DNA methylation analysis

GENOME-WIDE

Immunoprecipitation

All species
(All methylated regions)

- MeDIP-seq
- MeDIP qPCR
- hMeDIP-seq
- hMeDIP qPCR
- MethylCap (MBD capture)

Plant compatible

WGBS

All species
(25 M CpGs in humans)

- Whole Genome Bisulfite Sequencing

Plant compatible

TARGETED GENOME

Target region

All species
(Client specifies # of CpGs)

- Amplicon sequencing
- Custom capture probes

Plant compatible

REDUCED GENOME

RRBS

All vertebrates
(~7-8M CpGs in human)

- Reduced Representation Bisulfite Sequencing

EPIC

Human
(850K CpGs)

- Illumina Infinium Methylation EPIC Array

Mouse
(285K CpGs)

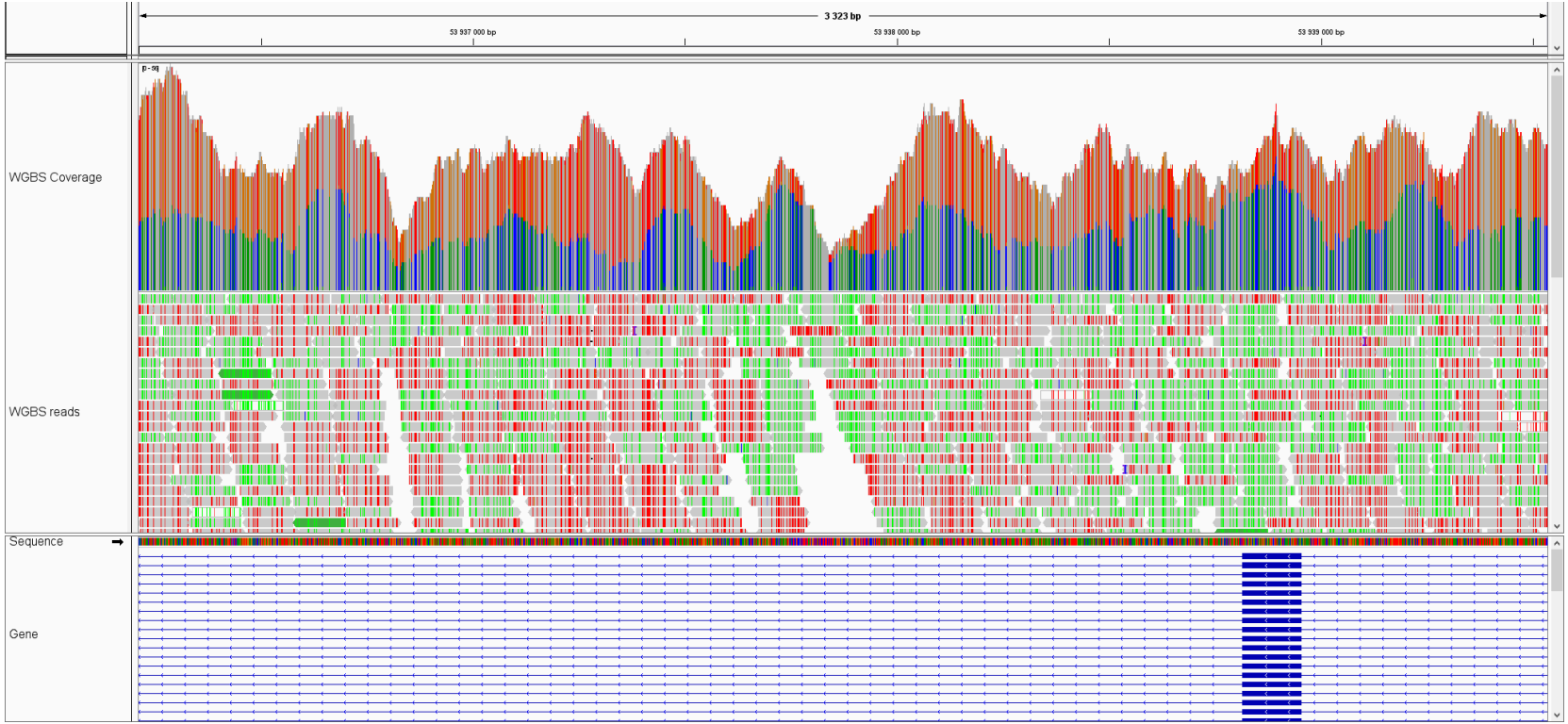
- 5-hmC profiling available
- Infinium Mouse Methylation Array

100-500 bp resolution

Single base resolution

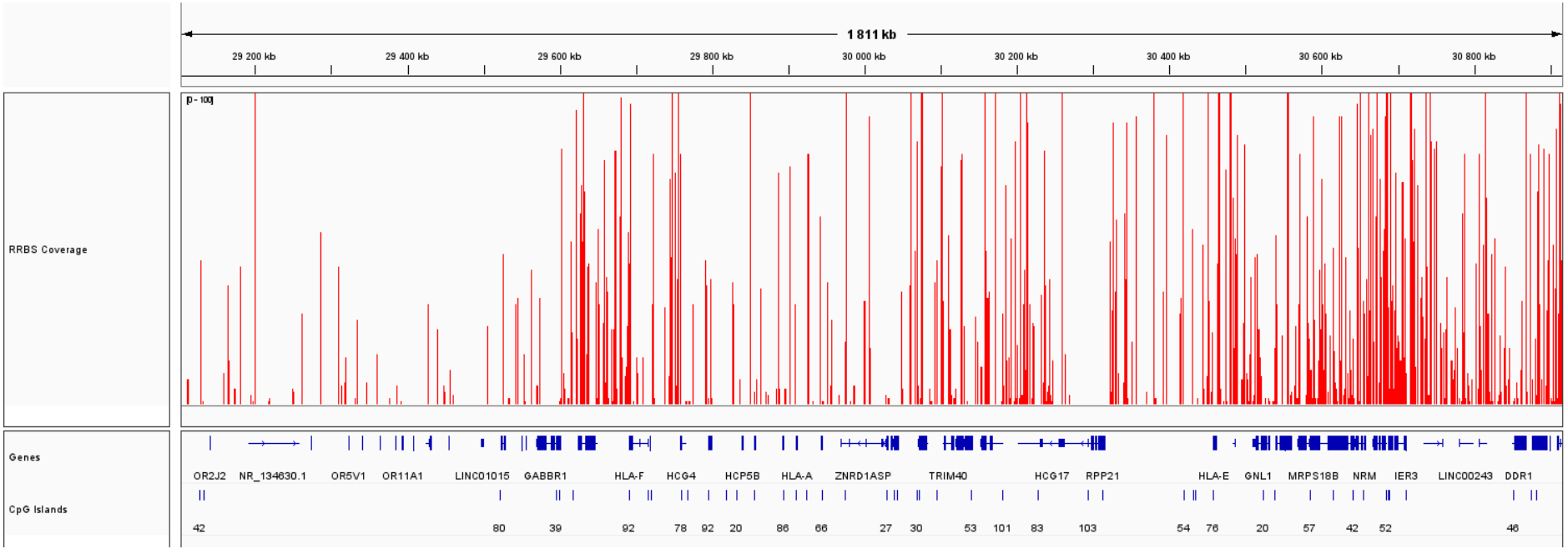
- Service
- Kit

Coverage: WGBS



Visualization of WGBS data on Integrative Genomic Viewer Browser (IGV)

Coverage: RRBS

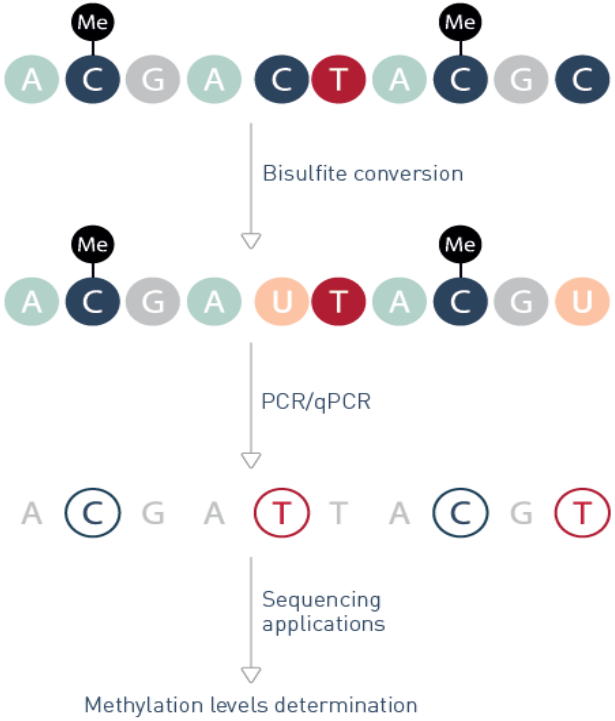


Visualization of RRBS data on Integrative Genomic Viewer Browser (IGV)

Bisulfite conversion

- **BS conversion:**
  - Cytosine (C) is converted to uracil (U) and next to thymine (T) via DNA amplification
  - Methylated C remain unchanged

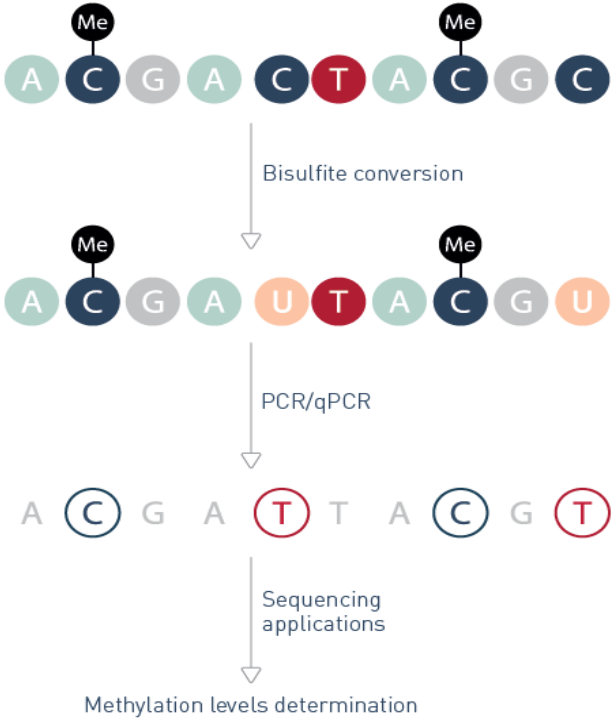
- **NGS analysis:**
  - Single nucleotide resolution
  - Genome-wide and region-specific analyses
  - Identification of differentially methylated CpG-sites/regions



Bisulfite conversion

- **Conversion efficiency is critical**
  - variations in efficiency affect CpG site coverage
  - incomplete conversion -> false positive methylation call

- **QC with spike-in controls**
  - assess the level of under/over-conversion events



What to expect with RRBS?

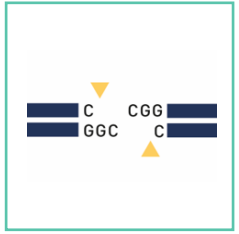
Reduced Representation Bisulfite Sequencing (RRBS) offers a cost-effective solution to perform genome-scale DNA methylation analysis at single nucleotide level

	WGBS	RRBS	Targeted BS-seq
Coverage	Whole genome	CpG-rich regions	Specific regions
Resolution	Single-nucleotide		
# CpG (human)	~25-30 M	~8 M	TBD
# CpG at ≥10X coverage	~25M	~4 M	TBD
Sequencing cost	High	Medium	Low
Computational cost	High	Low	Medium

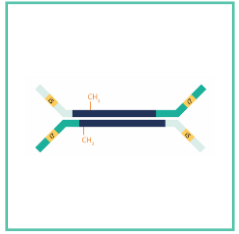
Workflow: Reduced Representation Bisulfite Sequencing

Sample preparation

Enzymatic Digestion (Msp1)

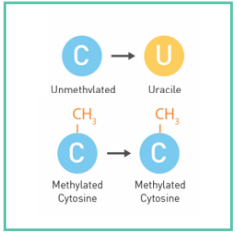


Library preparation

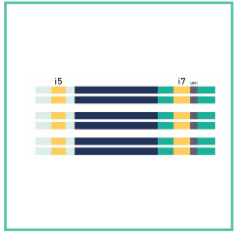


Size selection

Bisulfite conversion



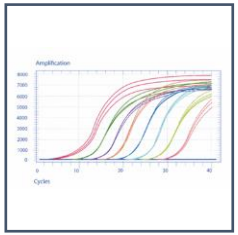
Library amplification



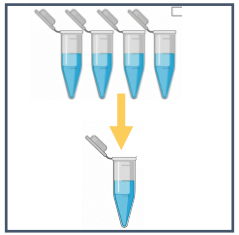
NGS analysis

DIAGENODE

Early UDI indexing & UMI tagging

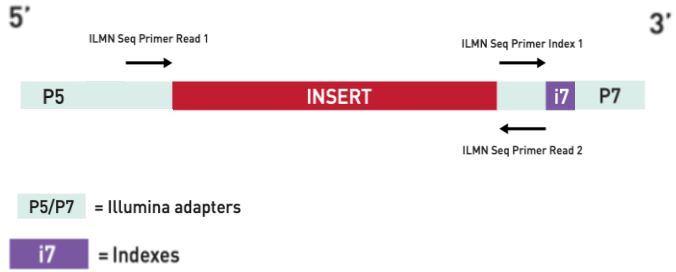
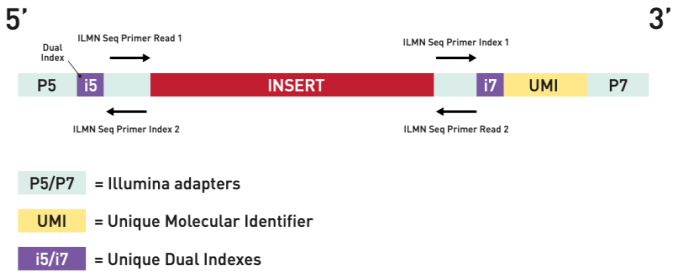


Quantification



Sample pooling

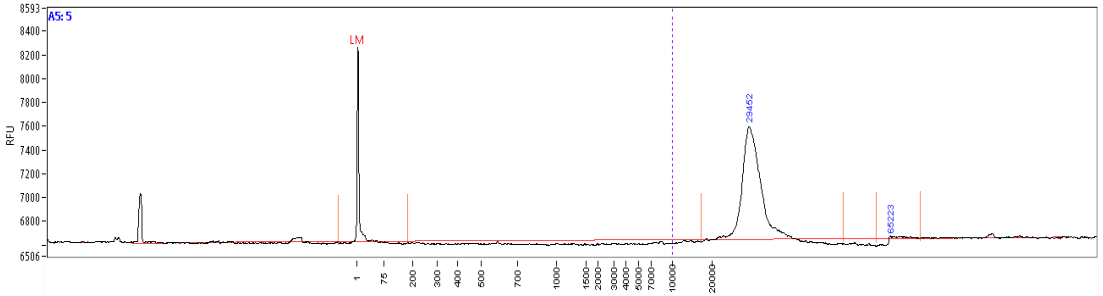
Premium RRBS V2 upgrade

Premium RRBS V1	Premium RRBS V2
Starting amount: 100 ng gDNA	Starting amount: 25 - 100 ng gDNA
	
Single Indexing (24 SI)	Unique Dual Indexing (2 sets 24 UDI)
No duplicates removal from data	UMI-duplicates removal

Cell lysis & gDNA extraction

- **Starting amount for RRBS:** 25 ng – 100 ng of gDNA

- **gDNA quality control**



High integrity gDNA

- **Cultured cells**

  Recommended isolation kit: XL GenDNA Extraction Module (Diagenode)

- **Other templates (e.g. tissue)**

  Flexibility to use any kits yielding high integrity gDNA

  Ensure good gDNA quality and purity

Enzymatic digestion

- MspI enzyme cuts the genome at 5’-CCGG-3’ sites
  - Enrichment in CpG-rich regions

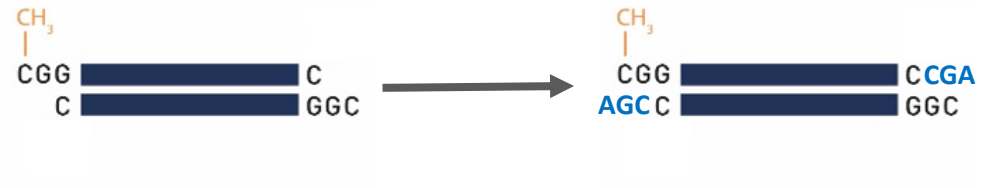
- Highly relevant for vertebrates: methylation mostly in CpG contexts

- Less relevant for plants: methylation in various contexts (45% CpG, 25% CHH, 30% CGH)

CpG-rich regions focus

Early Library preparation

- **End Repair**



- **DNA spike-in controls for BS conversion efficiency QC**
  - Synthetic sequence without homology to any model species
  - No interference with DNA sample of interest

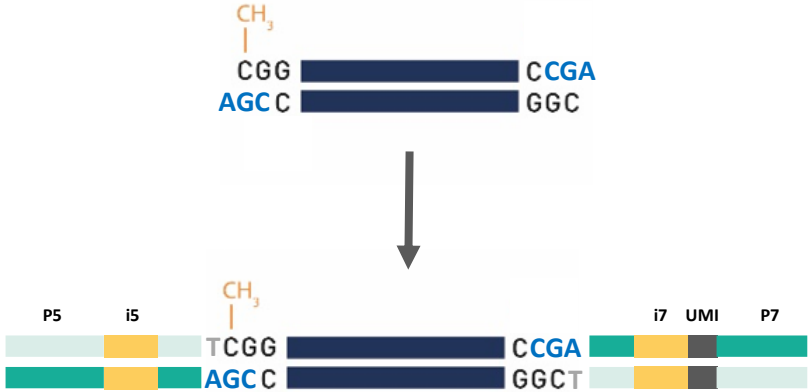
  - **Methylated (positive) spike-in**: detection of over-conversion
  - **Unmethylated (negative) spike-in**: detection of under-conversion

spike-in controls

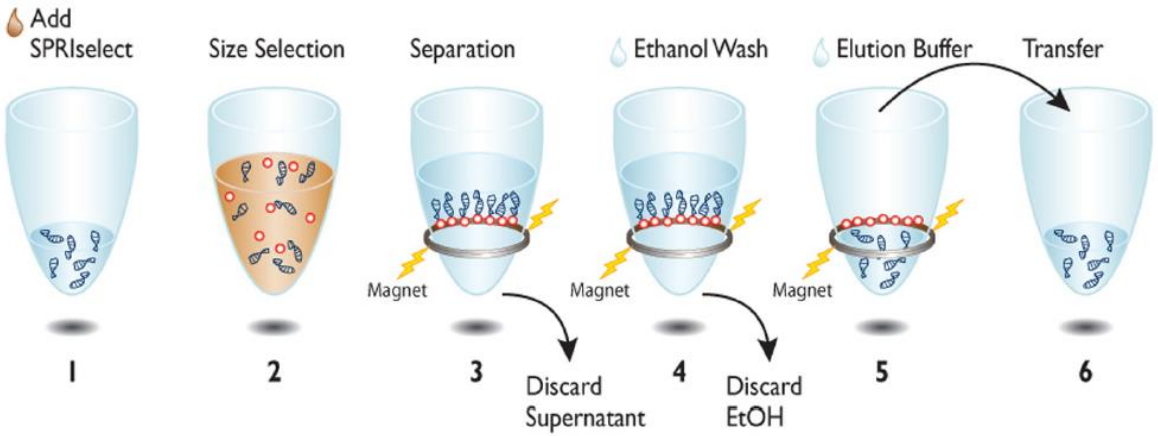
Included in the Diagenode RRBS kit + data analysis manual

Early Library preparation

- Adaptor ligation on dsDNA **before BS conversion**
  - Allows early multiplexing
  - Not possible on ssDNA after BS conversion
- Reduce adapter-dimer formation
  - Adjust adapters concentration to DNA starting amount
- Use of specific **methylated UDI-UMI-adapters**
  - **Unique dual indexes** (UDI) identifies read misassignment
  - **Unique molecular identifier** (UMI) identifies PCR-duplicates

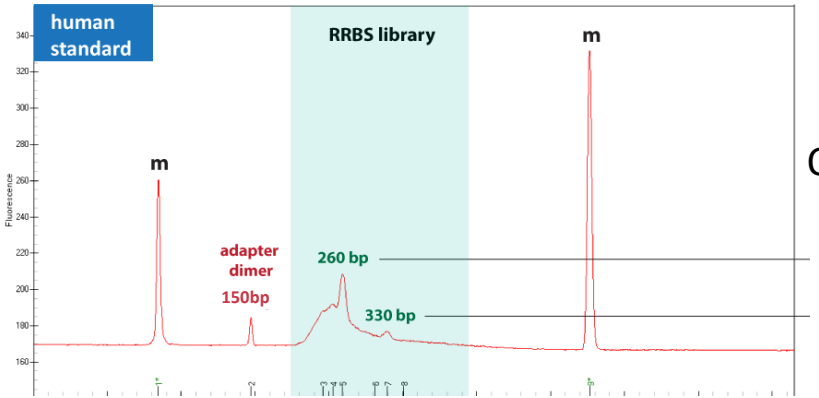


Bead-based size selection

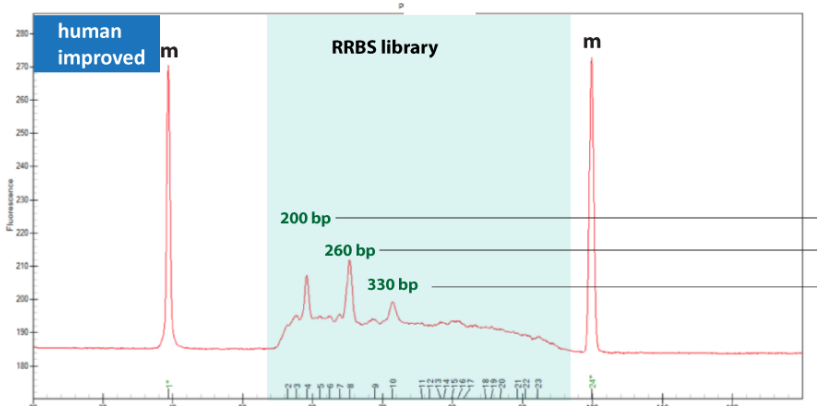


Adapter-excess/dimer removal

Bead-based size selection



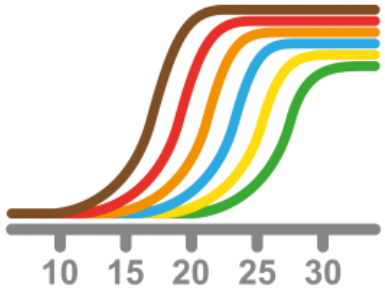
Optimization

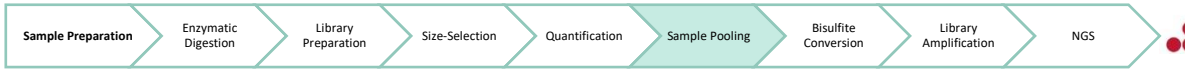


Careful optimized bead-based size selection provides the best sequencing results and high genomic coverage.

qPCR Quantification

- check correct ligation
- Find optimal pooling conditions



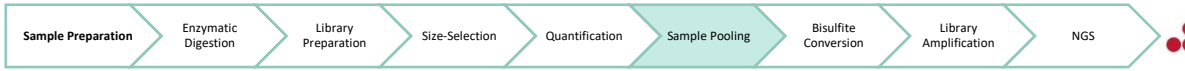


Early Sample Pooling

How many samples per pool?

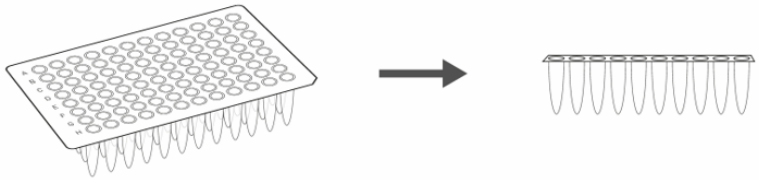
- Consider **# reads per lane** depending on the flow cell
- **# reads needed for each sample** to properly cover CpG islands
  - Human/Mouse: ~30-40 million raw reads per sample
- **mapping efficiency for the genome of interest**
- **# of barcodes available** (48 UDI available)

Suitable for large sample numbers



Early Sample Pooling

- RRBS libraries with compatible barcodes are pooled based on **qPCR results**
  - Reduces costs and handling time
  - Reduces PCR amplification
  - Software for Intelligent Pooling available soon

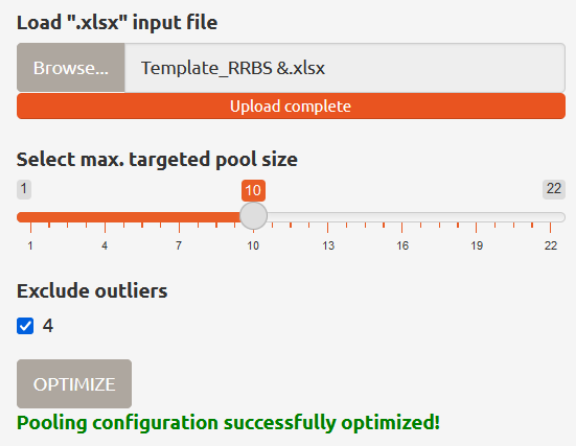


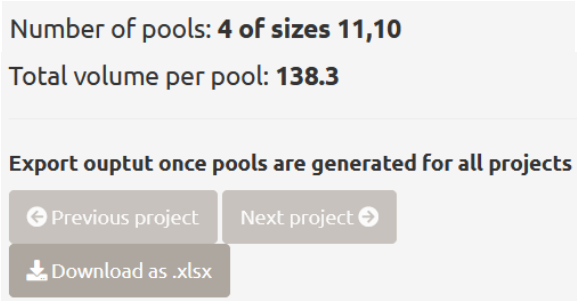
Early Sample Pooling

	A	B	C	D	E	F	G	H	I	J	K
1	Well	Sample name	**Adaptor ID**	Ct1	Ct2	**Average Ct**	delta ct	**pooling vol**	**water vol**	**pool name**	**total vol**
2						**"=AVERAGE(D3:E3)"**	"=F$12-F3"	**"=17*2^(-G3)"**	**"=(150-SUM(H3:H12))"**		
3	D08	Sample 44	**20**	5.4	5.19	**5.295**	0.76	10.0	**10.3**	**pool 1**	150
4	D05	Sample 41	**17**	5.68	5.23	**5.455**	0.6	11.2			
5	D06	Sample 42	**18**	5.69	5.38	**5.535**	0.52	11.9			
6	A09	Sample 9	**9**	5.56	5.92	**5.74**	0.315	13.7			
7	C10	Sample 34	**10**	5.67	5.85	**5.76**	0.295	13.9			
8	A11	Sample 11	**11**	5.83	5.76	**5.795**	0.26	14.2			
9	C12	Sample 36	**12**	6.09	5.66	**5.875**	0.18	15.0			
10	A06	Sample 6	**6**	5.82	6.13	**5.975**	0.08	16.1			
11	D07	Sample 43	**19**	5.82	6.25	**6.035**	0.02	16.8			
12	D04	Sample 40	**16**	5.91	6.2	**6.055**	0	17.0			
13											
14	D02	Sample 38	**14**	6.26	6.22	**6.24**	0.76	**10.0**	**16.5**	**pool 2**	**150**
15	A07	Sample 7	**7**	6.21	6.33	**6.27**	0.545	**11.7**			
16	D03	Sample 39	**15**	6.58	5.99	**6.285**	0.53	**11.8**			
17	A02	Sample 2	**2**	6.33	6.37	**6.35**	0.465	**12.3**			

Balanced library pooling

Software for Intelligent Pooling





Balanced library pooling

Features & Benefits

- **Time-saving** – Avoid complex calculations
- **Highest pooling efficiency** - based on qPCR quantification
- **Powerful** – Incorporates advanced aspects such as number of samples per pool required, the separation between projects
- **Accurate** – Identify outliers

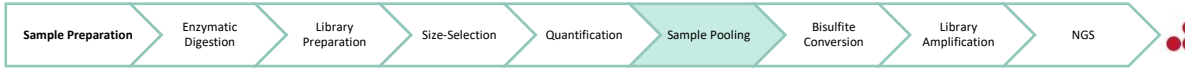
Software for Intelligent Pooling

A	B	C	D	E	F	G
source_well	**sample_name**	**adaptor_ID**	**notes**	**pool**	**vol_sample_ul**	**vol_water_ul**
Project: 10						
B01	Sample 13	13		1	17	0
B02	Sample 14	14		1	16,3	
A01	Sample 1	1		1	12,9	
A05	Sample 5	5		1	12,2	
A07	Sample 7	7		1	12,1	
B05	Sample 17	17		1	11,6	
B07	Sample 19	19		1	11,5	
B08	Sample 20	20		1	10,9	
A11	Sample 11	11		1	8,7	
A09	Sample 9	9		1	8,4	
B03	Sample 15	15		2	17	13
B04	Sample 16	16		2	13,2	
A03	Sample 3	3		2	12,8	
A02	Sample 2	2		2	12,6	
A08	Sample 8	8		2	12,4	
A12	Sample 12	12		2	11,3	
B06	Sample 18	18		2	10,5	
A06	Sample 6	6		2	9,7	
A10	Sample 10	10		2	9	

Balanced library pooling

Features & Benefits

- **Time-saving** – Avoid complex calculations
- **Highest pooling efficiency** - based on qPCR quantification
- **Powerful** – Incorporates advanced aspects such as number of samples per pool required, the separation between projects
- **Accurate** – Identify outliers



Early Sample Pooling

After pooling, column-based concentration to reduce pool volume.

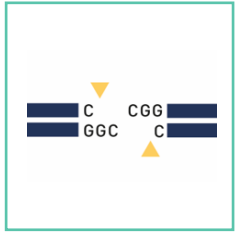
Optimized to reduce sample loss:

- Reduced library amplification cycles
- Reduced % duplicates in data
- Allows for lower DNA starting amount handling

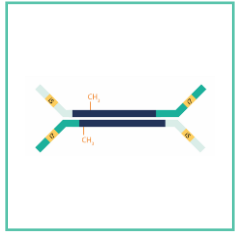
Workflow: Reduced Representation Bisulfite Sequencing

Sample preparation

Enzymatic digestion

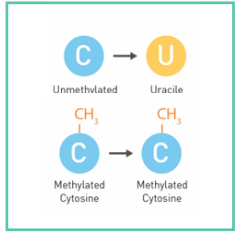


Library preparation

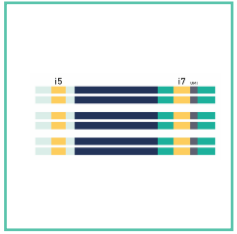


Size selection

Bisulfite conversion



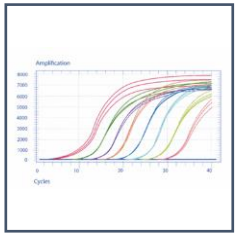
Library amplification



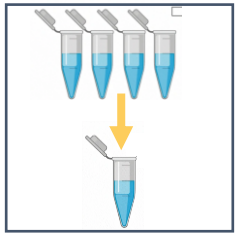
NGS analysis

DIAGENODE

Early UDI indexing & UMI tagging



Quantification

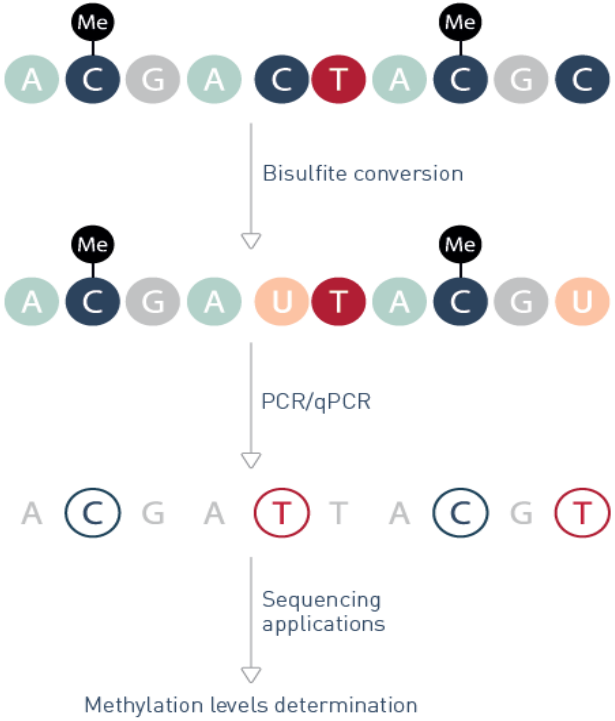


Sample pooling

Bisulfite conversion

- Optimized protocol to:
- Ensure efficient conversion
- Reduce BS-induced fragmentation

Optimized
BS Conversion



Library amplification: Optimal PCR cycles

Minimize # PCR cycles

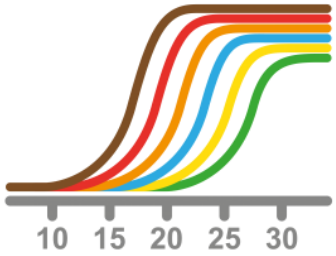
Minimization of PCR cycles

- Limits the risk of introducing bias during PCR
- Minimizes # of PCR duplicates

Determine optimal # of PCR cycles for each sample on a small aliquot

- **Optimal cycle number** $= Ct \quad (rounded\ to\ the\ nearest\ whole\ number)$

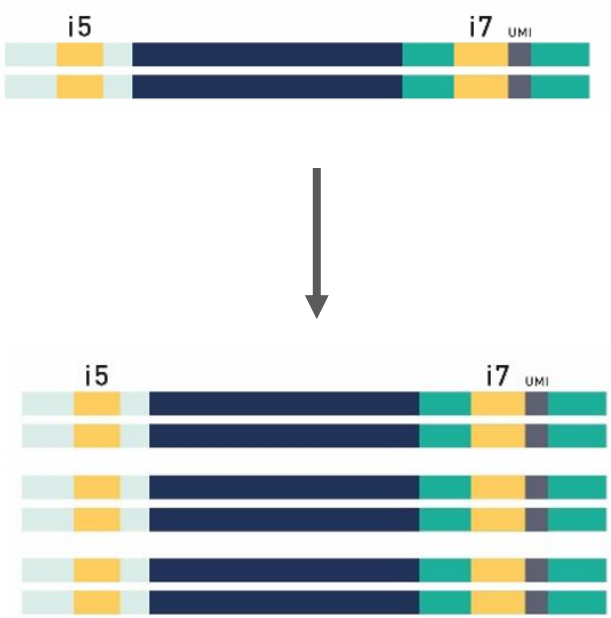
e.g. $Ct = 6.82$ → 7 amplification cycles
$Ct = 9.25$ → 9 amplification cycles



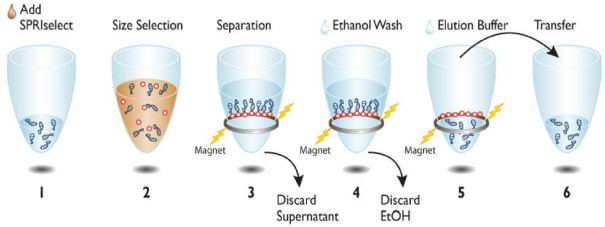
Library amplification

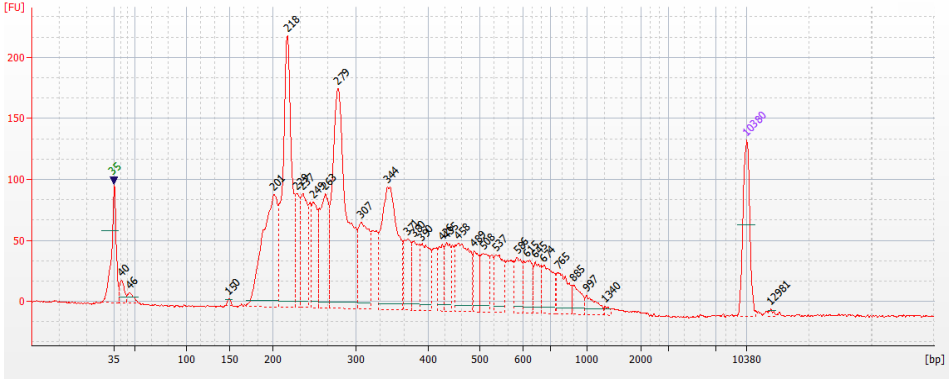
- RRBS amplification Master Mix with Taq polymerase dedicated to bisulfite converted DNA

Efficient amplification

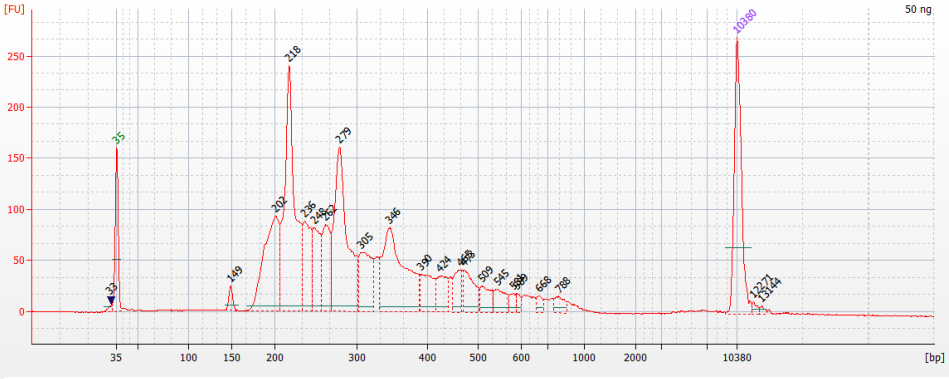


Library amplification: Clean-up



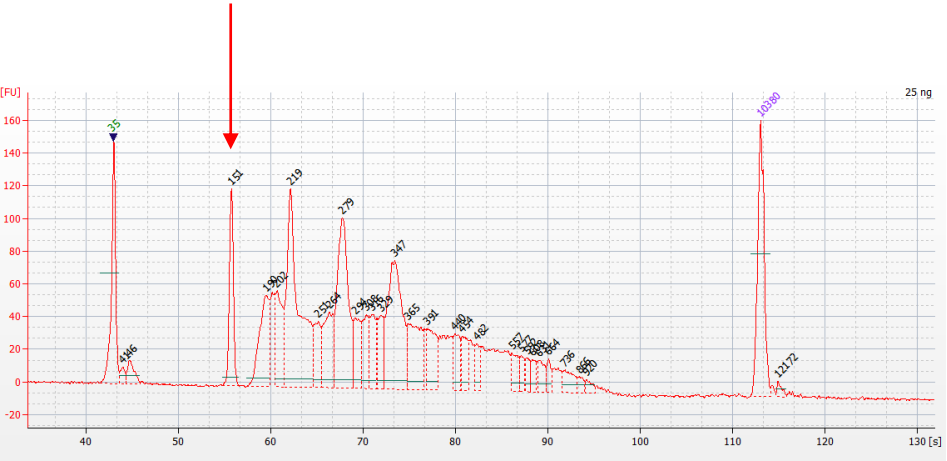


*Final Premium RRBS V2 libraries prepared from **100 ng** human gDNA (pooled by 5 samples)*

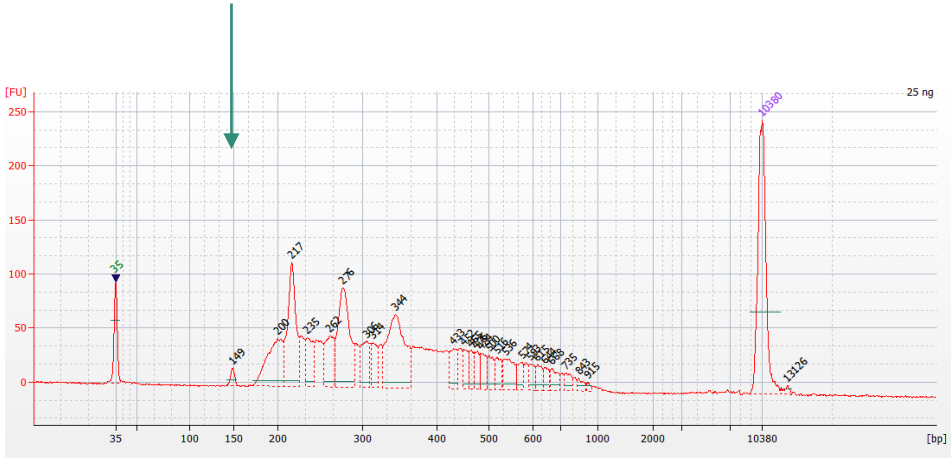


*Final Premium RRBS V2 libraries prepared from **50 ng** human gDNA (pooled by 5 samples)*

Library amplification: Final clean-up



*Final Premium RRBS V2 libraries prepared from **25 ng** human gDNA (pooled by 5 samples) <u>before</u> additional clean-up*



*Final Premium RRBS V2 libraries prepared From **25 ng** human gDNA (pooled by 5 samples) <u>After</u> 2nd clean-up*

Sequencing recommendations

- Sequencing 2x 50bp
- Specific run mode to read UMI: cycle 50-8-(i5) and 17-(i7)-50
- RRBS are low diversity libraries (MspI restriction motif + BS conversion)
  - Reduce library concentration for better clustering
  - Include other libraries to increase diversity (PhiX spike-in)
  - Update the version of the Control Software from Illumina

Step	MiSeq	HiSeq2000/2500	HiSeq3000/4000	NextSeq500/550	NovaSeq6000
Percentage of Illumina PhiX control	5%	5%	5%	20%	15%
Cluster density	Aim at a 30% beneath the optimal range for the chemistry version and platform used				
Softwatre version	RTA 1.17.28 or newer	HCS 2.2.38 or newer	HCS 3.4.0.38 or newer	NCS 1.3 or newer	NSCS 1.6 or newer

- Read depth depending on organism (Human/mouse 30-40M reads/sample – compare WGBS: 500M)

RRBS V2: Data Analysis Pipeline

Reads Processing

- Demultiplexing
- Trimming
- Alignment
- PCR duplicates removal with UMI

Spike-in Processing

- Calculate BS conversion rates
- Spike-in FASTA & user manual

Methylation Analysis

- Methylation extraction
- Identify DMRs
- DMRs annotation & visualization

RRBS: Standard Analysis

	Before UMI processing and PCR duplicates removal			After UMI processing and PCR duplicates removal		
	No. of CpGs detected with coverage ≥1x	No. of CpGs detected with coverage >10x	Average CpG coverage	No. of CpGs detected with coverage ≥1x	No. of CpGs detected with coverage >10x	Average coverage
100 ng	7.6 M	4.4 M	14	7.6 M	3.7 M	11
50 ng	7.4 M	4.3 M	13	7.4 M	2.7 M	9
25 ng	7.4 M	3.8 M	13	7.4 M	1.9 M	7

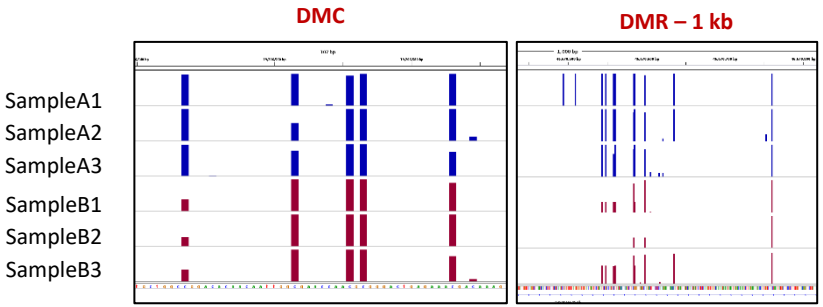
- Detect more CpGs
- Get Better CpG coverage
- Remove any artefacts due to PCR for an accurate coverage of your CpGs

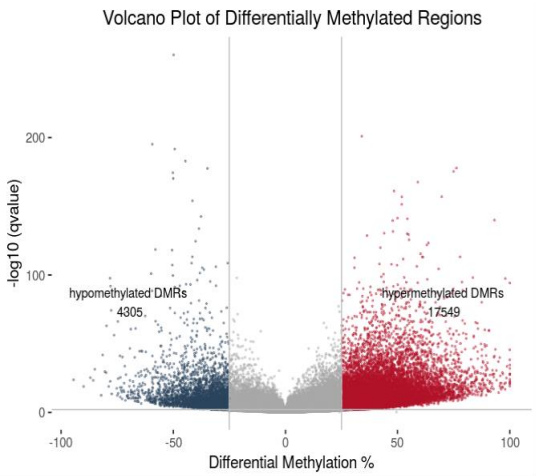
RRBS: Advanced Analysis

1. Differential methylation analysis
   - Methylation level analysis
   - Differentially Methylated CpGs (DMCs) analysis
   - Differentially Methylated Regions (DMRs) analysis
   - Annotation of DMCs and DMRs for genomic regions
   - Clustering analysis

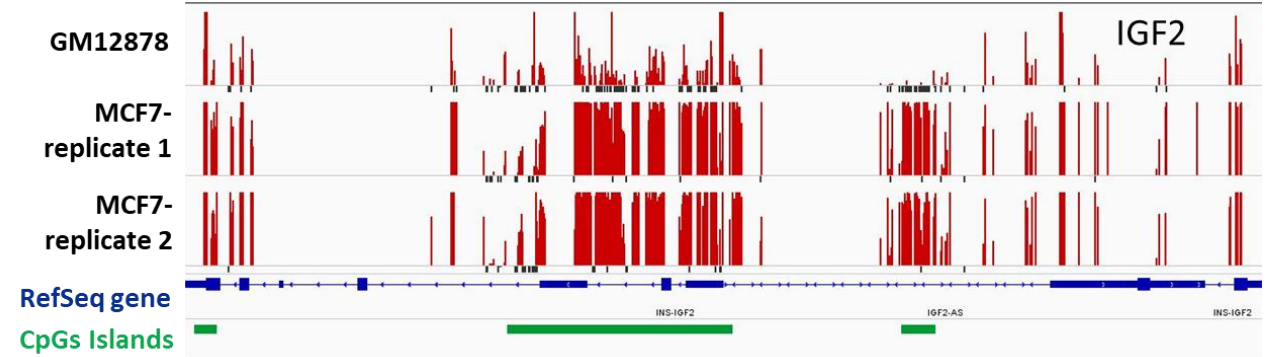
2. Gene ontology terms analysis
   - Enrichment analysis on gene associated with DMCs and DMRs
   - Get functional insights

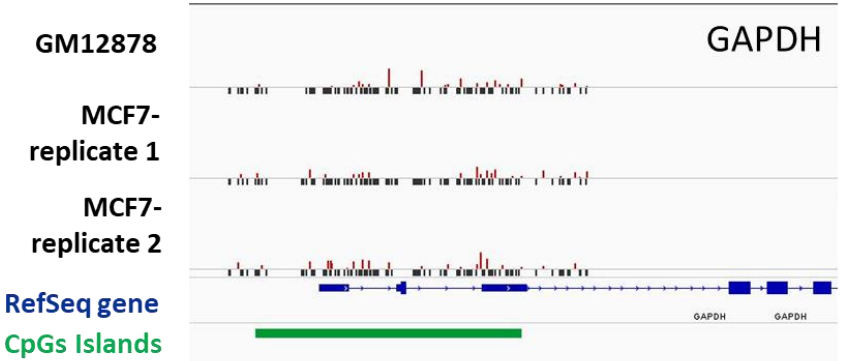
3. Pathway analysis
   - Identification of biological pathways in which genes associated with DMCs and DMRs may be over-represented (or under-represented)
   - Get mechanistic insights





RRBS: Advanced Analysis

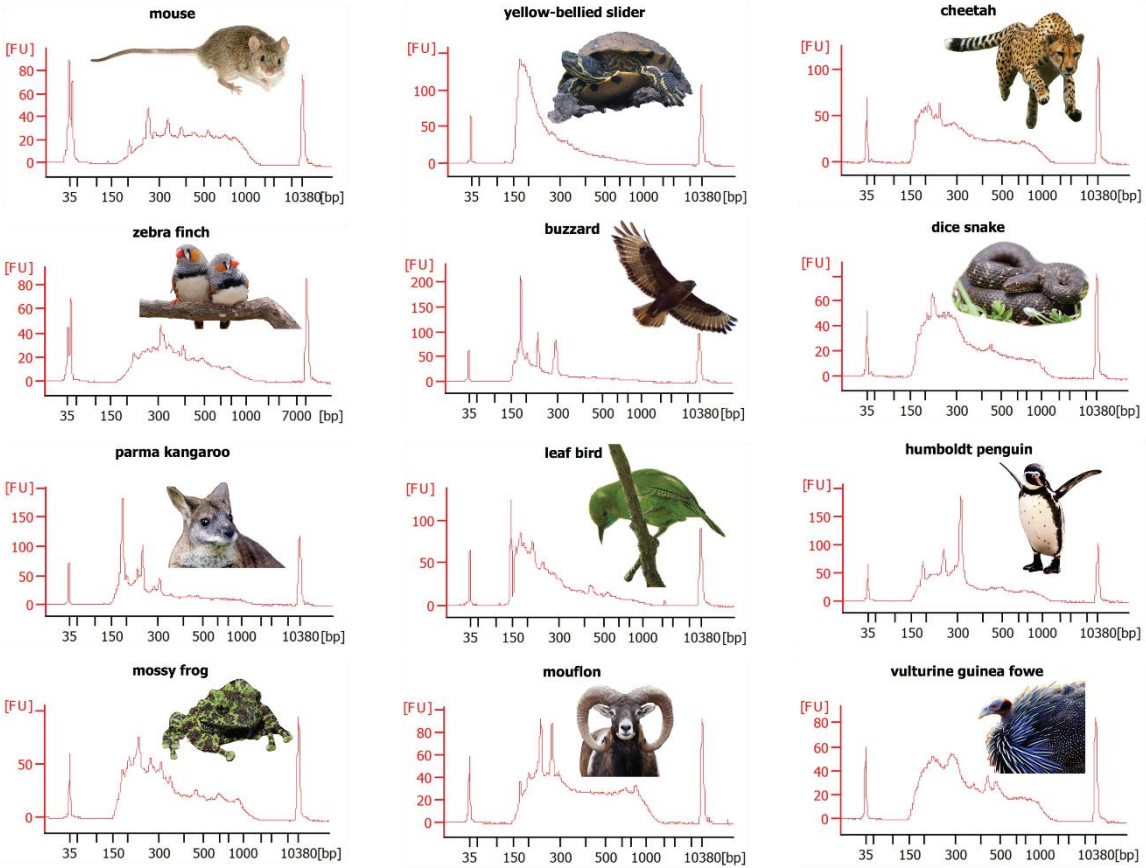




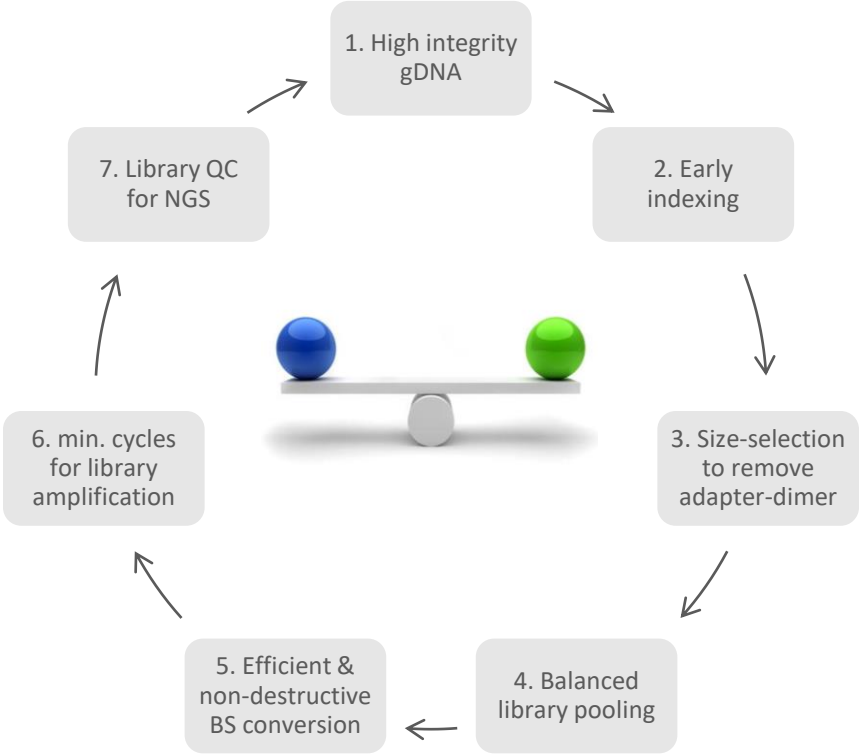
- In red: DNA methylation >1%
- In grey : DNA methylation <1%

RRBS versatility

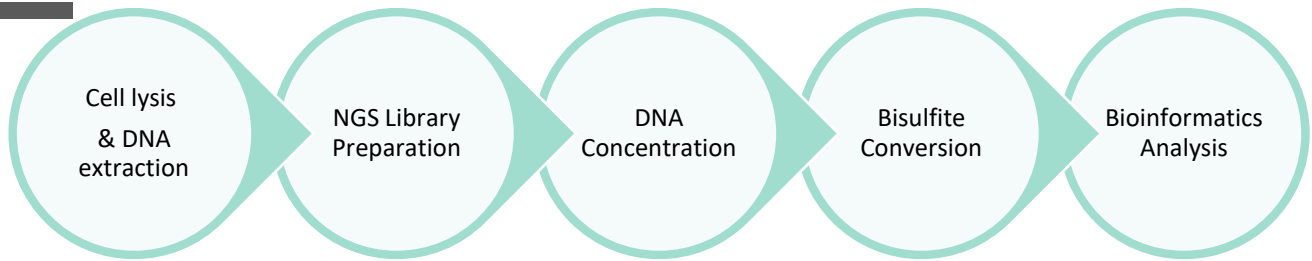
All vertebrate species



Summary – Tips for Good RRBS Assay



Summary – Diagenode Support



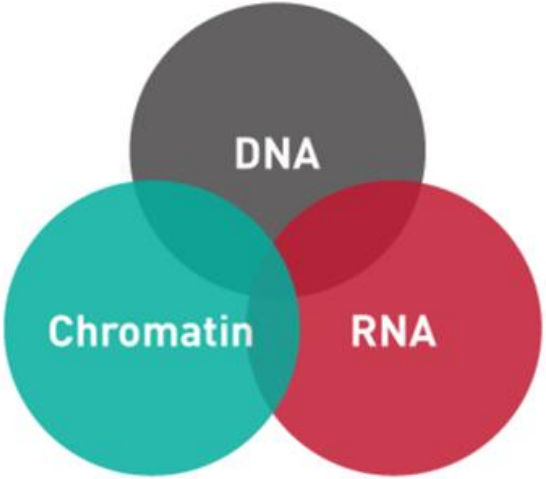
Reagents:	XL GenDNA Extraction Module	Spike-in controls	DiaPure Columns	Bisulfite conversion reagent for RRBS	
Kits:		Premium RRBS Kit V2			
Services:	Epigenomic Profiling Services				
					Data Analysis Service

www.diagenode.com

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



THANK YOU!

Thank you for taking part in our DNA methylation workshop!

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 :-))

www.diagenode.com