

Supplementary Information For
Regulatory Modules in Collombet Model
(CisView Database)

March 24, 2018

Contents

1	List of Regulatory Modules (CRM) Found With CisView Database	3
2	Transcription Factor Binding Sites (TFBSs)	6
3	Interactions in CRN	36
4	Results	38
4.1	Statistics	38
4.2	Figures	39

1 List of Regulatory Modules (CRM) Found With CisView Database

All sequence positions refer to the Mouse Genome (mm9) of UCSC[2]. Only "high-quality" CRMs, as assessed by the database[3], present in a range of 2kb upstream or 250bp downstream of the gene Transcription Start Site, are listed.

Gene	Position	CisView Region Annotations	CRM names & positions
Mac1	(U008827) chr7:135,206,153-135,271,676	R00037338	(CM07024162)chr7:135,206,027-135,206,277
Gfi1	(U026521) chr5:108145673-108155056	R00035357	(CM05020166)chr5:108,152,989-108,154,363 (CM05020167)chr5:108,155,008-108,156,058
Egr1	(U018551) chr18:35,020,515-35,024,635	R00022016	(CM18007036)chr18:35,019,224-35,019,288 (CM18007038)chr18:35,019,861-35,020,911 (CM18007040)chr18:35,021,897-35,022,048
Egr2	(U011618) chr10:66,998,222-67,004,936	R00026017 & R00026018	(CM10012847)chr10:66,997,886-66,998,387 (CM10012848)chr10:66,999,497-67,000,290 (CM10012850)chr10:67,000,893-67,001,193
PU1 (SPI1)	(U043484) chr15:102,236,319-102,266,835	R00039665	(CM15019435)chr15:102,235,930-102,237,349
Id2	(U033848) chr12:25,778,662-25,780,960	R00042718 & R00042719	(CM12004700)chr12:25,778,632-25,779,894 (CM12004701)chr12:25,780,239-25,780,494 (CM12004702)chr12:25,780,875-25,781,813
Cebpb	(U002854) chr2:167,514,414-167,515,918	R00006580	(CM02038440)chr2:167,513,224-167,513,430

Table 1: List of CRMs associated with each gene present in the model (1)

Gene	Position	CisView Region Annotations	CRM names & positions
Cebpa	(U007965) chr7:35,904,311-35,906,947	R00036394	(CM07005327)chr7:35,903,311-35,904,437
Csf1r	(U018789) chr18:61,245,049-61,292,124	R00022284	(CM18013129)chr18:61,265,306-61,265,556

Table 2: List of CRMs associated with each gene present in the model (2)

Gene	Position	CisView Region Annotations	CRM names & positions
Runx1	(U037315) chr16:92,602,139-92,826,319	R00012990 & R00012992	(CM16019077)chr16:92,696,349-92,699,375 (CM16019121)chr16:92,824,562-92,824,836 (CM16019122)chr16:92,825,634-92,825,730 (CM16019123)chr16:92,825,962-92,826,709
Ikzf1	(U012319) chr11:11,584,982-11,672,929	R00014597 & R00014598	(CM11001391)chr11:11,584,803-11,585,142 (CM11001392)chr11:11,585,199-11,586,404 (CM11001393)chr11:11,586,177-11,586,295
Flt3	(U026965) chr5:148,142,316-148,212,065	R00035871	(CM05029465)chr5:148,210,928-148,211,228 (CM05029466)chr5:148,211,841-148,212,463
Mef2c	(U015066) chr13:83,643,032-83,807,266	R00018519 & R00018520	(CM13016542)chr13:83,641,733-83,642,221 (CM13016543)chr13:83,642,470-83,643,509 (CM13016544)chr13:83,643,896-83,644,606 (CM13016553)chr13:83,661,868-83,665,534
Ets1	(U010230) chr9:32,443,798-32,565,405	R00032801 & R00032802	(CM09005608)chr9:32,443,606-32,444,027 (CM09005636)chr9:32,503,484-32,503,651 (CM09005638)chr9:32,504,834-32,505,134

Table 3: List of CRMs associated with each gene present in the model (3)

Gene	Position	CisView Region Annotations	CRM names & positions
Il7r	(U036001) chr15:9,435,542-9,459,928	R00004036	(CM15001293)chr15:9,459,511-9,459,811
E2A (Tcf3)	(U032010) chr10:79,871,857-79,896,393	R00045128 & R00045129	(CM10015455)chr10:79,883,498-79,883,607 (CM10015456)chr10:79,884,495-79,884,795 (CM10015460)chr10:79,895,598-79,897,177
Foxo1	(U003222) chr3:52,072,258-52,157,143	R00026889	(CM03009529)chr3:52,071,321-52,072,723

Table 4: List of CRMs associated with each gene present in the model (4)

Gene	Position	CisView Region Annotations	CRM names & positions
Ebf1 (Olf1)	(U012493) chr11:44,430,817-44,820,533	R00014807 & R00014808 & R00014809 & R00014812 & R00014817 & R00014822 & R00014824	(CM11009346)chr11:44,429,416-44,429,665 (CM11009348)chr11:44,430,613-44,431,743 (CM11009361)chr11:44,445,294-44,445,388 (CM11009362)chr11:44,445,943-44,446,252 (CM11009363)chr11:44,446,316-44,446,741 (CM11009372)chr11:44,467,940-44,468,415 (CM11009373)chr11:44,468,941-44,469,399 (CM11009374)chr11:44,469,489-44,469,845 (CM11009375)chr11:44,469,985-44,471,311 (CM11009429)chr11:44,551,179-44,551,883 (CM11009430)chr11:44,552,058-44,552,709 (CM11009431)chr11:44,553,030-44,553,675 (CM11009482)chr11:44,639,417-44,639,752 (CM11009483)chr11:44,640,296-44,640,596 (CM11009484)chr11:44,641,176-44,641,337 (CM11009506)chr11:44,682,205-44,682,664

Table 5: List of CRMs associated with each gene present in the model (5)

Gene	Position	CisView Region Annotations	CRM names & positions
Pax5	(U025012) chr4:44,544,253-44,723,534	R00019057	(CM04008155)chr4:44,723,280-44,724,185
Cd19	(U029242) chr7:133,551,961-133,558,386	R00009128	(CM07023775)chr7:133,558,250-133,558,582
Csf1	(U024503) chr3:107,543,965-107,563,387	R00040533	(CM03020953)chr3:107,562,979-107,563,622

Table 6: List of CRMs associated with each gene present in the model (6)

2 Transcription Factor Binding Sites (TFBSs)

Only TFBSs which had a conservation score ≥ 0.5 , and a mismatch score ≤ 0.2 , were listed. CRMs are classified in two main types in CisView: either Promoter, or either distal (Distal), e.g. enhancer, insulator, silencer. TFs in bold are those which are present in the original model.

Gene	CRM name	TFBSs & Strand Orientation	Type
Mac1	CM07024162	SP1 (-), LEF (+)	Type: promoter.
Gfi1	CM05020166	MAZ (+/-), MyoD (+), Egr1 (-)	Type: promoter.
Gfi1	CM05020167	Irf1 (+), AML (+), GABP (-), ETS (+/-), HSF (-), ELK (-), ARP1 (-), MEIS (+), AP1 (+), AP2A (-), FAC1 (-), CEBP (+), NFAT (+), NKX25 (-), CDP (-), MEF2 (+), HAND (-), Nanog1 (+), OLF1 (+), E4BP4 (+), HLF (-), ZID (-), AP2 (+), GABP (-)	Type: promoter.
Egr1	CM18007036	FOXD1 (+), FOXP1 (-), POLYA (+), FOXA (-), CART, NKX61 (-), POU3F (-), MEF2 (-), PU1 (-), HMG (+), LEF (-)	Type: DCRM.

Table 7: Bindings in each detected CRM (1)

Gene	CRM #	TFBSs & Orientation	Notes
Egr1	CM18007038	TAXCREB1 (+/-), ELK (+/-), ETS (+/-), SRF (+/-), YY1 (-), GABP (-), LYF (+), MAZ (+), SP1 (+), ATF6 (-), CREB (-), ATF1 (-), E4F1 (-), PAX3 (+), LRH (-), MEF2 (+), SMAD (+)	Type: promoter.
Egr1	CM18007040	IPF1 (-), HNF3 (+), MAF (+), TAXCREB1 (-), CREB (-), AP4 (+), STAT (+), CDP (+), POU6F (+), HNF1 (-), SRF (+), SF1B (-)	Type: DCRM.

Table 8: Bindings in each detected CRM (2)

Gene	CRM name	TFBSs & Strand Orientation	Type
Egr2	CM10012847	NFKB (-), SP1 (-), EGR1 (-), ARNT (-), MYC (+)	Type: promoter.
Egr2	CM10012848	PAX2 (+), MAZ (+), NCX (+), EF1D (-), GABP (+), ELK (+), TAL1 (-), EGR1 (-), AP4 (+), MAZR (+), SP1 (+), E4BP4 (-), HLF (+), MEF2 (+), RSRFC4 (+/-), NKX25 (-), FOXO (+/-), NKX3A (-), FOXD (+), FOXA (-), SRY (+), FOXJ1 (+), HNF3 (-), TATA (+), NFAT (-), GATA (-), TST (-), NFY (+), IRF (+), NKX61 (+)	Type: DCRM.
Egr2	CM10012850	AREB6 (+), ER (+), NF1 (-), SF1A (+), LRH (+), Gfi (+), UBP2 (+), AP4 (+/-), MyOD (-), MEIS (+), YY1 (+), CEBP (-), MZF (-), AIRE (+), LEF (+), KLF (+), PAX4 (9), RREB (-)	Type: DCRM.

Table 9: Bindings in each detected CRM (3)

Gene	CRM name	TFBSs & Strand Orientation	Type
PU1 (SPI1)	CM15019435	YY1 (-), NFY (-), Nanog1 (+/-), ATF6 (+), CREB (-), ATF1 (-), PAX4 (-), PBX (-), CTCFB (-), HMX (-), HNF4 (+), TEF1 (-), SP1 (+/-), GABP (-), ETS (+/-), TEL (-), PU1 (-), GCM (+), FOXD1 (+), Gata (+), PITX (+), NCX (-), MAZ (+/-), ATF1 (-), HELIOS (-), MAZR (+/-), MYB (+), GCNF (-), HEN1 (-), AP4 (+), KLF (+), PAX4 (+), ZF5 (+), AP2 (-), AP2A (+), LF1 (+), ZIC (-)	Type: promoter.

Table 10: Bindings in each detected CRM (4)

Gene	CRM name	TFBSs & Strand Orientation	Type
Id2	CM12004700	AIRE (-), ATF (+), POLYA (-), POU3F (+), MRF2 (+), FOXD1 (+), FOXA (+/-), CDC5 (-), PBX (+/-), GFI (+), HMG (+/-), LEF (+/-), FOXD (+), FOXO (+), PAX4 (-), POU6F (+), BRN (-), TATA (+/-), FXR (+), IRF (+/-), OCT4A (-), NKX3A (+), CREB (-), ATF1 (-), Oct4/Sox2 (+/-), CDC5 (+), CART (+), CHX (+), LHX (+/-), NKX61 (-), COUP (-), Pax8 (-), E4BP4 (+), HLF (-), NKX3A (+), STAF (-), UBP2 (-), HAND (+), Pax6 (+), Pax8 (-), CEBP (+), STAF (-), IK (+), MYB (+), CDX (-), TBX5 (-), Myod (-), GABP (-), BARBIE (-), SF1B (+), ...	Type: 3'UTR-CRM.

Table 11: Bindings in each detected CRM (5)

Gene	CRM name	TFBSs & Strand Orientation	Type
Id2	CM12004700	... GR (+), AR (+), ARNT (+), GATA (+), LYF (+), PU1 (+), RP58 (-), SRF (+), NKX61 (-), BRN (-), Oct (+), MEF2 (+/-), MZF (+), HELIOS (+), LYF (-), XVENT (+), HLF (-), FOXJ2 (+), Sox9 (+), COMP (+), PBX (-), NFAT (+), EVI (+), Pax3 (-), CART (+), Gata (+), LHX (+), CHX (-), BRACH (-)	Type: 3'UTR-CRM.
Id2	CM12004701	POU6F (+), IRF (-)	Type: promoter.

Table 12: Bindings in each detected CRM (5')

Gene	CRM name	TFBSs & Strand Orientation	Type
Id2	CM12004702	SP1 (+), WHN (-), NFY (-), MEF3 (+), NF1 (-), NFY (+), Pax4 (+), MAZ (+), HEN1 (+), AP4 (+), HOXA5 (+/-), IPF1 (+), CDP (+), SF1B (+), SF1 (-), MTF1A (+), Oct (-), POU3F (-), E4BP4 (-), LRH (-), GCNF (-), Nanog (+), POU6F (+), BRN (-), LHX (+), CHX (+/-), Oct4/Sox2 (+), TEF (-), TEF1 (+)	Type: promoter.
Cebpb	CM02038440	IK (+), MAZ (-), MAZR (-), PAX4 (-), KLF (-), LRH (+), TGIF (+), SP1 (+), MTF1B (-), SF1 (-), LXR2 (+), SREBP (+), GATA (+), CEBP (+/-), Nanog1 (-), ARP1 (-)	Type: DCRM.
Cebpa	CM07005327	HMX (+), AP2REP (+), ZID (+), TATA (+), Nanog1 (+)	Type: promoter.

Table 13: Bindings in each detected CRM (6)

Gene	CRM name	TFBSs & Strand Orientation	Type
Csflr	CM18013129	LEF (+), HMG (-), Sox18 (-), SRY (-), Foxo (+), NFAT (-), MZF (+), RFX1 (+), RFX (+)	Type: promoter.
Runx1	CM16019077	Cdx (+), MEIS (-), ER (-), FOXO (+/-), MRF2 (+), FOXP1 (-), FOXD1 (+), FOXA (-), TATA (+), POU6F (-), FAC1 (+), FOXP1 (-), CDP (+), Nanog (+), IPF1 (+), FOXJ2 (-), HNF3 (+), NKX3A (+), LRH (-), HMG (+), LEF (-), HEN1 (-), AP4 (-), LYF (+), IK (+), AML (-), AIRE (+), ZID (+/-), MAF (+/-), AP1 (+/-), PU1 (+/-), Myod (+), Egr1 (-), RFX1 (+), TGIF (+), FOXD (+), Sox9 (+), IRF (+), FXRDR1 (+), ...	Type: promoter.

Table 14: Bindings in each detected CRM (7)

Gene	CRM name	TFBSs & Strand Orientation	Type
Runx1	CM16019077	..., BARBIE (+), PIT (+/-), MEF3 (+), CP2 (-), TST (+), NFAT (+), Nanog1 (+/-), ARNT (+), NFY (+/-), GFI (+), ZF5 (-), Stat (+), CTCFA (+)	Type: promoter.
Runx1	CM16019121	TEF (+), FAC1 (-), LEF (+), HMG (-), Sox18 (-), LEF (-), FOXD1 (+), AP1 (-), FOXP3 (-), POU6F (-), FOXJ2 (-), AP1 (+), SRY (+), FOXO (-)	Type: DCRM.
Runx1	CM16019122	MEF3 (-), XVENT (-), FOXJ2 (-), TST (-), MAZ (-), PU1 (-)	Type: DCRM.
Runx1	CM16019123	FOXA (-), SRY (+), FOXO (+/-), FOXP3 (-), XVENT (-), CEBP (+), Sox9 (-), HNF1 (+), NFAT (+/-), IRF (+), AML (-), NRSF (+), COUP (+), FOXD (-), CART (+), ZIC (-), NCX (-), MZF (-), NFY (-), ATF6 (-), SREBP1 (+), SREBP (-), ARNT (+), CREB (-), ...	Type: promoter.

Table 15: Bindings in each detected CRM (8)

Gene	CRM name	TFBSs & Strand Orientation	Type
Runx1	CM16019123	..., MYC (+), AP1 (+), ATF (+), PPAR (-), Myod (-), AP2 (+), FAC1 (+), FOXD (+), HNF3 (-), FOXA (-), FOXP1 (-), GR (+), XVENT (+), RREB (-), AML (-), OCT4B (+), PU1 (+), FOXD1 (-), GATA (+)	Type: promoter.
Ikzf1	CM11001393	E2F (-), Sox9 (+), Sox (-), HMG (+), PU1 (+), CEBP (+), BARBIE (+), IK (+), HMX (+)	Type: DCRM.
Ikzf1	CM11001391	∅	Type: DCRM.
Ikzf1	CM11001392	NFY (-), RREB (-), AIRE (+), LYF (+), IRF (+), HELIOS (+/-), PAX4 (-), KLF (+/-), PU1 (+), SMAD (-), E2F (-), Sox9 (+), Sox (-), HMG (+), CEBP (+), BARBIE (+), IK (+), HMX (+)	Type: promoter.

Table 16: Bindings in each detected CRM (9)

Gene	CRM name	TFBSs & Strand Orientation	Type
Flt3	CM05029465	∅	Type: promoter.
Flt3	CM05029466	CP2 (+), MEF3 (+), AP2 (+), AP2A (+), HEN1 (+), Myod (+), WHN (-), E4F1 (-), AML (+), ARNT (-), MYC (+)	Type: promoter.
Mef2c	CM13016542	XVENT (-), IPF1 (+), FAC1 (-), FOXD1 (+), Oct4/SOx2 (-), PU1 (+), MAF (-), Sox9 (+), HMG (+), FOXD (+/-), FOXA (+), PAX4 (-), AFP1 (-), CHX (+), FOXP1 (-), POLYA (+/-), HMX (-), AP1 (+/-), GR (-), AIRE (+), STAF (-), UBP2 (-), CP2 (-), FXRDR3 (+), GATA (-), ZID (-), FOXO (+/-), LEF (-), FOXP1 (+), MEF2 (+), RORA (-)	Type: DCRM.

Table 17: Bindings in each detected CRM (10)

Gene	CRM name	TFBSs & Strand Orientation	Type
Mef2c	CM13016543	<p> Nanog1 (-), Oct4/Sox2 (-), BRN (+/-), SRF (-), LYF (+), PBX (-), AP1 (-), PITX (-), HNF1 (-), FOXJ2 (-), E2F (-), XVENT (-), HMX (+/-), NFAT (+/-), NKX61 (-), IPF1 (+), IRF (+/-), HOXA5 (+), HOXA4 (+), LHX (+), CHX (-), FOXP1 (-), NKX25 (-), HSF (-), Stat (+/-), PU1 (+), TATA (-), GABP (-), KLF (+), Pax4 (+/-), COUP (-), LEF (-), SF1B (+), LXR2 (-), Myod (+/-), TAXCREB1 (+), Pax8 (-), Pax3 (-), ARP1 (-), HNF4 (-), HMG (+), AP4 (+), HAND (+), MEISHOX (-), AREB6 (-), TGIF (+), MEIS (-), Oct4B (+), HMG (+), Sox18 (+), CREB (+), ETS (+/-), TEL (-), MEF3 (+), ... </p>	Type: promoter.

Table 18: Bindings in each detected CRM (11)

Gene	CRM name	TFBSs & Strand Orientation	Type
Mef2c	CM13016543	..., CEBP (+), NFY (+), NKX3A (-)	Type: promoter.
Mef2c	CM13016544	CHX (+), LHX (-), POU6F (+/-), AFP1 (+), FOXJ2 (-), NKX61 (-), TEF1 (-), E4BP4 (-), NFAT (+/-), HOXA5 (+), FOXD1 (+), E4BP4 (-), MRF2 (-), FOXA (+), FOXP1 (+), POLYA (-), IRF (-), CIZ (-), AR (-), CART (+), NKX61 (+), CHX (+), LHX (+/-), MEF2 (+/-), NKX25 (+), GCNF (+), ARP1 (-), LRH (+), Sox9 (-), HMG (-), IRF (+), FOXD1 (-), LEF (-), Oct4/Sox2 (-), HNF3 (-), YY1 (+), NKX3A (+), CDP (+), HMG (-), TEF1 (-), Oct4A (+), TATA (+), CDP (-), NFY (+), AP1 (-), FOXD1 (+), CDX (-), IRF1 (-)	Type: promoter.

Table 19: Bindings in each detected CRM (11')

Gene	CRM name	TFBSs & Strand Orientation	Type
Mef2c	CM13016553	Mef2 (+/-), POU3F (-), NCX (+/-), MSX (-), RFX (-), UBP2 (-), CART (+), LHX (+/-), NKX61 (+), CHX (+), HSF (-), PU1 (-), TEF1 (-), EF1D (-), NFAT (+), HNF1 (-), GATA (-), ETS (+), AP4 (+), LXR1 (+), ER (-), LYC (+), FOXO (+), Nanog1 (+/-), RORA (-), FXRDR3 (+), BRACH (+), EF1D (-), TBX5 (+), PX (-), IRF (-), IPF1 (-), Nanog (-), SP1 (+), Pax6 (+), Pax4 (+/-), Pax8 (+/-), AP1 (-), NFY (-), NFKB (+), SMAD (+), TGIF (+), TEF1 (-), MEISHOX (+/-), Pax2 (+), LXR2 (+), MAF (-), STAF (+), CHX (+), Oct4/Sox2 (-), NKX25 (+), PBX (+/-), ...	Type: promoter.

Table 20: Bindings in each detected CRM (11")

Gene	CRM name	TFBSs & Strand Orientation	Type
Mef2c	CM13016553	... GFI (+/-), TEF1 (+), FOXD1 (+), BRN (+), LHX (-), POU6F (+/-), NKX61 (-), CHX (+), SREBP (-), KLF (+), FOXD (-), HNF3 (+), FOXJ1 (-), SRY (-), FOXA (+), POLYA (-), RFX (-), SP1 (-), MZF (-), MAZ (+/-), RFX1 (+), RFX (+), HIF1 (+), MAZ (-), HELIOS (-), GABP (-), STAT (-), HNF1 (+), NKX3A (-), PITX (-), NKX25 (-), ZF5 (+/-), Egr1 (-), HMX (-), COMP (+), LEF (+/-), HMG (-), HSF (-), ARP1 (+), BARBIE (-), HMG (-), AP1 (+), LRH (+), HNF4 (-), FOXJ2 (+), HOXA5 (-), SRF (-), E4BP4 (+), NKX3A (+), FOXP1 (-), MRF2 (+), NFY (-), CDC5 (+), ...	Type: promoter.

Table 21: Bindings in each detected CRM (11")

Gene	CRM name	TFBSs & Strand Orientation	Type
Mef2c	CM13016553	... FAC1 (-), TEF1 (+), IRF (+), TAL1 (-) , RP58 (+), GABP (-), MEIS (+/-), SRY (+/-), LXR1 (+), HNF1 (+), MEIS (-), OCT (+/-), OCT4A (-), TEF1 (-), ARNT (+), GR (+), AR (+), ER (+), ARP1 (+), RORA (-), SF1 (+), SF1A (-), LRH (-), GCNF (-), HE- LIOS (-), CDC5 (+), BARBIE (+), CHX (-), HOXA4 (-), FOXJ1 (+), HNF3 (-), AP1 (-), IRF (-), HLF (+), LEF (-), IRF (-), Sox9 (-), TATA (+), POU6F (-), HNF3 (+), MSX (+), LRH (-), HNF4 (-), MSX (+), PITX (-), MYOD (-), LRH (-)	Type: promoter.

Table 22: Bindings in each detected CRM (11")

Gene	CRM name	TFBSs & Strand Orientation	Type
Ets1	CM09005608	FoxJ2 (+), MRF2 (-), AREB6 (+), NFAT (+)	Type: promoter.
Ets1	CM09005636	IRF (-), LYF (+), NFY (-), SP1 (-)	Type: DCRM.
Ets1	CM09005638	NFAT (+), Sox9 (+), FOXO (-), PBX (+)	Type: promoter.

Table 23: Bindings in each detected CRM (12)

Gene	CRM name	TFBSs & Strand Orientation	Type
Il7r	CM15001293	AIRE (-), AML (+), COUP (+), HMX (+), HAND (-), ETS (+/-)	Type: promoter.
E2A (Tcf3)	CM10015455	CR (-), AR (-), MYB (+), CP2 (+), ELK (-), ETS (-), RFX1 (+), IRF (+)	Type: DCRM.
E2A (Tcf3)	CM10015456	\emptyset	Type: promoter.
E2A (Tcf3)	CM10015460	SREBP (-), HNF4 (-), ZF5 (+), TAX- CREB1 (-), SREBP1 (+), MYOD (-), EGR1 (+), SP1 (-), PAX4 (-), KLF (-), GATA (+)	Type: promoter.

Table 24: Bindings in each detected CRM (13)

Gene	CRM name	TFBSs & Strand Orientation	Type
Foxo1	CM03009529	AREB6 (+), EGR1 (+), SP1 (+), FOXO (+/-), SRY (+), CTCFA (-), NKX25 (-), TST (+), HNF1 (-), HOXA5 (-), MRF2 (+), KLF (-), MAZR (-), OLF1 (+), LF1 (-), Nanog1 (+), NFY (-), ZF5 (+), SMAD (+), MYB (+), WHN (-), MAZ (+)	Type: promoter.
Ebf1 (Olf1)	CM11009346	CREB (-), TATA (-), PBX (+), Gfi1 (+/-), HMG (+), ARP1 (-), LEF (-), HAND (+), NFY (-), YY1 (+), BRN (-), PU1 (-)	Type: DCRM.

Table 25: Bindings in each detected CRM (14)

Gene	CRM name	TFBSs & Strand Orientation	Type
Ebf1 (Olf1)	CM11009348	COUP (+), IRF (+/-), NF1 (+), HOXA5 (-), POU3F (+), IPF1 (-), GCM (-), COMP (+), NFY (+/-), TAXCREB1 (+), ATF1 (+), CREB (+), HLF (+), OLF1 (+), Oct (+), MEF2 (+), MAZR (-), RFX1 (+), WHN (-), FOXD (+), PU1 (+), NFAT (+), CEBP (+), NFAT (+), IRF (+), MEISHOX (+), FOXO (+), TGIF (+), ZIC (+), SREBP1 (-), NFKB (-), GR (+), AR (+), RREB (+), ZIC (-), MYB (+), CIZ (-), NFKB (+), HELIOS (-), MEF3 (+), Oct4/Sox2 (-), ARP1 (-), AML (-), HMG (+), LEF (-), MAZ (+)	Type: promoter.

Table 26: Bindings in each detected CRM (14')

Gene	CRM name	TFBSs & Strand Orientation	Type
Ebf1 (Olf1)	CM11009361	Stat (-), HNF1 (-), POLYA (-), Sox9 (+), HLF (+), FOXJ2 (+), MRF2 (-)	Type: DCRM.
Ebf1 (Olf1)	CM11009362	MEF2 (+/-), TST (+/-), FOXP1 (+), POLYA (-), MRF2 (-), CART (+), TAL1 (+), MYOD (-), IRF (-), LHX (+), CHX (-), NKX61 (+), Nanog (-), HOXA4 (-), NFY (-)	Type: promoter.
Ebf1 (Olf1)	CM11009363	AREB6 (+), HSF (+), Oct4B (+), NCX (+), RORA (+), COUP (-), PAX2 (+), GCNF (+), LRH (+), FOXD (-), FOXO (+/-), STAF (-), LXR2 (-), TGIF (-), COUP (+), MEISHOX (+), OCT (-)	Type: DCRM.

Table 27: Bindings in each detected CRM (15)

Gene	CRM name	TFBSs & Strand Orientation	Type
Ebf1 (Olf1)	CM11009372	FXR (+), HNF1 (+), SREBP1 (+), TAX-CREB1 (+), CP2 (-), RFX1 (-), CTCFC (-), AREB6 (+), SMAD3B (-), ETS (+/-), TEL (+), FOXD1 (-), OCT (-), POU3F (-), Oct4/Sox2 (-), NFAT (-), MEISHOX (-), TGIF (+), PAX2 (+), MEF3 (+), HOXA5 (-), SRF (-), CART (+), CHX (-), NKX61 (+), HOXA4 (-), ZID (-), GABP (-), GR (-)	Type: DCRM.

Table 28: Bindings in each detected CRM (16)

Gene	CRM name	TFBSs & Strand Orientation	Type
Ebf1 (Olf1)	CM11009373	TATA (+), Nanog (-), LHX (+), POU6F (+), CART (-), NKX61 (+), POLYA (-), STAT (+/-), NKX25 (-), CHX (+), PBX (+), TST (-), ARP1 (-), HMX (+), PITX (+), FOXD1 (-), FOXA (+), FOXO (+), NFAT (-), HEN1 (-), AP4 (+), RP58 (+), MSX (-), BAR- BIE (+), HMG (+), LEF (-), RFX1 (+), HLF (-), Sox9 (+)	Type: promoter.
Ebf1 (Olf1)	CM11009374	HNF1 (+), MTF1A (-) , RP58 (-), SREBP (-), FOXJ2 (-), Sox9 (-), HMG (-), Nanog (+), NKX61 (-), IPF1 (+), CART (+), FOXD1 (-), Oct4A (-), Gfi (-)	Type: DCRM.

Table 29: Bindings in each detected CRM (16')

Gene	CRM name	TFBSs & Strand Orientation	Type
Ebf1 (Olf1)	CM11009375	Oct4/Sox2 (-), BRN (+/-), FOXP1 (-), POLYA (+/-), MEF2 (+/-), POU3F (-), OCT (+/-), NFAT (-), FOXD (-), FOXA (+/-), HLF (-), SRY (-), FOXO (-), FOXD1 (+), TATA (+), TST (+), NERF (-), ETS (+/-), TEL (-), COUP (-), PPAR (+), HNF4 (-), HMG (+), LEF (+/-) , SF1B (+), SF1 (+/-), ER (+/-), GCNF (+), RORA (-), HOXA5 (+), MAF (+), FXRIR1 (+), MZF (+), AP1 (+/-), CDP (+), MSX (-), MYOD (+/-), Nanog (+), NKX61 (+/-), IPF1 (+), NKX3A (+), NCX (+), CART (-), POU3F (+/-), OCT4A (+), CHX (+/-), ²⁰ E1X (+/-), FOXJ1 (-), PU1 (-), HNF3 (-), ...	Type: promoter.

Table 30: Bindings in each detected CRM (17)

Gene	CRM name	TFBSs & Strand Orientation	Type
Ebf1 (Olf1)	CM11009375	..., NKX3A (+/-), NKX25 (+/-), CEBP (+), FOXJ2 (-), HSF (+), TST (-), RSRFC4 (-), PBX (-), CDP (-), ARP1 (-), LEF (-), FXR (+), CDC5 (-), LHX (-), MRF2 (+), E4BP4 (-), TEF1 (+), STAT (+), MEF3 (-), IRF (-), MSX (+), HOXA4 (-), HNF4 (-), RREB (+), MIF (-), RFX (+), FAC1 (+), LRH (+), SRF (+), GCNF (-), STAT (-), SMAD3B (-)	Type: promoter.
Ebf1 (Olf1)	CM11009429	MEF2 (-), TATA (-), POU6F (+/-), NKX25 (+), CART (+/-), LHX (-), BRN (+/-), IPF1 (-), FOXJ2 (+/-), POU3F (+/-), MRF2 (+/-), GCNF (-), E4BP4 (+/-), NKX3A (+), IRF (-), RFX1 (-), FOXO (+/-), ³⁰ HOXA5 (+), E4BP4 (-), LHX (+), MRF2 (+), CDX (+), ...	Type: DCRM

Table 31: Bindings in each detected CRM (17')

Gene	CRM name	TFBSs & Strand Orientation	Type
Ebf1 (Olf1)	CM11009429	..., FOXA (+), HNF3 (+), FOXJ1 (-), SRY (-), FOXD (-), OCT (+), POLYA (-), NFY (-) , EVI (+), NFAT (+), Sox9 (+), NKX61 (+), SRF (+), AREB6 (-), GFI (+), HE- LIOS (-), GCM (+), TST (-), STAT (-), RFX (-), RP58 (+), MRF2 (-), ATF1 (+), CREB (+), E4BP4 (+), HLF (-), TBX5 (-), BRACH (-), EF1D (+), HSF (+), IRF (+)	Type: DCRM

Table 32: Bindings in each detected CRM (17")

Gene	CRM name	TFBSs & Strand Orientation	Type
Ebf1 (Olf1)	CM11009430	MRF2 (-), LEF (+), HMG (-), Sox18 (-), HNF1 (+), FOXD1 (-), TST (-), TATA (-), FOXO (+), FOXD (+/-), HNF3 (+), HNF1 (+), Nanog1 (-), IPF1 (-), MEISHOX (-), TGIF (+), MEIS (-), IK (-), MEF2 (+), NFAT (+/-), PU1 (+), NKX25 (+), TEF1 (-), NFY (+), GR (-), AR (-), Oct4/Sox2 (+)	Type: DCRM.
Ebf1 (Olf1)	CM11009431	BRN (+), TATA (-), POU6F (-), LHX (+), CHX (-), NKX61 (+), HOXA4 (-), AR (-), XVENT (+/-), GFI (+/-), HAND (+), ER (-), NFAT (-), Sox9 (-), HMG (-), HLF (+), MEL18 (-), MEF3 (-), Oct4/Sox2 (+/-), HNF4 (+), CDC5 (+/-), HOXA5 (+), HELIOS (+), 323 PBX (+), CDP (-), MRF2 (+), IPF1 (+), ...	Type: promoter

Table 33: Bindings in each detected CRM (18)

Gene	CRM name	TFBSs & Strand Orientation	Type
Ebf1 (Olf1)	CM11009431	..., E4BP4 (-), AP2REP (-), FAC1 (+), FOXD (+), FOXA (+/-), SRY (+), FOXO (+/-), FOXJ1 (+), HNF3 (-), SRY (+), FOXP1 (-), MSX (+), NKX61 (+), HOXA4 (-), NFAT (+), HAND (-), FOXA (+), SRY (-), POLYA (-), MEF2 (+/-), BRN (-), MIF (-), RFX (-), Oct4/Sox2 (+)	Type: promoter

Table 34: Bindings in each detected CRM (18')

Gene	CRM name	TFBSs & Strand Orientation	Type
Pax5	CM04008155	PU1 (+/-), IRF (-), NFAT (+/-), SRY (-), FOXO (+), FOXA (+), POLYA (-), FOXD1 (-), EVI (-), POU6F (-), GABP (-), FXRIR1 (+), EF1D (-), WHN (-), AP1 (-), MYC (-), ATF6 (+), TATA (-), MYB (-), CTCFB (+), SP1 (-)	Type: promoter.
Cd19	CM07023775	TEF1 (+), PU1 (+), LF1 (-), AP2 (+), AP2A (-), STAF (-), CP2 (-), Nanog1 (-), ZIC (+)	Type: promoter
Csf1	CM03020953	HEN1 (-), AP4 (+), IRF (+/-), MZF (-), PAX3 (+), MYC (-), IK (+), FOXD (+), HNF3 (-), ZID (+), NFkB (+), LRH (+)	Type: promoter

Table 35: Bindings in each detected CRM (19)

Gene	CRM name	TFBSs & Strand Orientation	Type
Ebf1 (Olf1)	CM11009482	MYOD (+/-), AP4 (+), FOXD1 (+), FOXA (-), FOXP3 (-), POU6F (-), OCT4B (+), SREBP1 (+), SMAD3B (+), PU1 (+), MZF (-), NFAT (+), HNF1 (+), AP1 (+), CEBP (-), FOXP1 (-), IRF (+), EVI (-), GATA (-), ETS (+/-), BARBIE (-), HMG (+), Sox9 (-), FOXO (+)	Type: DCRM.
Ebf1 (Olf1)	CM11009483	∅	Type: promoter
Ebf1 (Olf1)	CM11009484	Nanog1 (-), CDP (-), NF1 (+), ARP1 (+), LRH (-), GFI (+/-)	Type: promoter
Ebf1 (Olf1)	CM11009506	CDP (+), SREBP1 (+), FXRDR1 (-), MYOD (-), IK (+), CHX (+), LF1 (-), LEF (+), HMG (-), Sox18 (-), SRY (-), FOXO (+), NKX25 (+), NFAT (-), MSX (+), CEBP (+), TEF1 (-), ETS (-), ATF6 (+), SREBP1 (-), ARNT (-), MYC (+)	Type: promoter

Table 36: Bindings in each detected CRM (20)

3 Interactions in CRN

Gene	TF Bindings to ≥ 1 CRM	Regulators in Original Model[1]	Number of Matches Overall TEs
Mac1	\emptyset	Spi1	$\frac{0}{1}$
Gfi1	Egr1, Ets1, Cebpa, Cebpb, Mef2c, Ebf1	Pax5, Cebpa, Egr2, Ikzf1	$\frac{1}{4}$
Egr1	Mef2c, PU1, Ets1	\emptyset	1

Table 37: Interactions in the resulting CRN (1)

Gene	TF Bindings to ≥ 1 CRM	Regulators in Original Model[1]	Number of Matches Overall TEs
Egr2	Egr1, Mef2c, Foxo1, Gfi1, Ceba, Cebpb	Spi1, Gfi1, Pax5	$\frac{1}{3}$
PU1 (Spi1)	Ets1, PU1	Spi1, Runx1, Cebpa, Ikzf1, Gfi1, Cebpb, Foxo1	$\frac{1}{7}$
Id2	Gfi1, Foxo1, Cebpa, Cebpb, Ikzf1 (IK1), PU1, Mef2c, Runx1 (EVI1)	Cebpa, Cebpb, Spi1, Ebf1, Gfi1	$\frac{4}{5}$

Table 38: Interactions in the resulting CRN (2)

Gene	TF Bindings to ≥ 1 CRM	Regulators in Original Model[1]	Number of Matches Overall TFs
Cebpb	Ikzf1, Cebpa, Cebpb	Spi1, Cepba, Cebpb	$\frac{2}{3}$
Cebpa	\emptyset	Runx1, Foxo1, Cebpa	$\frac{0}{3}$
Csf1r)	Foxo1	Spi1, Pax5	$\frac{0}{2}$

Table 39: Interactions in the resulting CRN (3)

Gene	TF Bindings to ≥ 1 CRM	Regulators in Original Model[1]	Number of Matches Overall TFs
Runx1	Foxo1, Ikzf1, PU1, Egr1, Gfi1, Cebpa, Cebpb	\emptyset	1
Ikzf1	PU1, Cebpa, Cebpb, Ikzf1	Mef2c, Pax5	$\frac{0}{2}$

Table 40: Interactions in the resulting CRN (4)

Gene	TF Bindings to ≥ 1 CRM	Regulators in Original Model[1]	Number of Matches Overall TFs
Flt3	\emptyset	Ikzf1, Spi1, Pax5	$\frac{0}{3}$
Mef2c	Mef2c, Foxo1, PU1, Ets1, Cebpa, Cebpb, Gfi1, Egr1	Spi1, Cebpa, Cebpb	1
Ets1	Foxo1	E2A	$\frac{0}{1}$

Table 41: Interactions in the resulting CRN (5)

Gene	TF Bindings to ≥ 1 CRM	Regulators in Original Model[1]	Number of Matches Overall TFs
Il7r	Ets1	Spi1, Ef2c, Cebpa, Cebpb, Ebf1, Foxo1	$\frac{0}{6}$
E2A	Ets1, Egr1	Ikzf1, Spi1, Ebf1, Pax5, E2A	$\frac{0}{5}$
Foxo1	Egr1, Foxo1, Ebf1	Cebpa, Cebpb, E2A, Foxo1	$\frac{1}{4}$

Table 42: Interactions in the resulting CRN (6)

Gene	TF Bindings to ≥ 1 CRM	Regulators in Original Model[1]	Number of Matches Overall TFs
Ebf1 (Olf1)	Gfi1, PU1, Mef2c, Ebf1, Cebpa, Cebpb, Foxo1, Ets1, Runx1 (EVI1), Ikzf1	Ebf1, E2A, Foxo1, Ets1, Runx1, Spi1, Cebpa, Cebpb, Pax5	$\frac{7}{9}$
Pax5	PU1, Foxo1, Runx1	Ebf1, Cebpa, Cebpb, Pax5, Id2	$\frac{0}{5}$
Cd19	PU1	Pax5, Cebpa	$\frac{0}{2}$
Csf1	Ikzf1	\emptyset	1

Table 43: Interactions in the resulting CRN (7)

4 Results

4.1 Statistics

- Number of TF matches overall all TF bindings in the original model: 29.4%
- Average percentage of TF matches overall TF bindings for each gene: 23.4%

4.2 Figures

See script *rm_analysis.R* in the code section.

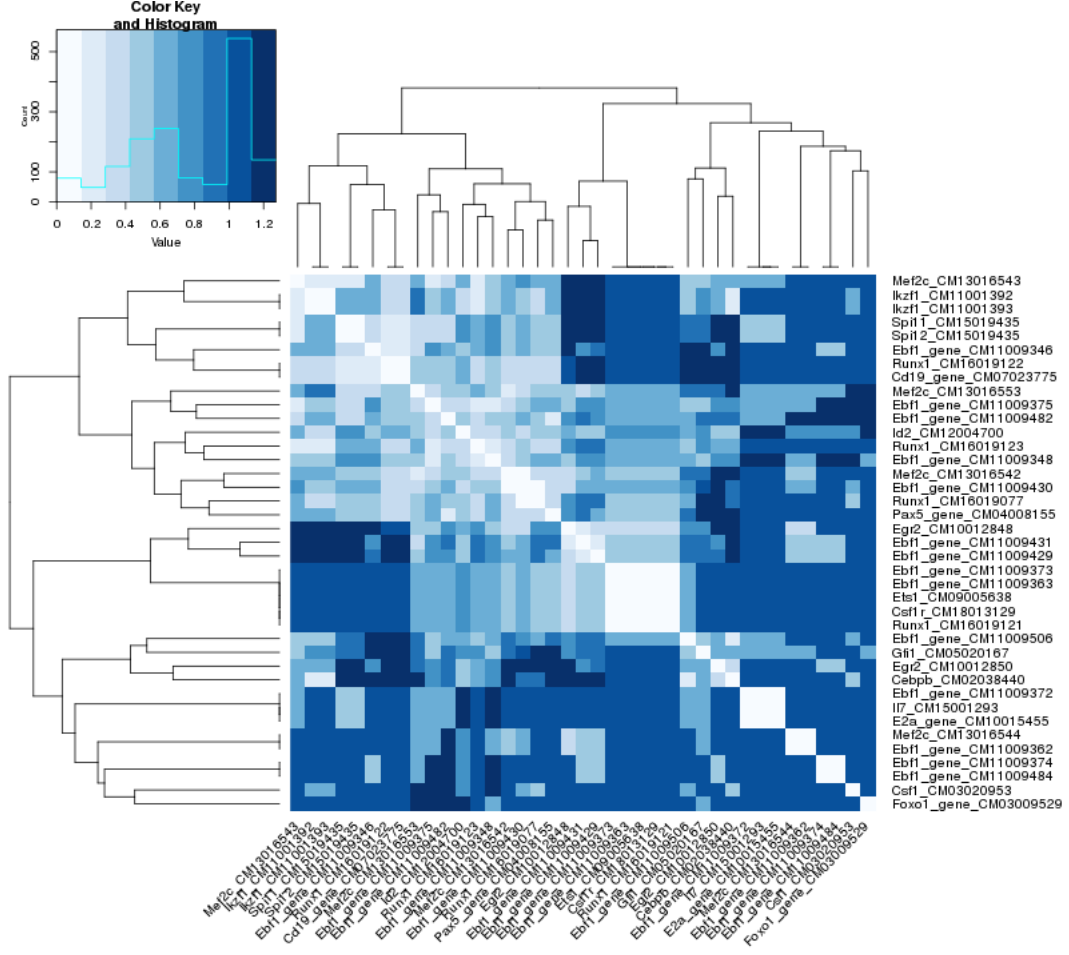


Figure 1: Spearman's ρ correlation matrix between all the CRMs (for TFs involved in the original Krone model). Original matrix M is such as $M_{i,j} = \delta_{TF_j \in CRM_i}$ with δ being the Kronecker symbol (associated function: *heatmap_c*).

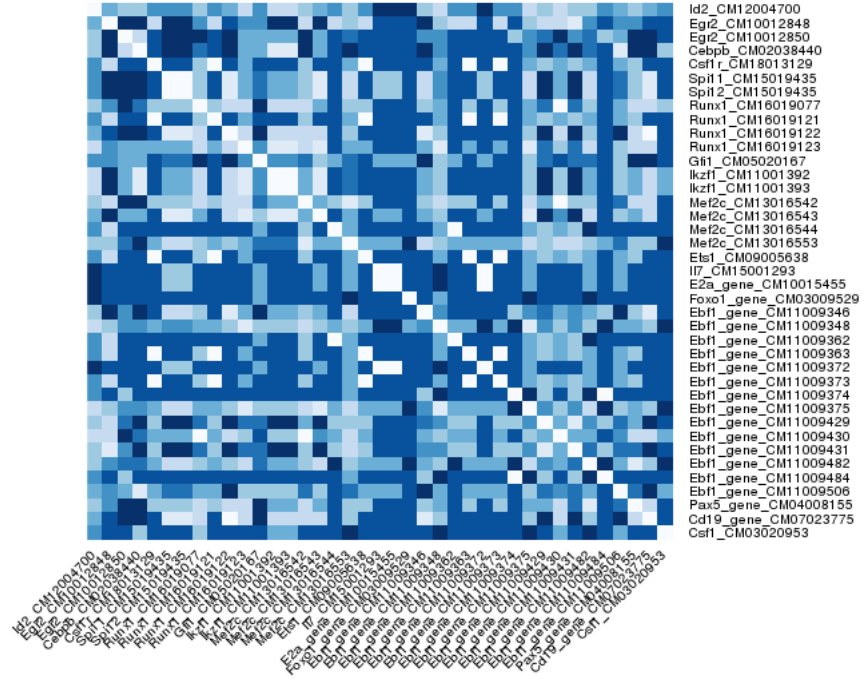
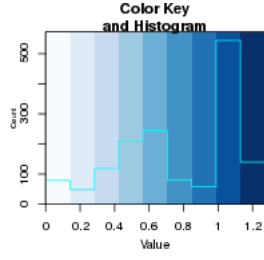


Figure 2: Spearman's ρ correlation matrix between all the CRMs (for TFs involved in the original model). Original matrix M is such as $M_{i,j} = \delta_{TF_j \in CRM_i}$ with δ being the Kronecker symbol (associated function: *heatmap_c2*).

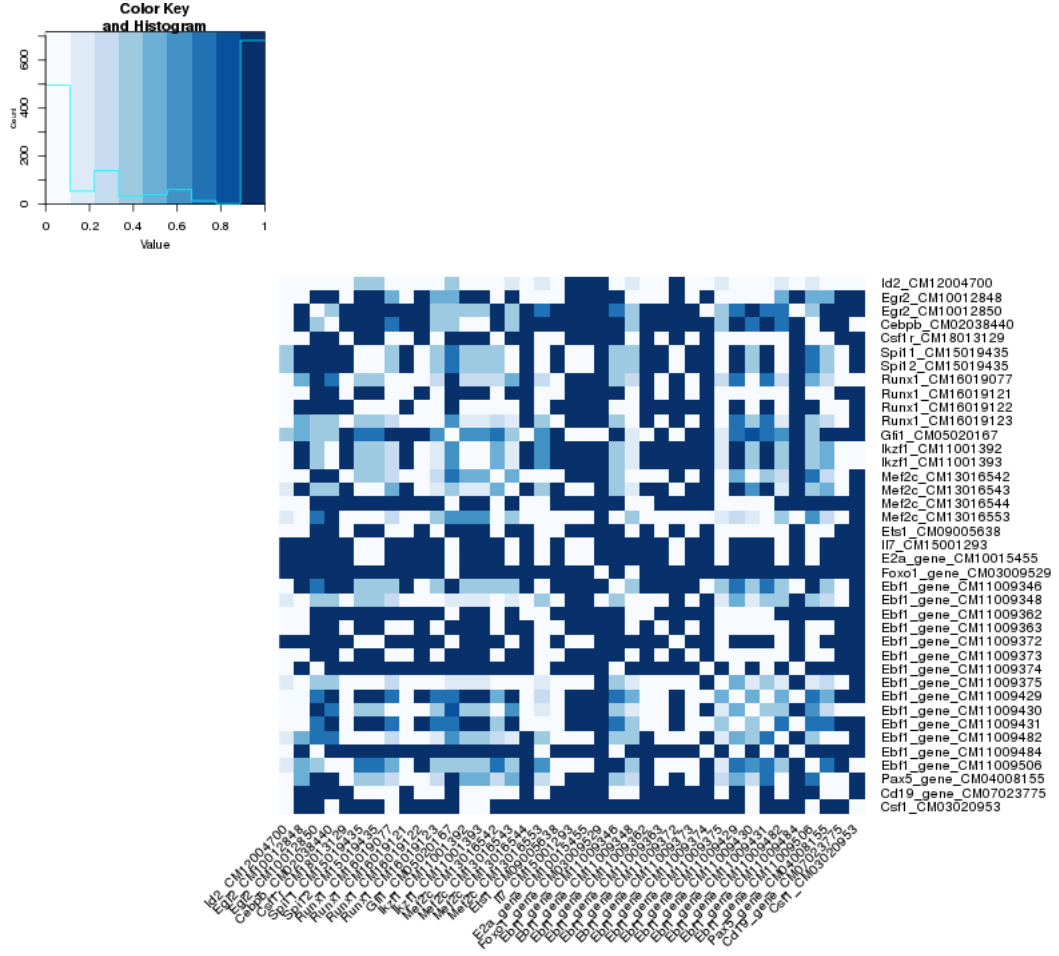


Figure 3: Overlap matrix between all the CRMs (for TFs involved in the original model). Original matrix M is such as $M_{i,j} = 1 - \frac{|CRM_i \cap CRM_j|}{\min(|CRM_i|, |CRM_j|)}$ (associated function: *overlapMatrix*).

References

- [1] Samuel Collombet, Chris van Oevelen, Jose Luis Sardina Ortega, Wasim Abou-Jaoudé, Bruno Di Stefano, Morgane Thomas-Chollier, Thomas Graf, and Denis Thieffry. Logical modeling of lymphoid and myeloid cell specification and transdifferentiation. *Proceedings of the National Academy of Sciences*, page 201610622, 2017.
- [2] Donna Karolchik, Robert Baertsch, Mark Diekhans, Terrence S Furey, Angie Hinrichs, YT Lu, Krishna M Roskin, Matthias Schwartz, Charles W Sugnet, Daryl J Thomas, et al. The ucsc genome browser database. *Nucleic acids research*, 31(1):51–54, 2003.
- [3] Alexei A Sharov, Dawood B Dudekula, and Minoru SH Ko. Cisview: a browser and database of cis-regulatory modules predicted in the mouse genome. *DNA research*, 13(3):123–134, 2006.

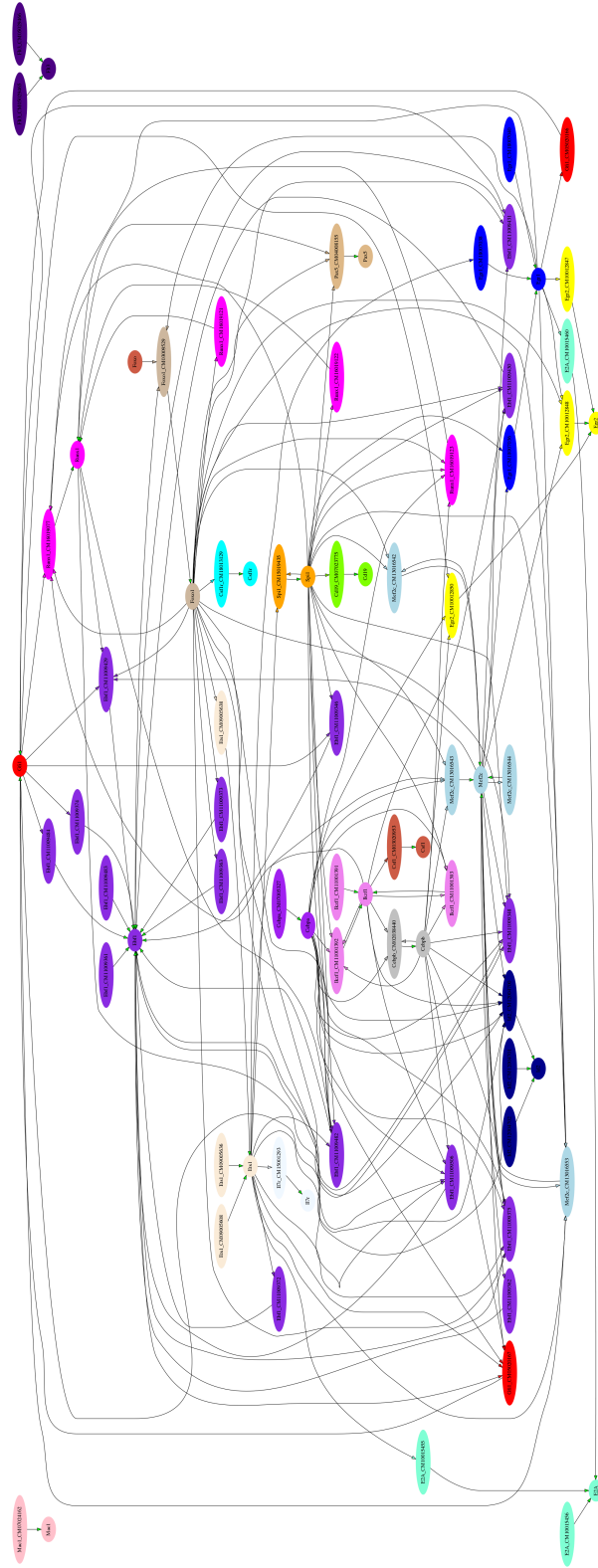


Figure 4: Resulting cis-regulatory network (with only "TF to CRM" and "CRM to gene" interactions).