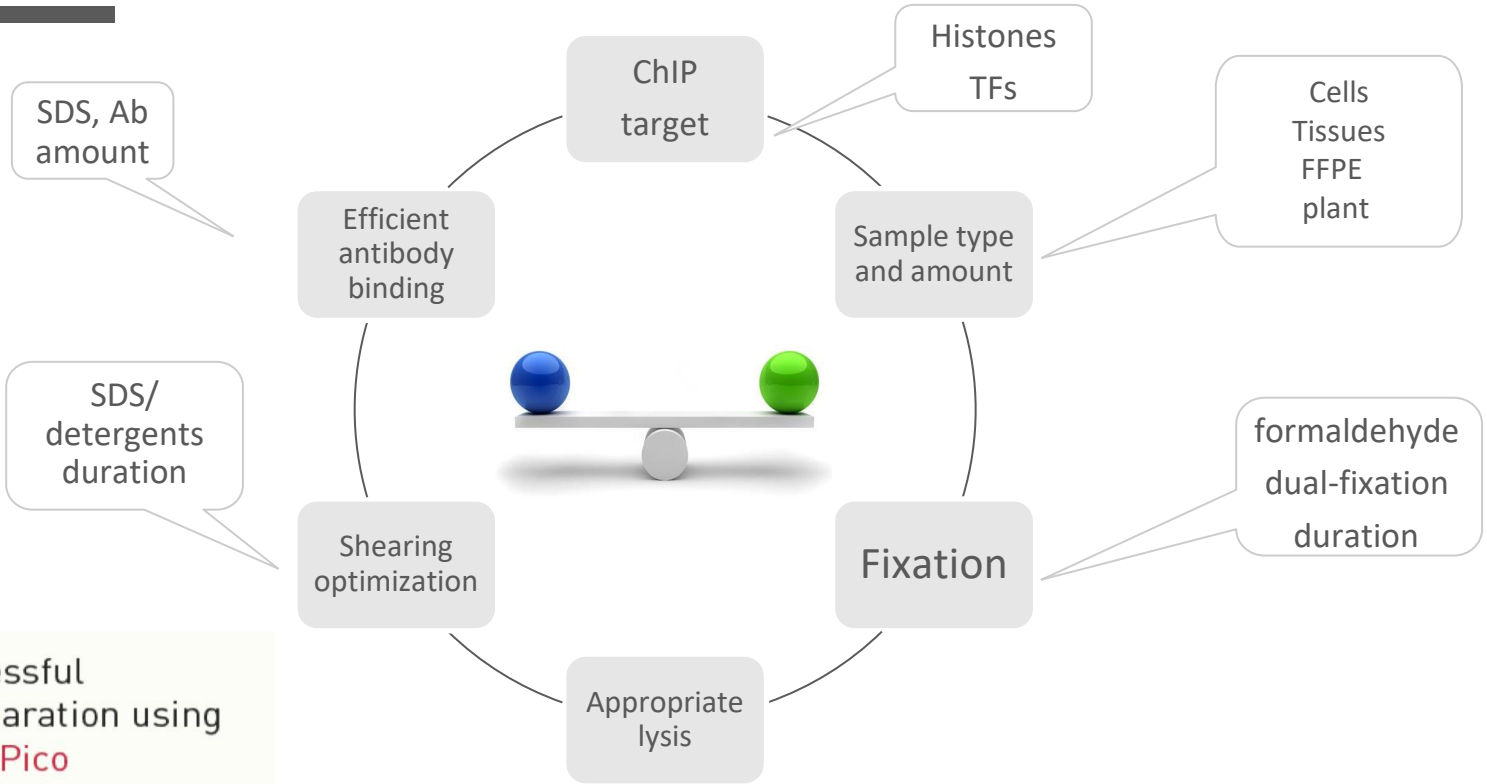


- Elution of the chromatin complexes from protein A/G-bound magnetic beads:
  elution buffer 30 min at RT

- Reversal of cross-links:
  Incubation for at least 4h at 65°C

- Isolation of the ChIP'd DNA:
  IPure magnetic beads
  Column purification (DiaPure columns for low elution volumes >6µl)

Summary – Tips to Prepare Good Chromatin



Guide for successful chromatin preparation using the Bioruptor® Pico

https://www.diagenode.com/files/protocols/bioruptor-pico-chromatin-preparation-guide.pdf

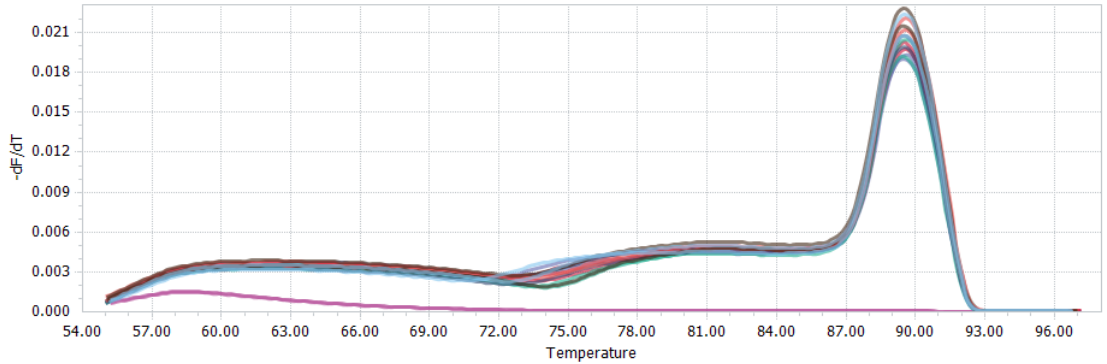
ChIP-qPCR

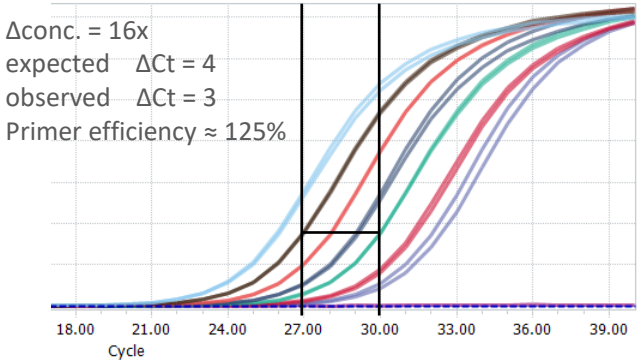
- Target & Primer selection is key for ChIP-qPCR
- Predict qPCR-targets from ChIP-seq data
- qPCR as QC prior ChIP-seq
- If no ChIP-seq data:
  estimate binding from similar data, biological function etc.
  -use multiple regions
- PCR program depends on Master Mix, qPCR system and primer pairs

Setting up ChIP-qPCR

- **Target primer design:**
  - Place primers around binding site
  - 50-150bp amplicons
  - 20-30 bp primers with a similar Tm between 55° and 60°C

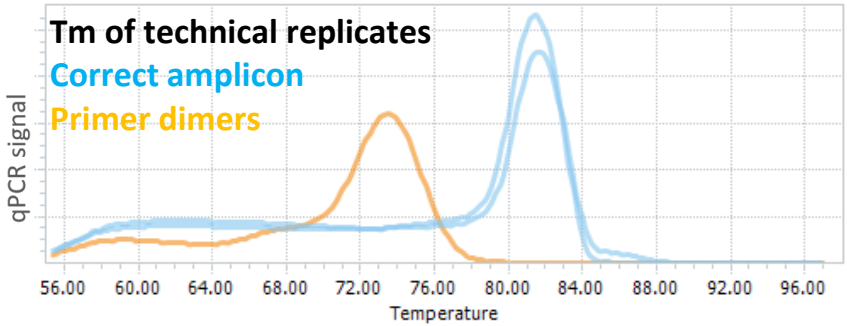
- **Primer pair validation:**
  - Check on gDNA/input for T_m profile
  - Check efficiency (95-105% acceptable)





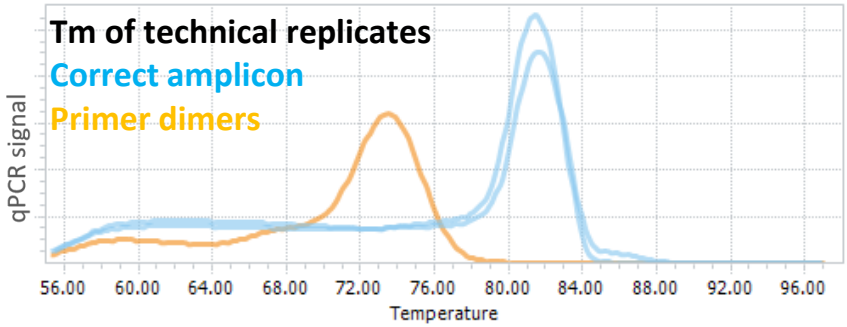
Setting up ChIP-qPCR

- **QC**
  - T_m: no second peaks, no primer dimers
  - technical replicates within 0.3 Ct
  - Ct values >>30 are often not reliable

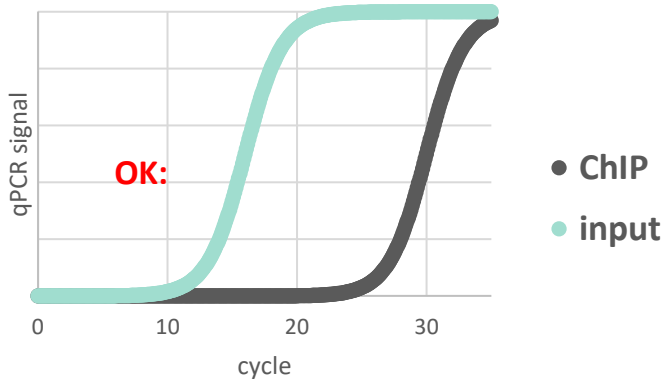


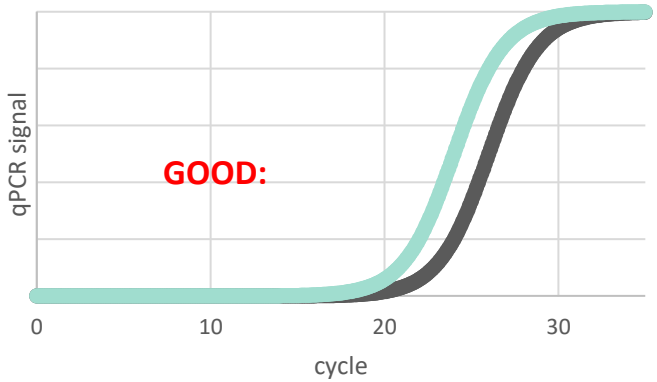
Setting up ChIP-qPCR

- **QC**
  - T_m: no second peaks, no primer dimers
  - technical replicates within 0.3 Ct
  - Ct values >>30 are often not reliable



- **ChIP and input samples**
  - Adjust amount of ChIP-sample/input to obtain comparable Ct values
  - Consider primer efficiency for high ΔCt





ChIP-qPCR analysis

ChIP recovery *R*:

- chromatin recovery as % of Input
- R should be minimal for the IgG control and high for the epitope of interest

$$R = \frac{2^{Ct\,(input)} - Ct(ChIP)}{100\,(input\ fraction)}$$

-> for each target separately

Input fraction is often corrected with a logarithmic compensatory factor, e.g. -6.64 Ct for 1% input

ChIP fold-enrichment *F* with ΔΔCt method:

- fold-enrichment of bound vs. epitope-"free" regions
- S varies depending on regions analyzed

$$F = \frac{R(positive\ region)}{R(negative\ region)}$$

-> main success parameter

Successful ChIP?

- If wrong regions targeted – risk of false negative result
- ***F*** > 2 for ChIP-qPCR analysis
- ***F*** > 4 for ChIP-seq

ChIP-qPCR Exercise

Conditions: untreated (A), treated (B)
Targets: positive control (e.g. GAPDH) (P), genes of interest (G1, G2) intergenic region (N)
Antibodies: IgG, H3K4me3

Ct values

A

	IgG		H3K4me3		Input (1% of sample)	
P	34.0	36.0	26.0	26.1	27.0	27.1
G1	35.0	35.0	27.0	27.2	26.0	25.7
G2	-	37.0	33.0	34.0	29.0	29.6
N	34.0	35.0	33.0	33.5	28.0	28.1

B

	IgG		H3K4me3		Input (1% of sample)	
	-	36.4	25.4	25.5	26.5	26.6
	35.4	36.4	26.4	26.6	25.4	25.1
	34.4	33.4	28.7	28.5	28.6	28.7
	35.4	-	32.4	33.4	27.4	27.5

ChIP-qPCR Exercise

Conditions: untreated (A), treated (B)
Targets: positive control (e.g. GAPDH) (P), genes of interest (G1, G2) intergenic region (N)
Antibodies: IgG, H3K4me3

Ct values

A

	IgG		H3K4me3		Input (1% of sample)	
P	34.0	36.0	26.0	26.1	27.0	27.1
G1	35.0	35.0	27.0	27.2	26.0	25.7
G2	-	37.0	33.0	34.0	29.0	29.6
N	34.0	35.0	33.0	33.5	28.0	28.1

B

	IgG		H3K4me3		Input (1% of sample)	
	-	36.4	25.4	25.5	26.5	26.6
	35.4	36.4	26.4	26.6	25.4	25.1
	34.4	33.4	28.7	28.5	28.6	28.7
	35.4	-	32.4	33.4	27.4	27.5

1. Technical sanity check

- values out of range
- high Ct-variation (>0.3) between technical replicates
- A/B inputs shifted
- Ct(H3K4me3) ≈ Ct(input)

ChIP-qPCR Exercise

Conditions: untreated (A), treated (B)
Targets: positive control (e.g. GAPDH) (P), genes of interest (G1, G2) intergenic region (N)
Antibodies: IgG, H3K4me3

2. **Averaging Technical replicates**

	A			B		
	IgG	H3K4me3	Input (1% of sample)	IgG	H3K4me3	Input
P	35.0	26.1	27.1	36.4	25.5	26.6
G1	35.0	27.1	25.9	35.9	26.5	25.3
G2	37.0	33.5	29.3	33.9	28.6	28.7
N	34.5	33.3	28.1	35.4	32.9	27.5

3. **Biological Sanity Check**

- Ct(P) < Ct(N) for H3K4me3
- Ct(H3K4me3) << Ct(IgG)

ChIP-qPCR Exercise

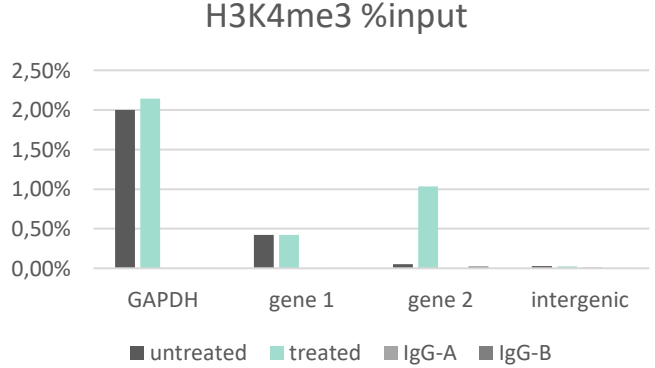
Conditions: untreated (A), treated (B)
Targets: positive control (e.g. GAPDH) (P), genes of interest (G1, G2) intergenic region (N)
Antibodies: IgG, H3K4me3

	A			B		
	IgG	H3K4me3	Input (1% of sample)	IgG	H3K4me3	Input
P	35.0	26.1	27.1	36.4	25.5	26.6
G1	35.0	27.1	25.9	35.9	26.5	25.3
G2	37.0	33.5	29.3	33.9	28.6	28.7
N	34.5	33.3	28.1	35.4	32.9	27.5

4. **Recovery (% of input)**

$$R = \frac{2^{Ct\,(input)} - Ct(ChIP)}{100\,(input\ fraction)}$$

	A	IgG-A	B	IgG-B
P	2.00%	0.00%	2.14%	0.00%
G1	0.42%	0.00%	0.42%	0.00%
G2	0.05%	0.00%	1.04%	0.03%
N	0.03%	0.01%	0.02%	0.00%



ChIP-qPCR Exercise

Conditions: untreated (A), treated (B)
Targets: positive control (e.g. GAPDH) (P), genes of interest (G1, G2) intergenic region (N)
Antibodies: IgG, H3K4me3

H3K4me3 % of input

	A	B
P	2.00%	2.14%
N	0.03%	0.02%

5. **ChIP fold-enrichment**

$$F_A = \frac{R(P_A)}{R(N_A)} = 74x\ enrichment$$

$$F_B = \frac{R(P_B)}{R(N_B)} = 94x\ enrichment$$

ChIP-qPCR Exercise

Conditions: untreated (A), treated (B)
Targets: positive control (e.g. GAPDH) (P), genes of interest (G1, G2) intergenic region (N)
Antibodies: IgG, H3K4me3

H3K4me3 % of input

	A	B
P	2.00%	2.14%
N	0.03%	0.02%

5. **ChIP fold-enrichment**

$$F_A = \frac{R(P_A)}{R(N_A)} = 74x\ enrichment$$

$$F_B = \frac{R(P_B)}{R(N_B)} = 94x\ enrichment$$

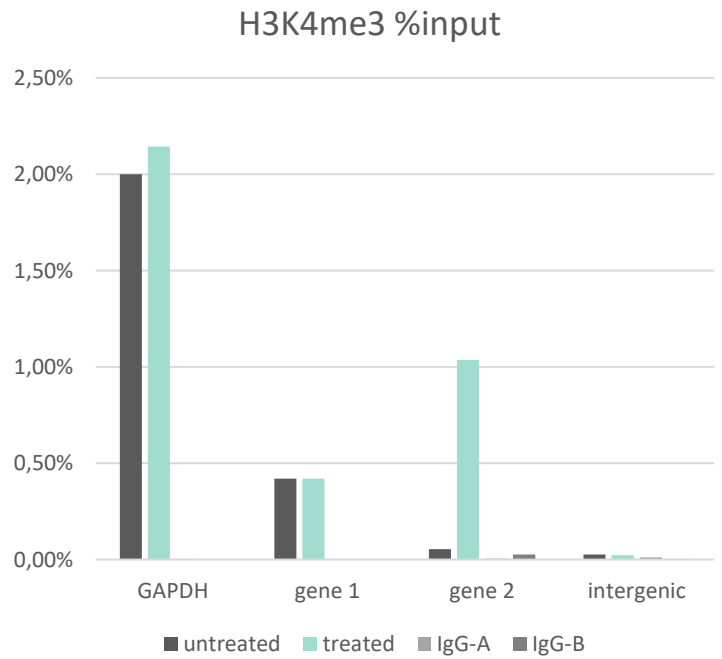
7. **Assessment**
   - ***F***>2 ChIP-qPCR qualified
   - ***F***>4 ChIP-seq qualified

8. **Optimization**
   - if ***R*** or ***F*** are low, optimize ChIP parameters

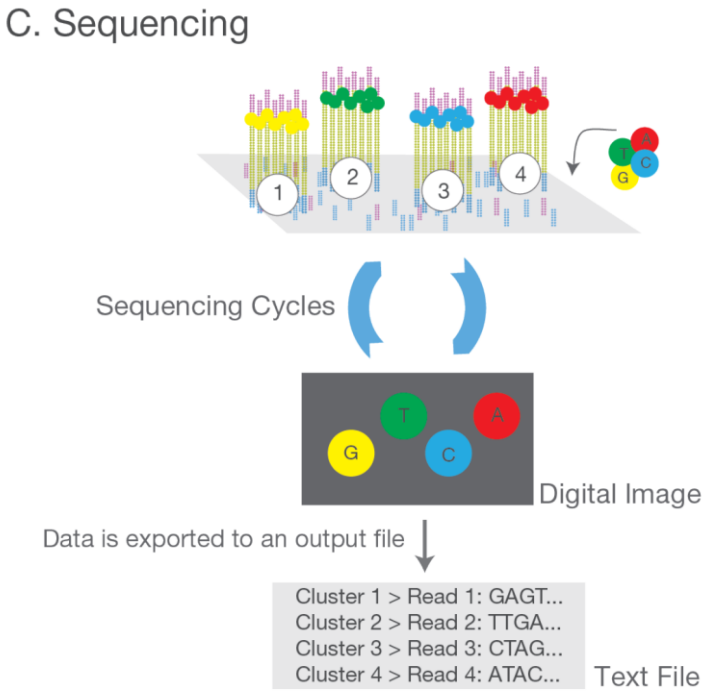
DIAGENODE

ChIP-qPCR Analysis summary

- Determine ChIP-recovery and fold-enrichment
- input used as reference to calculate ChIP-recovery
- Each ChIP-target requires specific control regions
- Suitable control regions can vary among samples



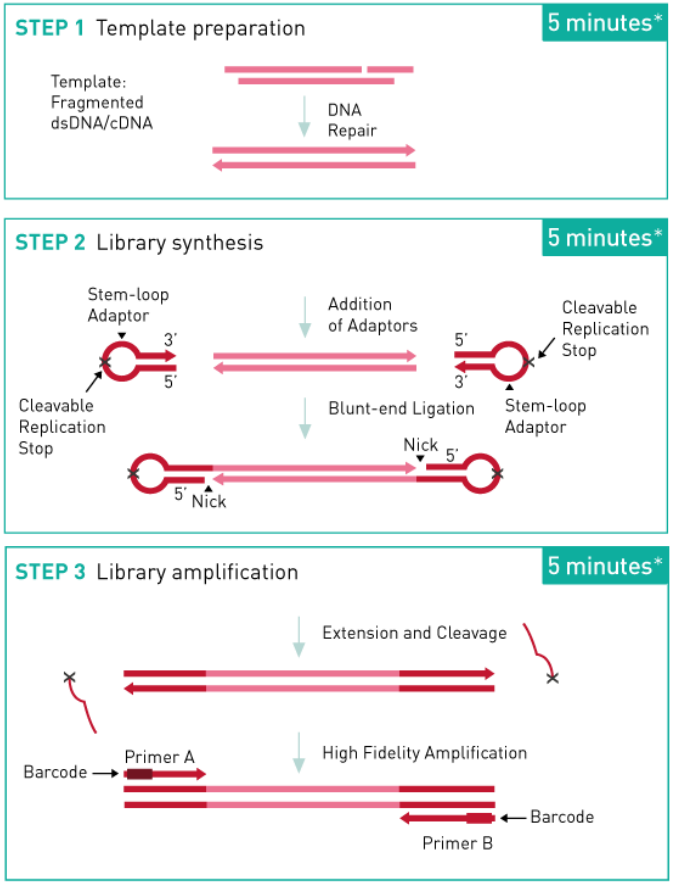
ChIP-seq: library prep



ChIP-seq: Library Prep

- Low input
- Minimal steps
  - To maximize recovery
- Sensitive
  - Minimal PCR amplification
- Suitable for pooling

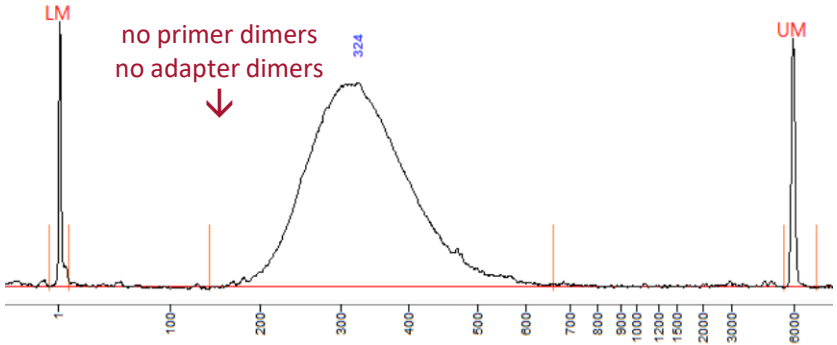
MicroPlex kit workflow



* hands-on-time

Library Pooling

- **Determine library size**
  - Bioanalyzer or Fragment Analyzer
  - Identify adapter dimers or unexpected library sizes
- **Quantify**
  - Qubit
  - qPCR –quantify sequencable library
  - Convert from ng/µl to nM using average library size
- **Dilute and Pool normalized libraries**
  - Same size for best clustering

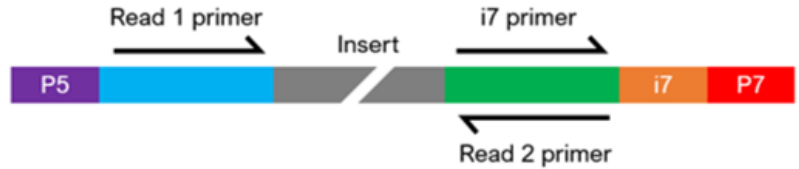


Benefits of Multiplexing

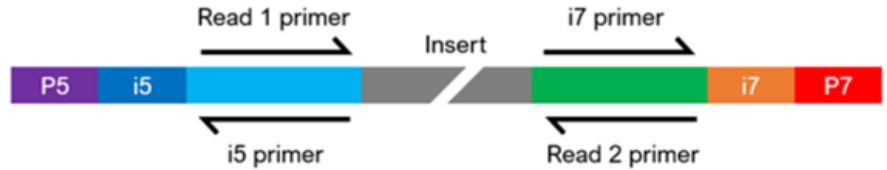
- **Fast High-Throughput Strategy:**
  - Large sample numbers can be simultaneously sequenced
- **Cost-Effective Method:**
  - Reduces time and reagent use
  - Cluster detection more efficient with different bases in beginning of read
- **Simplified Analysis:**
  - Automatic sample identification with "**barcodes**“ using Illumina software

Single and Dual-indexed Libraries

- **Single-index sequencing**
  - Low level of multiplexing
- **Dual indexing**
  - Higher multiplexing - more samples per lane possible
  - Higher accuracy of sample identification
- Unique dual indexing (UDI)
  - Allows filtering of index-hopping events



Dual-Indexed Sequencing



ChIP-seq: Sequencing Settings

Read length

- 50 bp sufficient for most ChIPs
- adjust fragment-size to read length

Sequencing depth

- mainly set by samples/flow cell and flow cell type
- 30 M reads for sharp peaking targets e.g. H3K4me3, H3K27ac
- 50 M for broadly distributed and abundant targets e.g. H3K27me3
- use same depth for input

Replicates

- ≥ duplicates
- increased replicate number will improve sensitivity of the downstream analysis

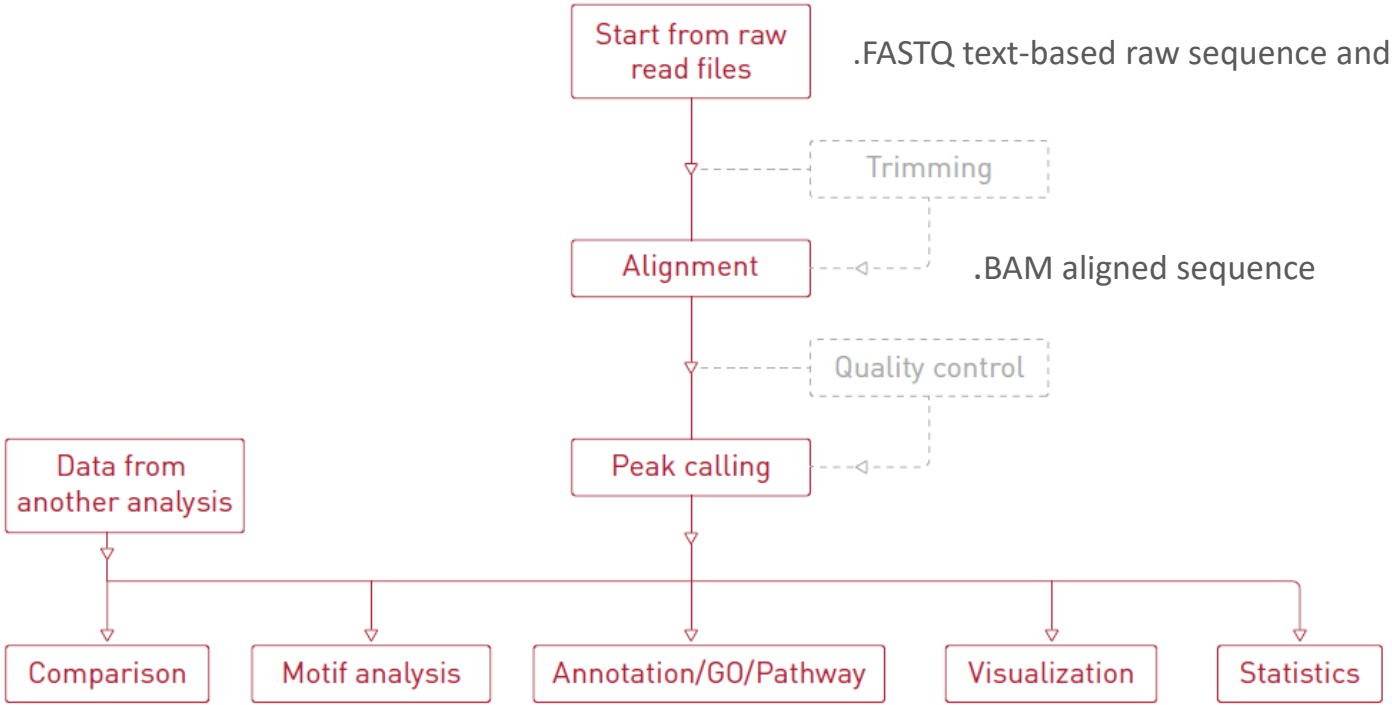
Input sequencing

- one input per sample is gold standard
- pooling inputs from replicates can often be considered

ChIP-seq: Analysis

Bioinformatician	R free-ware and online tool kits	https://www.r-project.org/ www.bioinformatics.babraham.ac.uk/projects/seqmonk/ biit.cs.ut.ee/gprofiler/gost
Wet-lab expert with free time	standard bio-informatic services free-ware and online tool kits	https://www.diagenode.com/en/categories/Services Comprehensive Multi-Omic and bio-info services www.bioinformatics.babraham.ac.uk/projects/seqmonk/ initial & advanced data analysis, genome browser, graphical presentation of data https://biit.cs.ut.ee/gprofiler/gost Functional profiling tool
no expertise or no free time	advanced bio-informatic services	https://www.diagenode.com/en/categories/Services Comprehensive Multi-Omic and bio-info services

ChIP-seq: Analysis



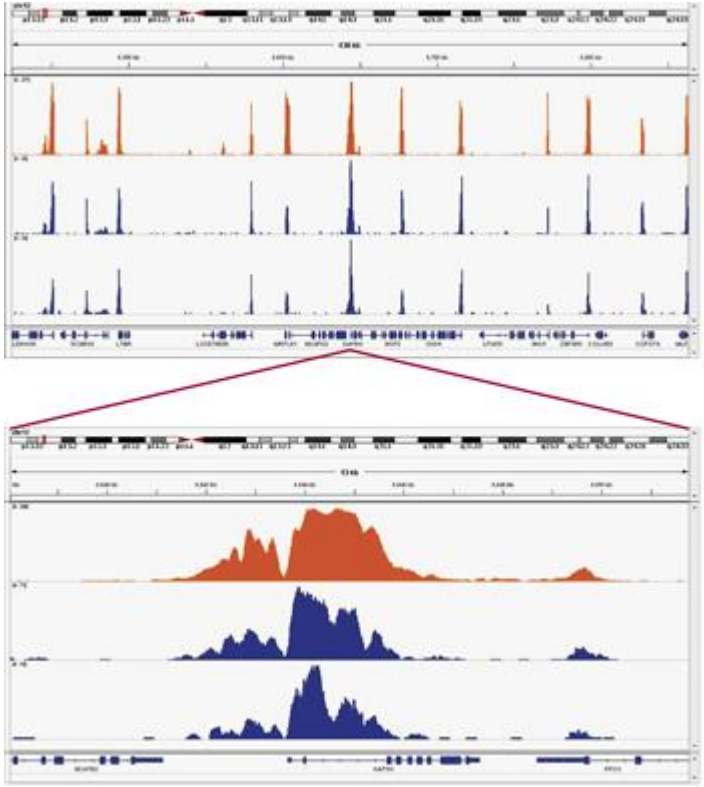
ChIP-seq: Analysis

Standard bioinformatic analysis:

- **alignment to reference genome**
- **peak calling**

Advanced bioinformatic analysis:

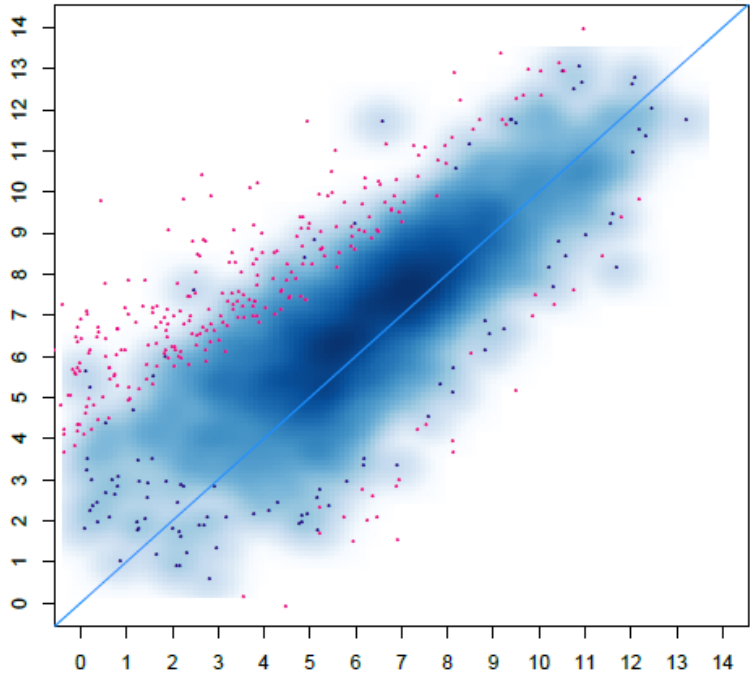
- **annotation of peaks and genes**
- differential analysis of peak/gene lists
- unsupervised analysis (PCA, clustering)
- functional enrichment analysis
  (e.g. Pathway analysis, Gene ontology)
- Machine learning
- integrative analysis
  (RNA-seq, ATAC-seq, more ChIP-seq targets)
- publication-ready Visualization of genomic regions



ChIP-seq: Analysis

Standard bioinformatic analysis:
- alignment to reference genome
- peak calling

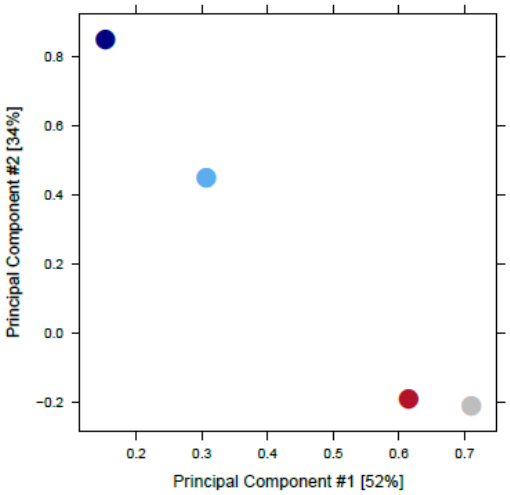
Advanced bioinformatic analysis:
- annotation of peaks and genes
- **differential analysis of peak/gene lists**
- unsupervised analysis (PCA, clustering)
- functional enrichment analysis
  (e.g. Pathway analysis, Gene ontology)
- Machine learning
- integrative analysis
  (RNA-seq, ATAC-seq, more ChIP-seq targets)
- publication-ready Visualization of genomic regions

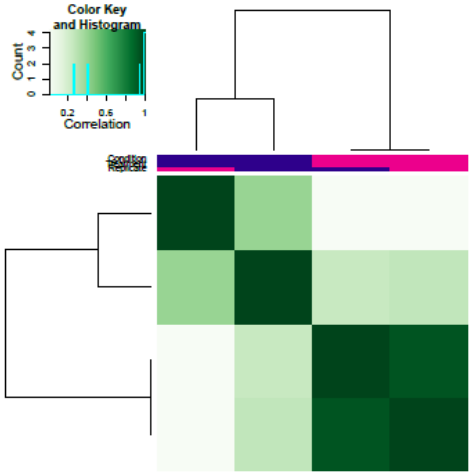


ChIP-seq: Analysis

Standard bioinformatic analysis:
- alignment to reference genome
- peak calling

Advanced bioinformatic analysis:
- annotation of peaks and genes
- differential analysis of peak/gene lists
- **unsupervised analysis (PCA, clustering)**
- functional enrichment analysis
  (e.g. Pathway analysis, Gene ontology)
- Machine learning
- integrative analysis
  (RNA-seq, ATAC-seq, more ChIP-seq targets)
- publication-ready Visualization of genomic regions

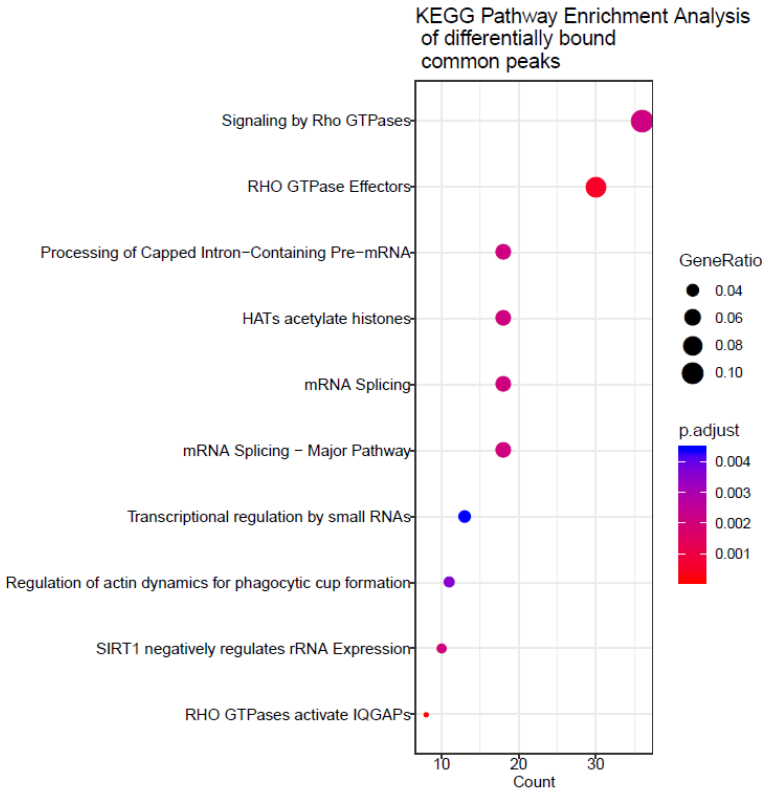




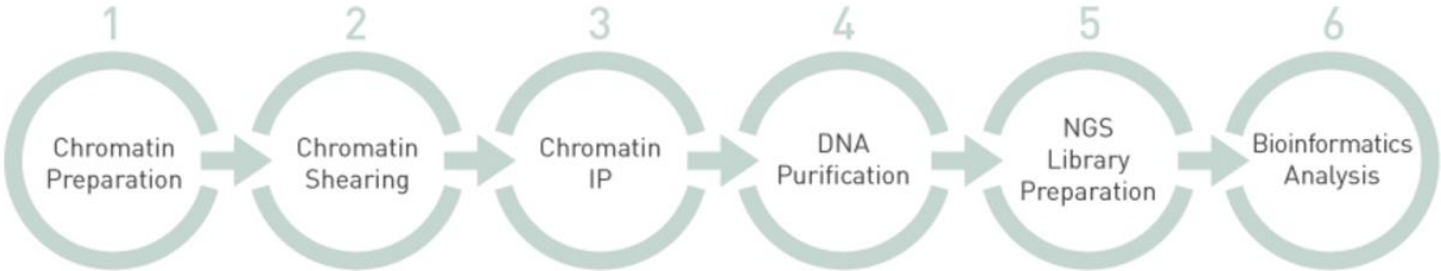
ChIP-seq: Analysis

Standard bioinformatic analysis:
- alignment to reference genome
- peak calling

Advanced bioinformatic analysis:
- annotation of peaks and genes
- differential analysis of peak/gene lists
- unsupervised analysis (PCA, clustering)
- **functional enrichment analysis**
  (e.g. Pathway analysis, Gene ontology)
- Machine learning
- integrative analysis
  (RNA-seq, ATAC-seq, more ChIP-seq targets)
- publication-ready Visualization of genomic regions



Summary



Solutions for each step of the way : Bioruptor, Kits, Antibodies, NGS library prep and Services

www.diagenode.com

Other Methods to Study Chromatin

- Methods to study the **interaction** between proteins and DNA (immuno- assays):
  - ChIPmentation
  - Cut&TAG

ChIPmentation™ & µChIPmentation™

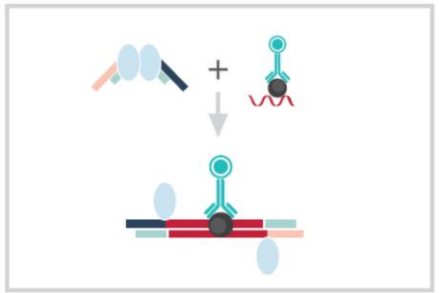
Easier and faster than classical ChIP-seq

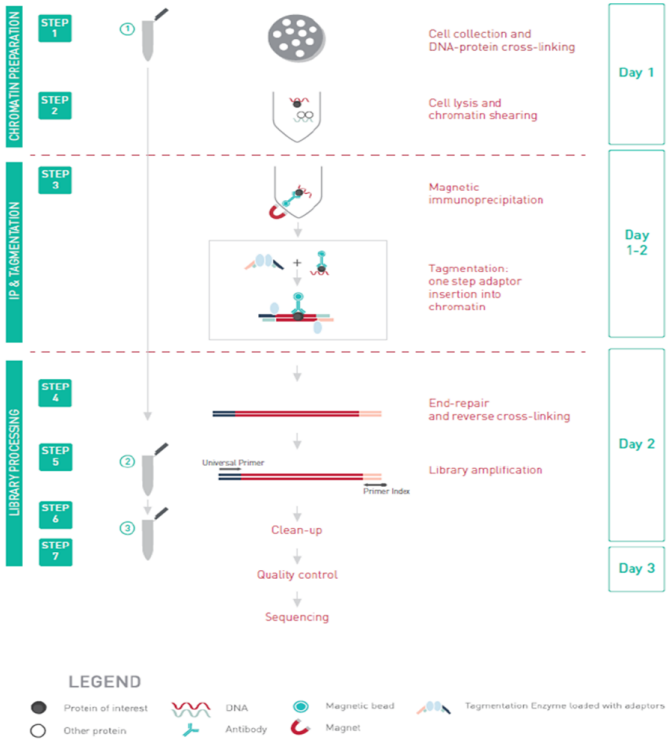
Validated for various histone marks

Ideal for analysis of large cohorts of samples (easy and fast)

Ideal for analysis of large number of marks on a unique sample

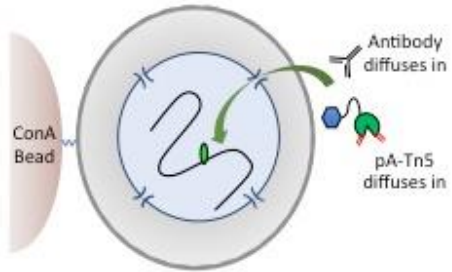
µChipmentation for 10,000 cells

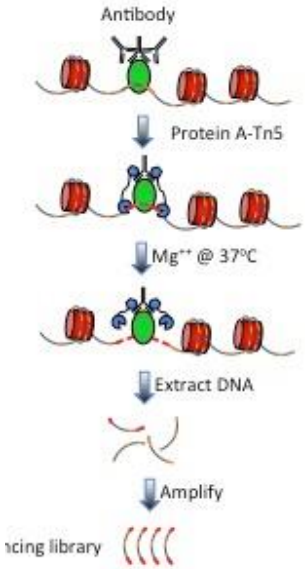




CUT&Tag: Cleavage Under Targets and Tagmentation

CUT&Tag (Cleavage Under Targets & Tagmentation)



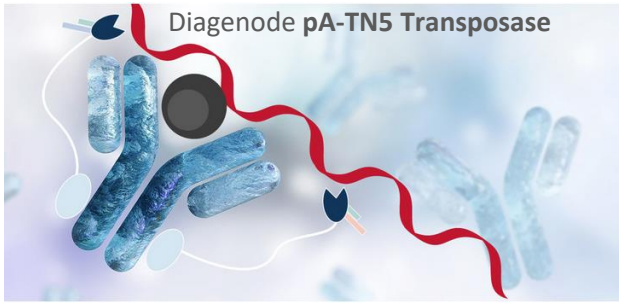


Key features:

Crucial reagent:
- proteinA-Tn5

Fast and easy protocol:
- fast tagmentation-based library prep
- No chromatin prep

Suitable for low cell numbers

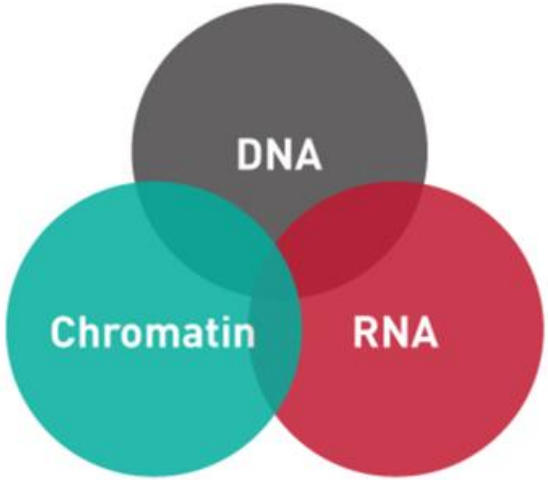


DIAGENODE SERVICE SOLUTIONS

EPIGENOMICS PROFILING SERVICES

15 YEARS OF EXPERTISE IN EPIGENETICS

- End-to-end epigenetic service and analysis
- Collaborative and customized project design
- Dedicated in-house expert for your project
- Presentation-quality data and graphs



CHIP WORKSHOP

THANK YOU!

Thank you for taking part in our ChIP workshops! (more coming)

Presentation will be sent to each participant

Watch for a little survey in your inbox – your feedback is invaluable

Stay in the know

Join our newsletter

@Diagenode

Spread the word
#EpiWorkshopsWithDiagenode
(more gifts for you :-))

customer.support@diagenode.com

www.diagenode.com