CEITEC

Central European Institute of Technology

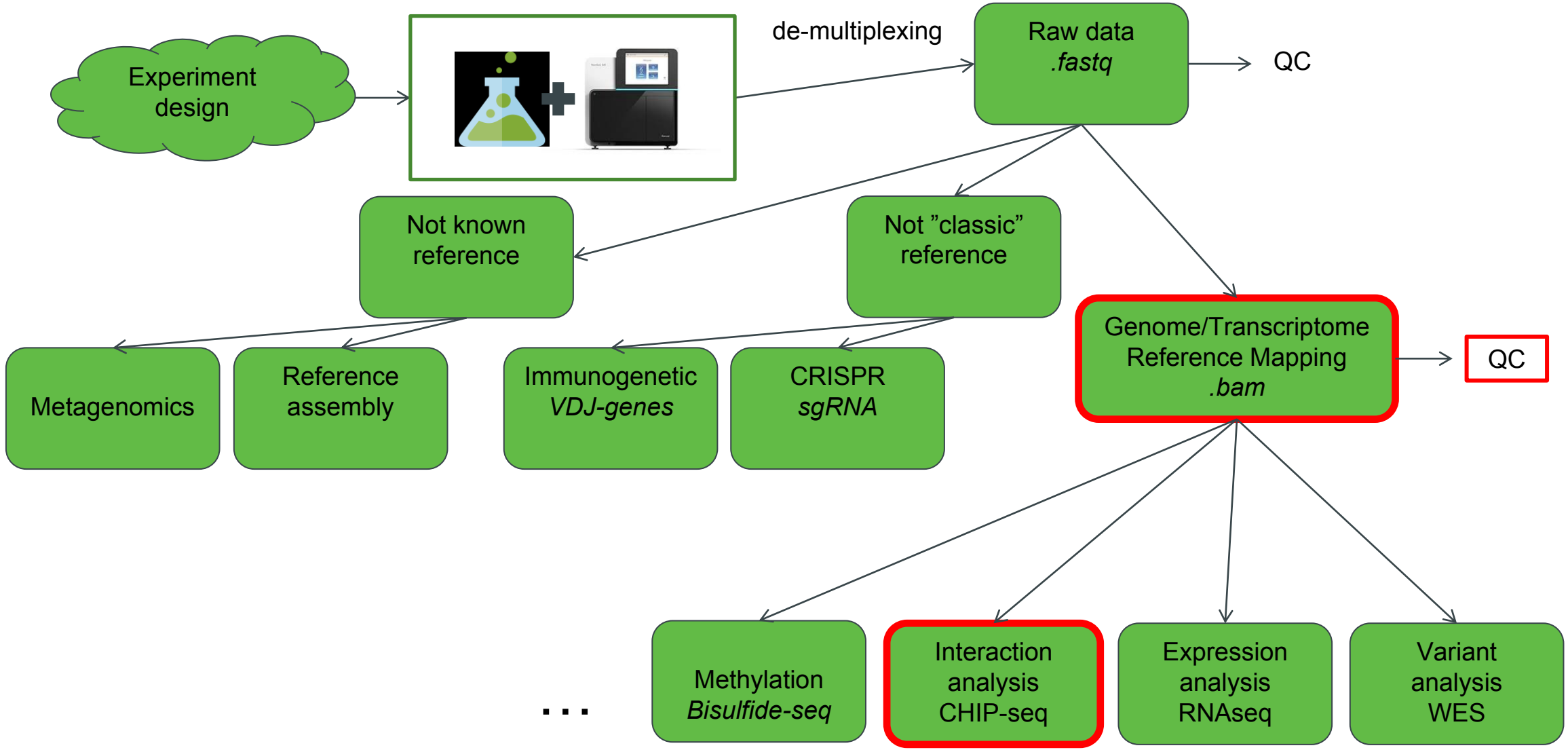
BRNO | CZECH REPUBLIC

Modern Genomic Technologies (LF:DSMGT01)

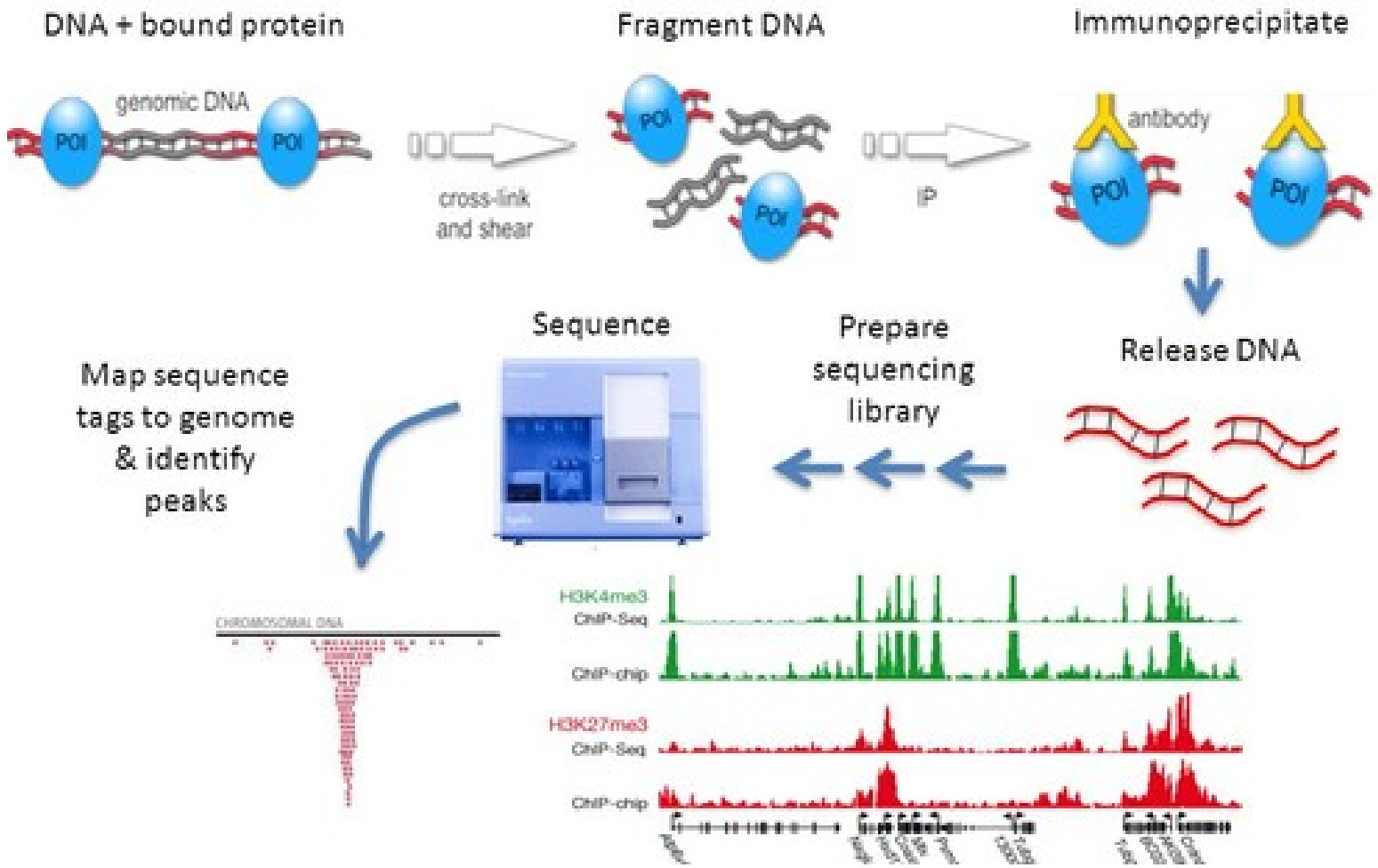
Lecture 5 : ChIP-seq

Vojta Bystry
vojtech.bystry@ceitec.muni.cz

NGS data analysis



ChIP-seq

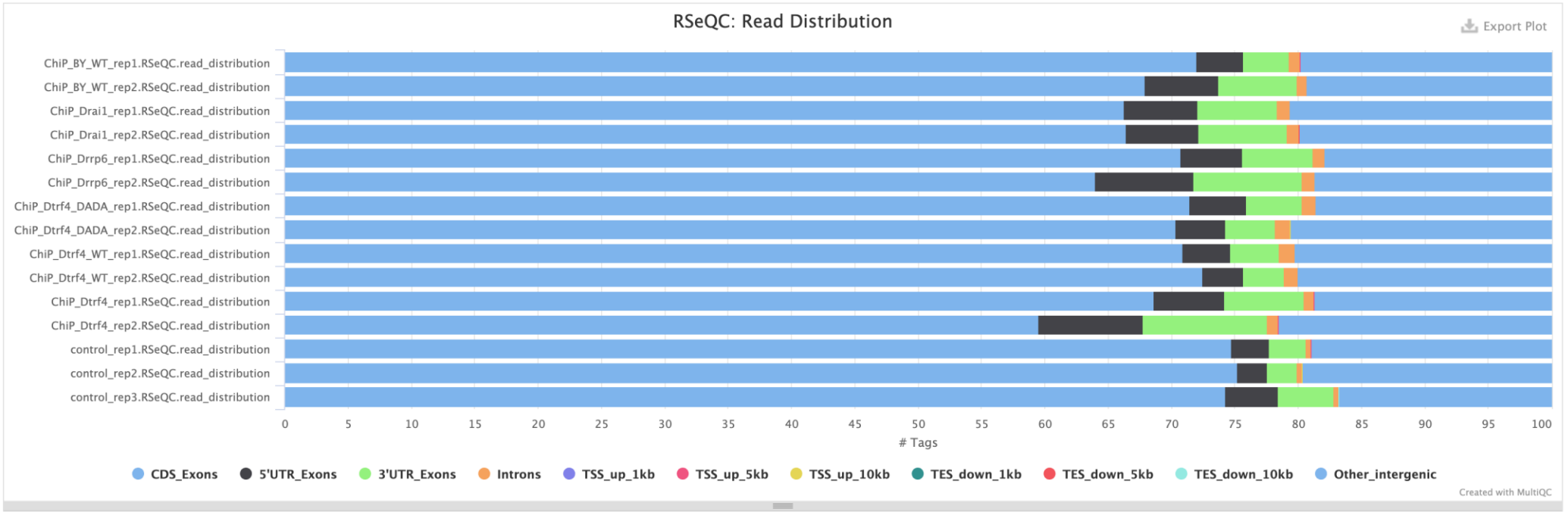


- Very laborious and hard laboratory preparation
- Not a good success rate

CEITEC

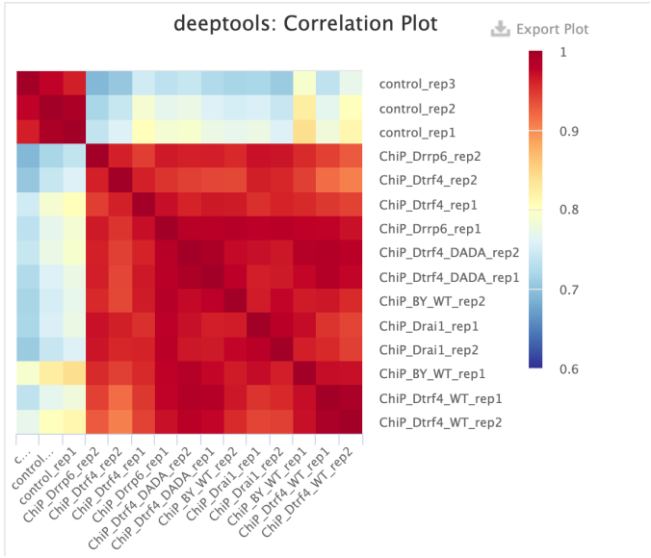
Primary analysis + QC

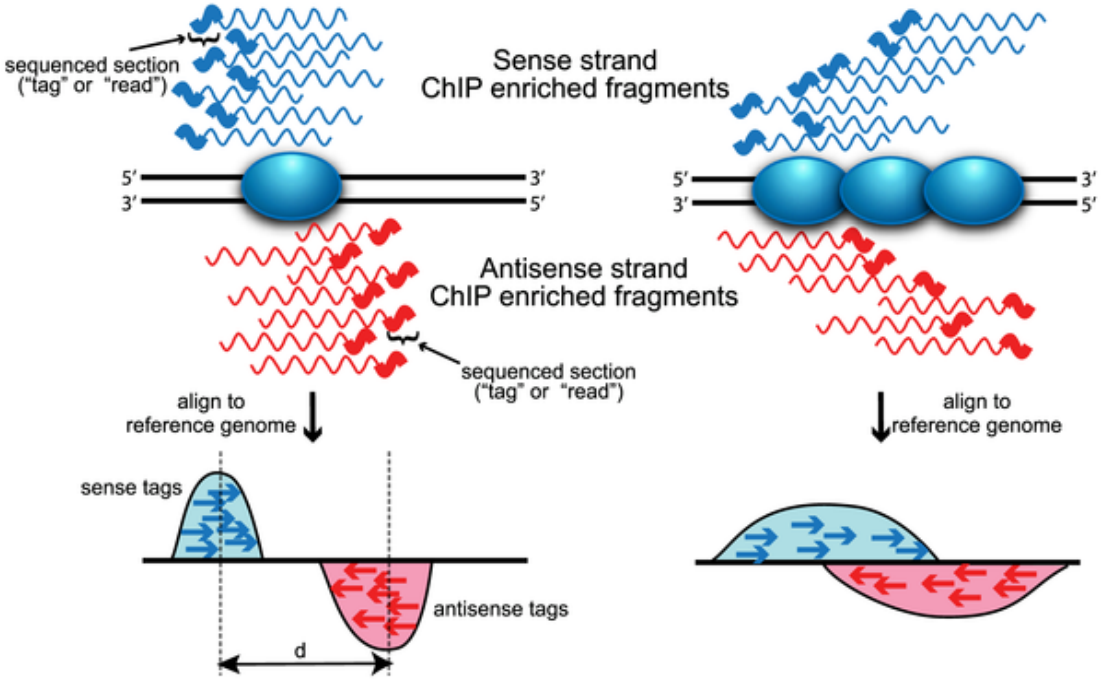
- Alignment – standard DNA (RNA for CLIP)
- QC standard
  - Check sequencing quality
  - RSeQC – Read Dstribution



Primary analysis + QC

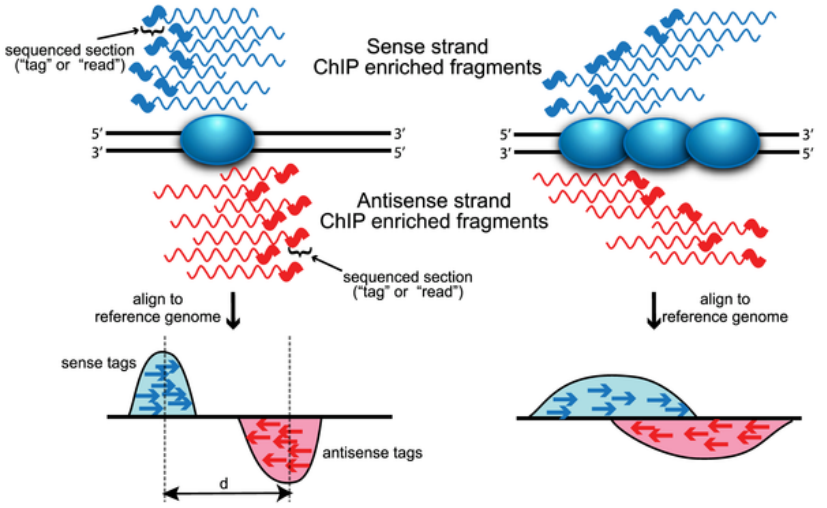
- IP experiment quality control
  - Sample correlation
    - Replicates control treatment
  - Strand cross-correlation
    - Shift of strand mapping
    - Shift should correlate with expected fragment size

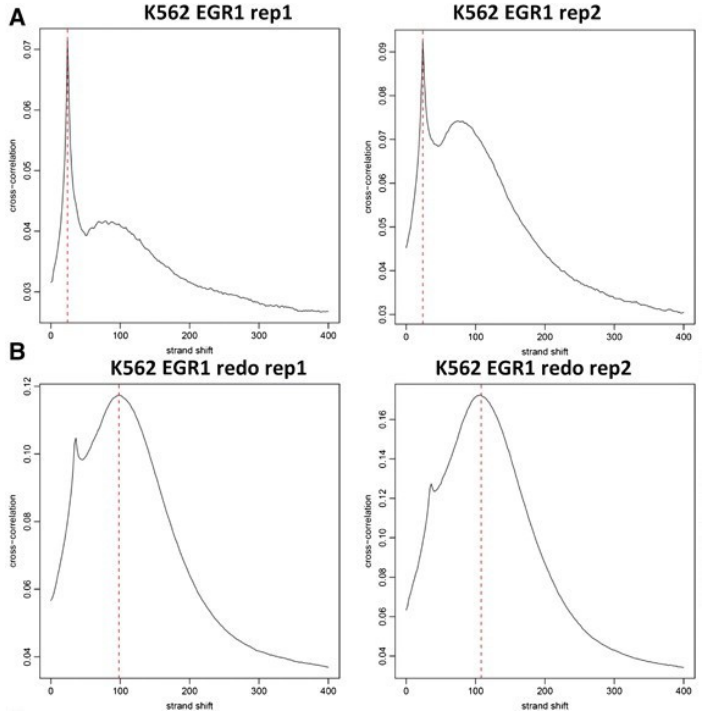




Primary analysis + QC

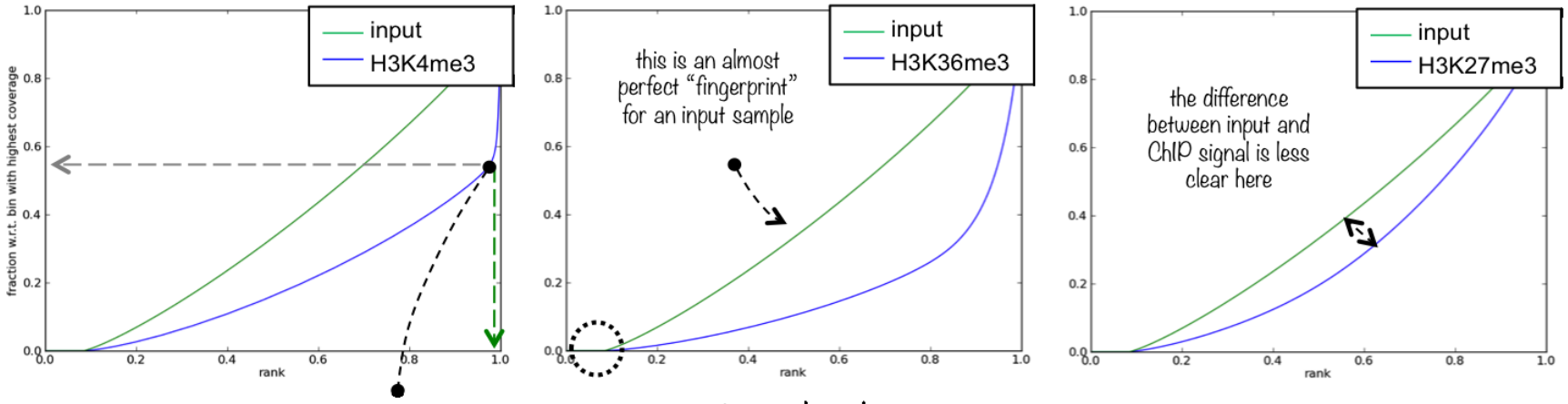
- IP experiment quality control
  - Sample correlation
    - Replicates control treatment
  - Strand cross-correlation
    - Shift of strand mapping
    - Shift should correlate with expected fragment size





Primary analysis + QC

- **Fingerprint profile**
  - **profile of cumulative read coverages**
  - how evenly are the reads distributed over the genome

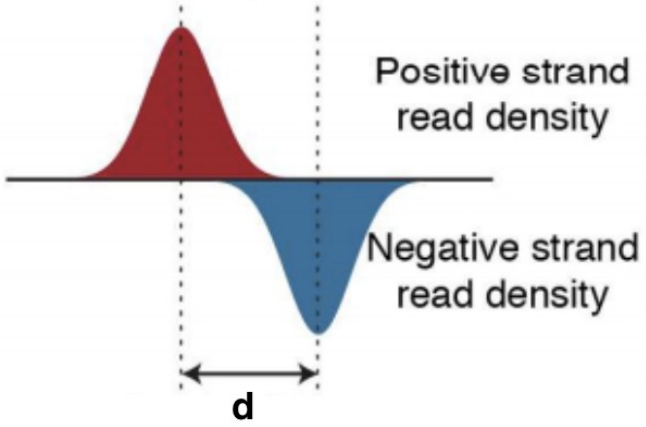


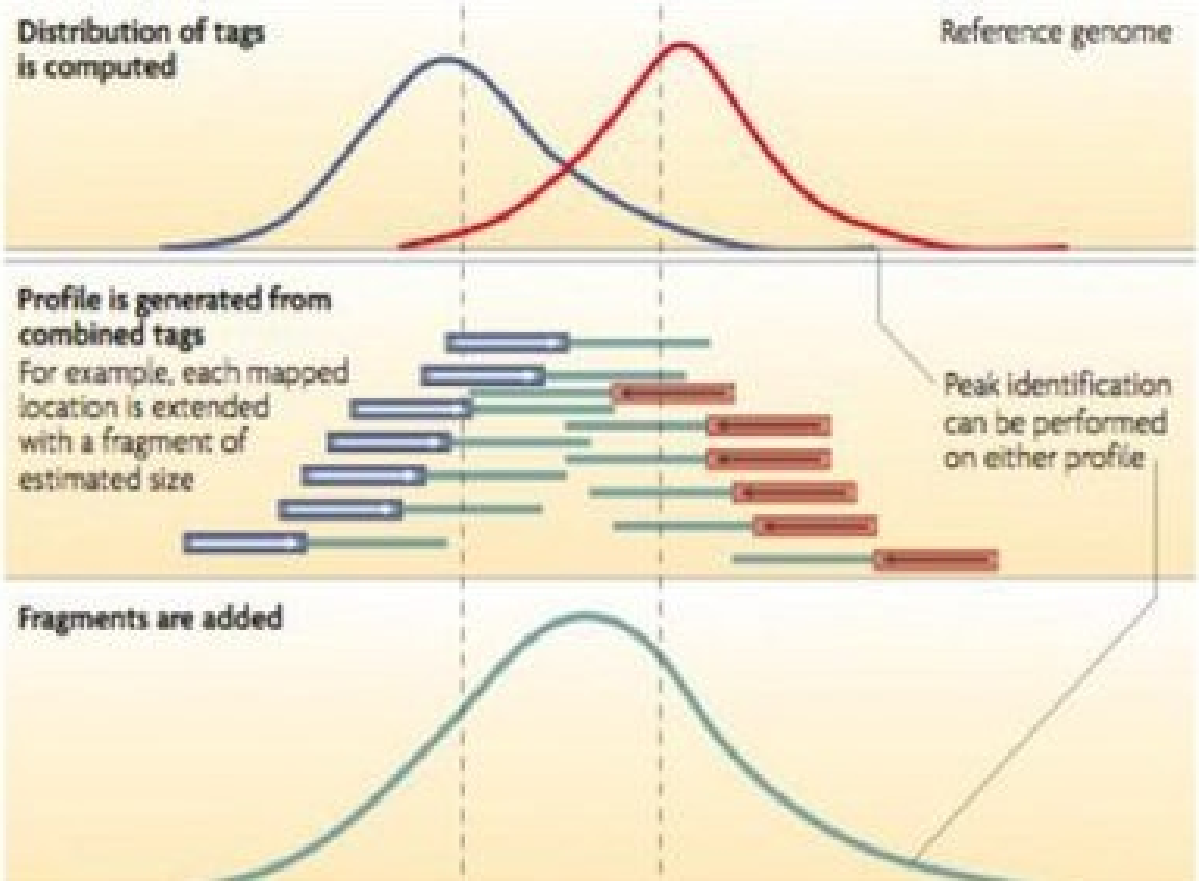
Deduplication

- Remove all duplicates!
  - It is very low probability that the fragment would be cut at exactly same position
  - Usually experiments without UMIs

Peak calling

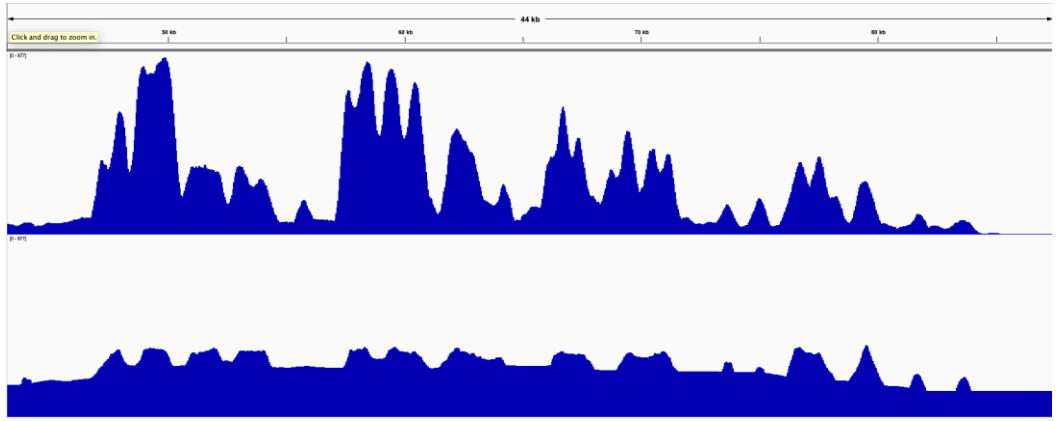
- Read extension





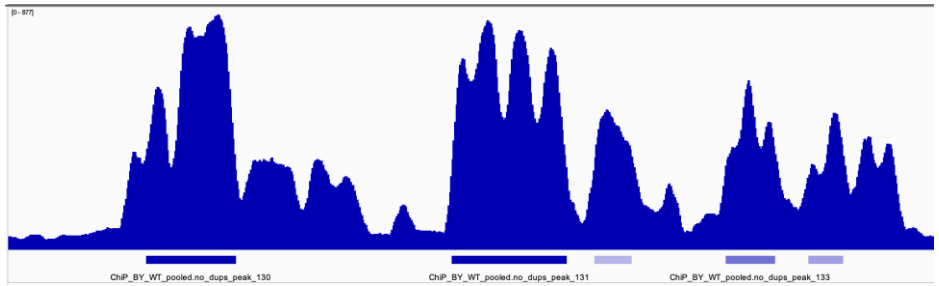
Peak calling

- Statistical assessment of peaks against background
- Background
  - Control sample – recommended
  - Model background from overall coverage of the sample

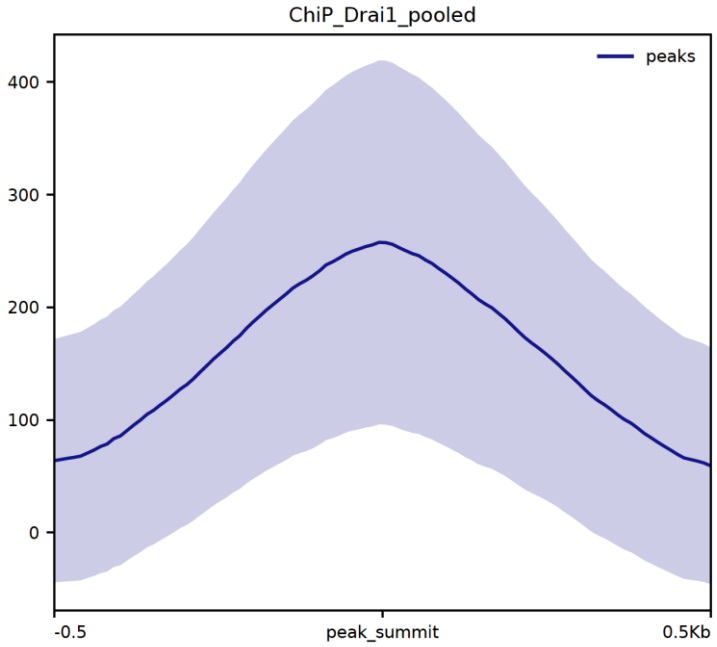


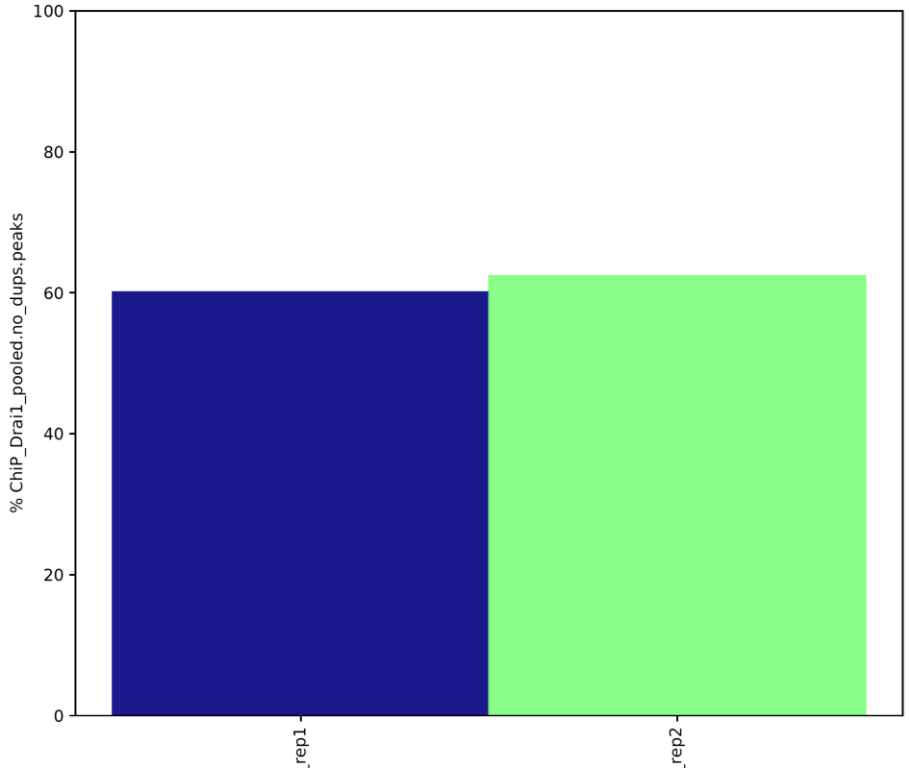
- Peak calling annotation
- Differential peak calling

Post peak calling QC



- FRIP score = fraction of reads in peaks
  - High number is good
  - However can be low in specific experiments and still the experiment be OK
- Average peak profile





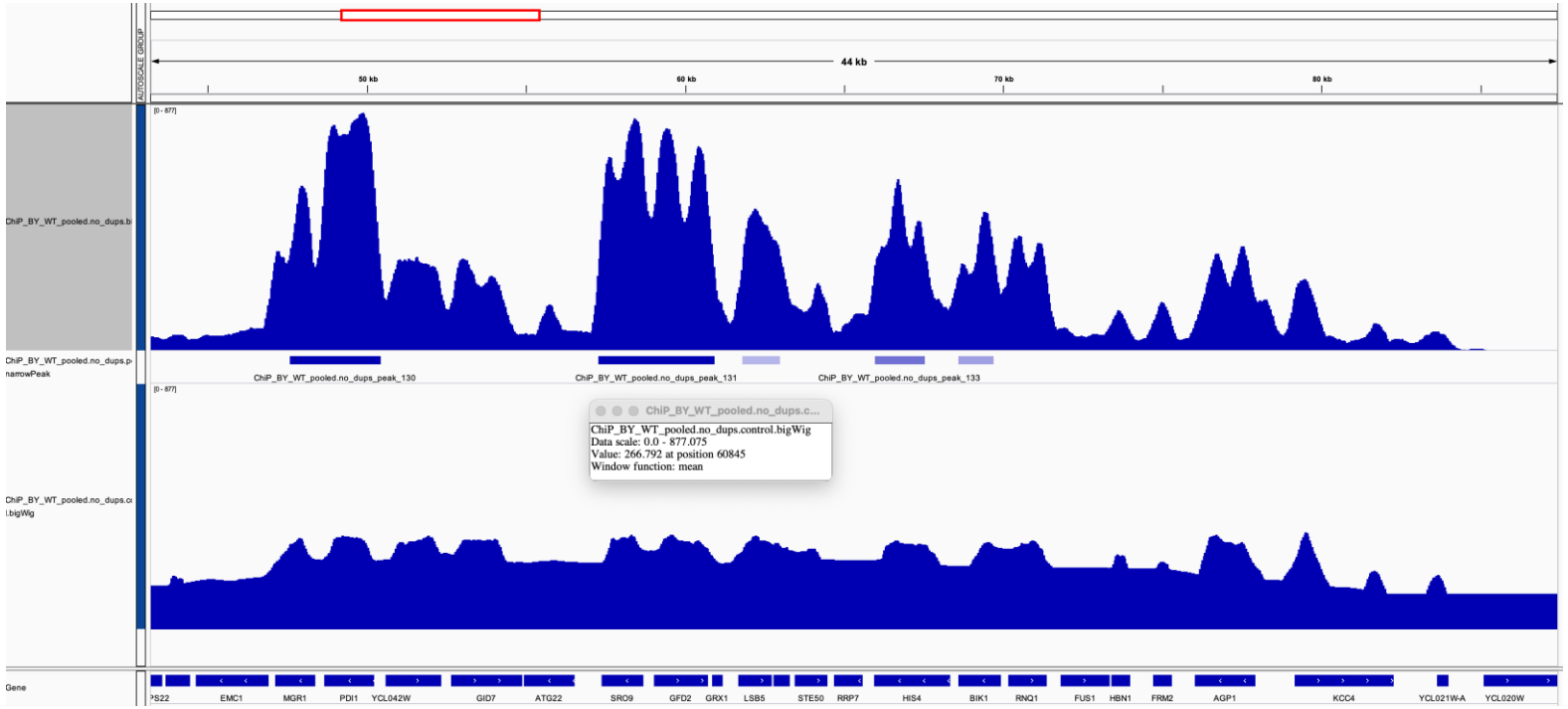
Peak calling results

chr	start	end	peak_ID	overall_score	strand	fold_change	-log(pavalue)	-log(qvalue)	relative_peak_summit	gene_name	gene_id
I	31479	33115	ChiP_BY_WT_pooled.no_dups_peak_1	429	.	1.88871	44.3616	42.9998	1232	GDH3	YAL062W
I	33537	34528	ChiP_BY_WT_pooled.no_dups_peak_2	610	.	2.08354	62.5989	61.0757	507	BDH2	YAL061W
I	35137	36342	ChiP_BY_WT_pooled.no_dups_peak_3	556	.	2.04075	57.1526	55.6747	425	BDH1	YAL060W
I	45839	46698	ChiP_BY_WT_pooled.no_dups_peak_4	126	.	1.43949	13.7207	Dec.75	433	FLC2	YAL053W
I	57192	60004	ChiP_BY_WT_pooled.no_dups_peak_5	854	.	2.40168	87.1869	85.4642	1022	BOL3,NA,BOL1,GCV3,PTA1	YAL046C,YAL045C,YAL044W-A,YAL044C,YAL043C
I	60315	63277	ChiP_BY_WT_pooled.no_dups_peak_6	704	.	2.15353	72.018	70.4181	1323	PTA1,YAL042C-A,ERV46,CDC24	YAL043C,YAL042C-A,YAL042W,YAL041W
I	66666	67791	ChiP_BY_WT_pooled.no_dups_peak_7	889	.	2.43399	90.6587	88.9061	755	CLN3	YAL040C
I	68347	69671	ChiP_BY_WT_pooled.no_dups_peak_8	820	.	2.25923	83.6998	82.0078	696	CYC3	YAL039C
I	71610	73588	ChiP_BY_WT_pooled.no_dups_peak_9	1183	.	2.44018	120.4	118.351	905	CDC19,NA	YAL038W,YAL037C-B,YAL037C-A
I	75651	76970	ChiP_BY_WT_pooled.no_dups_peak_10	860	.	2.41056	87.8185	86.0902	304	RBG1,FUN12	YAL036C,YAL035W
I	77324	77856	ChiP_BY_WT_pooled.no_dups_peak_11	398	.	1.93424	41.1426	39.8111	250	FUN12	YAL035W
I	79039	79494	ChiP_BY_WT_pooled.no_dups_peak_12	332	.	1.91858	34.5251	33.2593	247	FUN12,YAL034C-B	YAL035W,YAL034C-B
I	82712	84482	ChiP_BY_WT_pooled.no_dups_peak_13	113	.	1.42908	Dec.21	Nov.74	1469	POP5,PRP45	YAL033W,YAL032C
I	100120	100713	ChiP_BY_WT_pooled.no_dups_peak_14	807	.	2.47086	82.3941	80.7135	299	MAK16	YAL025C
I	106226	107428	ChiP_BY_WT_pooled.no_dups_peak_15	358	.	1.77564	37.1375	35.8455	501	PMT2	YAL023C
I	107973	109870	ChiP_BY_WT_pooled.no_dups_peak_16	771	.	2.26481	78.8375	77.1854	381	PMT2,FUN26	YAL023C,YAL022C
I	112214	114585	ChiP_BY_WT_pooled.no_dups_peak_17	520	.	Jan.32	53.4499	52.0045	840	CCR4,ATS1,NA	YAL021C,YAL020C,YAL019W-A
I	114751	116392	ChiP_BY_WT_pooled.no_dups_peak_18	350	.	1.78626	36.2902	35.0066	329	NA,FUN30	YAL019W-A,YAL019W
I	128402	132575	ChiP_BY_WT_pooled.no_dups_peak_19	1155	.	2.44079	117.568	115.556	2887	SYN8,DEP1,CYS3,SWC3	YAL014C,YAL013W,YAL012W,YAL011W
I	139243	139805	ChiP_BY_WT_pooled.no_dups_peak_20	69	.	1.34572	7.85378	6.90849	332	TRN1,SSA1	tP(UGG)A,YAL005C
I	142057	143930	ChiP_BY_WT_pooled.no_dups_peak_21	1811	.	3.16946	184.166	181.135	753	EFB1,SNR18,VPS8	YAL003W,snR18,YAL002W
I	166101	166567	ChiP_BY_WT_pooled.no_dups_peak_22	282	.	2.06498	29.4684	28.2548	244	TGA1	tA(UGC)A
I	169591	170278	ChiP_BY_WT_pooled.no_dups_peak_23	629	.	2.12958	64.5023	62.9641	326	ADE1	YAR015W
I	192608	193905	ChiP_BY_WT_pooled.no_dups_peak_24	384	.	1.78426	39.7611	38.4434	499	SWH1	YAR042W
II	36867	38491	ChiP_BY_WT_pooled.no_dups_peak_25	784	.	2.23904	80.1294	78.4676	482	NA,ATP1	YBL100C,YBL099W
II	43225	43782	ChiP_BY_WT_pooled.no_dups_peak_26	384	.	1.966	39.7836	38.4657	250	NA,MRX3	YBL096C,YBL095W,YBL094C
II	44181	44769	ChiP_BY_WT_pooled.no_dups_peak_27	596	.	2.15493	61.1141	59.6037	280	ROX3	YBL093C
II	45344	46996	ChiP_BY_WT_pooled.no_dups_peak_28	1527	.	3.1313	155.299	152.752	1118	RPL32,SCS22	YBL092W,YBL091C-A
II	59655	60610	ChiP_BY_WT_pooled.no_dups_peak_29	2000	.	4.40847	203.402	200.072	398	RPL23A	YBL087C
II	69809	71227	ChiP_BY_WT_pooled.no_dups_peak_30	260	.	1.65314	27.2126	26.0225	1005	NA,ALG3	YBL083C,YBL082C
II	72314	73020	ChiP_BY_WT_pooled.no_dups_peak_31	673	.	2.15795	68.9081	67.3341	344	NA	YBL081W
II	75150	75665	ChiP_BY_WT_pooled.no_dups_peak_32	520	.	2.1439	53.5167	52.0704	285	NUP170	YBL079W
II	87930	90492	ChiP_BY_WT_pooled.no_dups_peak_33	780	.	Feb.06	79.7143	78.0556	2285	NA,SNR56,RPS8A,KTI11	YBL073W,snR56,YBL072C,YBL071C-B,YBL071W-A,YBL071C
II	90761	91443	ChiP_BY_WT_pooled.no_dups_peak_34	314	.	1.84847	32.6653	31.4181	399	NA,AST1	YBL070C,YBL069W
II	111556	113226	ChiP_BY_WT_pooled.no_dups_peak_35	287	.	1.72466	30.0068	28.7874	851	SHP1,PTH2	YBL058W,YBL057C
II	113627	114157	ChiP_BY_WT_pooled.no_dups_peak_36	638	.	2.54546	65.4125	63.8668	279	PTC3	YBL056W
II	114710	115219	ChiP_BY_WT_pooled.no_dups_peak_37	388	.	2.13721	40.139	38.8175	233	PTC3	YBL056W
II	116347	117229	ChiP_BY_WT_pooled.no_dups_peak_38	231	.	1.81604	24.3357	23.1769	670	YBL055C	YBL055C
II	117472	118286	ChiP_BY_WT_pooled.no_dups_peak_39	1023	.	2.56761	104.225	102.348	446	TOD6	YBL054W
II	139205	140163	ChiP_BY_WT_pooled.no_dups_peak_40	327	.	1.91809	33.9732	32.7132	700	FUI1	YBL042C
II	141409	142016	ChiP_BY_WT_pooled.no_dups_peak_41	589	.	2.29474	60.4957	58.9904	258	PRE7	YBL041W
II	158759	159747	ChiP_BY_WT_pooled.no_dups_peak_42	692	.	2.11947	70.7935	69.2036	467	RIB1	YBL033C
II	160083	160915	ChiP_BY_WT_pooled.no_dups_peak_43	450	.	1.87536	46.4682	45.0864	409	HEK2	YBL032W

Peak calling results

chr	start	end	peak_ID	overall_score	strand	fold_change	-log(pvalue)	-log(qvalue)	relative_peak_summit	gene_name	gene_id
I	31479	33115	ChIP_BY_WT_pooled.no_dups_peak_1	429	.	1.88871	44.3616	42.9998	1232	GDH3	YAL062W
I	33537	34528	ChIP_BY_WT_pooled.no_dups_peak_2	610	.	2.08354	62.5989	61.0757	507	BDH2	YAL061W
I	35137	36342	ChIP_BY_WT_pooled.no_dups_peak_3	556	.	2.04075	57.1526	55.6747	425	BDH1	YAL060W
I	45839	46698	ChIP_BY_WT_pooled.no_dups_peak_4	126	.	1.43949	13.7207	Dec.75	433	FLC2	YAL053W
I	57192	60004	ChIP_BY_WT_pooled.no_dups_peak_5	854	.	2.40168	87.1869	85.4642	1022	BOL3,NA,BOL1,GCV3,PTA1	YAL046C,YAL045C,YAL044W-A,YAL044C,YAL043C
I	60315	63277	ChIP_BY_WT_pooled.no_dups_peak_6	704	.	2.15353	72.018	70.4181	1323	PTA1,YAL042C-A,ERV46,CDC24	YAL043C,YAL042C-A,YAL042W,YAL041W
I	66666	67791	ChIP_BY_WT_pooled.no_dups_peak_7	889	.	2.43399	90.6587	88.9061	755	CLN3	YAL040C
I	68347	69671	ChIP_BY_WT_pooled.no_dups_peak_8	820	.	2.25923	83.6998	82.0078	696	CYC3	YAL039C
I	71610	73588	ChIP_BY_WT_pooled.no_dups_peak_9	1183	.	2.44018	120.4	118.351	905	CDC19,NA	YAL038W,YAL037C-B,YAL037C-A
I	75651	76970	ChIP_BY_WT_pooled.no_dups_peak_10	860	.	2.41056	87.8185	86.0902	304	RBG1,FUN12	YAL036C,YAL035W
I	77324	77856	ChIP_BY_WT_pooled.no_dups_peak_11	398	.	1.93424	41.1426	39.8111	250	FUN12	YAL035W
I	79039	79494	ChIP_BY_WT_pooled.no_dups_peak_12	332	.	1.91858	34.5251	33.2593	247	FUN12,YAL034C-B	YAL035W,YAL034C-B
I	82712	84482	ChIP_BY_WT_pooled.no_dups_peak_13	113	.	1.42908	Dec.21	Nov.74	1469	POP5,PRP45	YAL033W,YAL032C
I	100120	100713	ChIP_BY_WT_pooled.no_dups_peak_14	807	.	2.47086	82.3941	80.7135	299	MAK16	YAL025C
I	106226	107428	ChIP_BY_WT_pooled.no_dups_peak_15	358	.	1.77564	37.1375	35.8455	501	PMT2	YAL023C
I	107973	109870	ChIP_BY_WT_pooled.no_dups_peak_16	771	.	2.26481	78.8375	77.1854	381	PMT2,FUN26	YAL023C,YAL022C
I	112214	114585	ChIP_BY_WT_pooled.no_dups_peak_17	520	.	Jan.32	53.4499	52.0045	840	CCR4,ATS1,NA	YAL021C,YAL020C,YAL019W-A
I	114751	116392	ChIP_BY_WT_pooled.no_dups_peak_18	350	.	1.78626	36.2902	35.0066	329	NA,FUN30	YAL019W-A,YAL019W
I	128402	132575	ChIP_BY_WT_pooled.no_dups_peak_19	1155	.	2.44079	117.568	115.556	2887	SYN8,DEP1,CYS3,SWC3	YAL014C,YAL013W,YAL012W,YAL011W
I	139243	139805	ChIP_BY_WT_pooled.no_dups_peak_20	69	.	1.34572	7.85378	6.90849	332	TRN1,SSA1	tP(UGG)A,YAL005C
I	142057	143930	ChIP_BY_WT_pooled.no_dups_peak_21	1811	.	3.16946	184.166	181.135	753	EFB1,SNR18,VPS8	YAL003W,snR18,YAL002W
I	166101	166567	ChIP_BY_WT_pooled.no_dups_peak_22	282	.	2.06498	29.4684	28.2548	244	TGA1	tA(UGC)A
I	169591	170278	ChIP_BY_WT_pooled.no_dups_peak_23	629	.	2.12958	64.5023	62.9641	326	ADE1	YAR015W
I	192608	193905	ChIP_BY_WT_pooled.no_dups_peak_24	384	.	1.78426	39.7611	38.4434	499	SWH1	YAR042W
II	36867	38491	ChIP_BY_WT_pooled.no_dups_peak_25	784	.	2.23904	80.1294	78.4676	482	NA,ATP1	YBL100C,YBL099W
II	43225	43782	ChIP_BY_WT_pooled.no_dups_peak_26	384	.	1.966	39.7836	38.4657	250	NA,MRX3	YBL096C,YBL095W,YBL094C
II	44181	44769	ChIP_BY_WT_pooled.no_dups_peak_27	596	.	2.15493	61.1141	59.6037	280	ROX3	YBL093C
II	45344	46996	ChIP_BY_WT_pooled.no_dups_peak_28	1527	.	3.1313	155.299	152.752	1118	RPL32,SCS22	YBL092W,YBL091C-A
II	59655	60610	ChIP_BY_WT_pooled.no_dups_peak_29	2000	.	4.40847	203.402	200.072	398	RPL23A	YBL087C
II	69809	71227	ChIP_BY_WT_pooled.no_dups_peak_30	260	.	1.65314	27.2126	26.0225	1005	NA,ALG3	YBL083C,YBL082C
II	72314	73020	ChIP_BY_WT_pooled.no_dups_peak_31	673	.	2.15795	68.9081	67.3341	344	NA	YBL081W
II	75150	75665	ChIP_BY_WT_pooled.no_dups_peak_32	520	.	2.1439	53.5167	52.0704	285	NUP170	YBL079W
II	87930	90492	ChIP_BY_WT_pooled.no_dups_peak_33	780	.	Feb.06	79.7143	78.0556	2285	NA,SNR56,RPS8A,KTI11	YBL073W,snR56,YBL072C,YBL071C-B,YBL071W-A,YBL071C
II	90761	91443	ChIP_BY_WT_pooled.no_dups_peak_34	314	.	1.84847	32.6653	31.4181	399	NA,AST1	YBL070C,YBL069W
II	111556	113226	ChIP_BY_WT_pooled.no_dups_peak_35	287	.	1.72466	30.0068	28.7874	851	SHP1,PTH2	YBL058W,YBL057C
II	113627	114157	ChIP_BY_WT_pooled.no_dups_peak_36	638	.	2.54546	65.4125	63.8668	279	PTC3	YBL056W
II	114710	115219	ChIP_BY_WT_pooled.no_dups_peak_37	388	.	2.13721	40.139	38.8175	233	PTC3	YBL056W
II	116347	117229	ChIP_BY_WT_pooled.no_dups_peak_38	231	.	1.81604	24.3357	23.1769	670	YBL055C	YBL055C
II	117472	118286	ChIP_BY_WT_pooled.no_dups_peak_39	1023	.	2.56761	104.225	102.348	446	TOD6	YBL054W
II	139205	140163	ChIP_BY_WT_pooled.no_dups_peak_40	327	.	1.91809	33.9732	32.7132	700	FUI1	YBL042C
II	141409	142016	ChIP_BY_WT_pooled.no_dups_peak_41	589	.	2.29474	60.4957	58.9904	258	PRE7	YBL041W
II	158759	159747	ChIP_BY_WT_pooled.no_dups_peak_42	692	.	2.11947	70.7935	69.2036	467	RIB1	YBL033C
II	160083	160915	ChIP_BY_WT_pooled.no_dups_peak_43	450	.	1.87536	46.4682	45.0864	409	HEX2	YBL032W

- peaks.annotated.bed
- control.bigWig
- bigWig



CEITEC

@CEITEC_Brno

Thank you for your attention!

Vojta Bystry
vojtech.bystry@ceitec.muni.cz