

Genome-wide in Silico Identification of New Conserved and Functional Retinoic Acid Receptor Response Elements (Direct Repeats Separated by 5 bp)

Sébastien Lalevée, Yannick N. Anno, Amandine Chatagnon, Eric Samarut, Olivier Poch, Vincent Laudet, Gérard Benoît, Odile Lecompte, Cécile

Rochette-Egly

► To cite this version:

Sébastien Lalevée, Yannick N. Anno, Amandine Chatagnon, Eric Samarut, Olivier Poch, et al.. Genome-wide in Silico Identification of New Conserved and Functional Retinoic Acid Receptor Response Elements (Direct Repeats Separated by 5 bp). Journal of Biological Chemistry, 2011, 286 (38), pp.33322 - 33334. 10.1074/jbc.M111.263681 . hal-02647235

HAL Id: hal-02647235 https://hal.inrae.fr/hal-02647235v1

Submitted on 29 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Genome-wide *in Silico* Identification of New Conserved and Functional Retinoic Acid Receptor Response Elements (Direct Repeats Separated by 5 bp)*^S

Received for publication, May 22, 2011, and in revised form, July 28, 2011 Published, JBC Papers in Press, July 29, 2011, DOI 10.1074/jbc.M111.263681

Sébastien Lalevée^{‡¶1,2}, Yannick N. Anno^{§2}, Amandine Chatagnon^{¶||}, Eric Samarut[‡], Olivier Poch[§], Vincent Laudet^{**}, Gerard Benoit^{¶||}, Odile Lecompte[§], and Cécile Rochette-Egly^{‡3}

From the [‡]Department of Functional Genomics and Cancer and [§]Department of Integrated Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire, INSERM U596 and CNRS UMR7104, Université de Strasbourg, F_67404 Illkirch Cedex, [¶]CNRS UMR5534, F-69622 Villeurbanne, [∥]Université Lyon 1, UMR5534, F-69622 Villeurbanne, and **Institut de Génomique Fonctionelle de Lyon, UMR 5242, Institut National de la Recherche Agronomique, Université de Lyon, Ecole Normale Supérieure de Lyon, 69364 Lyon Cedex 07, France

Background: Retinoic acid (RA) receptors regulate gene expression through binding-specific response elements (RAREs). **Results:** A collection of new DR5 RAREs located ± 10 kb from TSSs and conserved among 6 vertebrates species or more has been amassed. **Conclusion:** We provide a wider knowledge base for analyzing RA target genes.

Significance: The RA response of the conserved target genes differs between species and tissues.

The nuclear retinoic acid receptors interact with specific retinoic acid (RA) response elements (RAREs) located in the promoters of target genes to orchestrate transcriptional networks involved in cell growth and differentiation. Here we describe a genome-wide in silico analysis of consensus DR5 RAREs based on the recurrent RGKTSA motifs. More than 15,000 DR5 RAREs were identified and analyzed for their localization and conservation in vertebrates. We selected 138 elements located ±10 kb from transcription start sites and gene ends and conserved across more than 6 species. We also validated the functionality of these RAREs by analyzing their ability to bind retinoic acid receptors (ChIP sequencing experiments) as well as the RA regulation of the corresponding genes (RNA sequencing and quantitative real time PCR experiments). Such a strategy provided a global set of high confidence RAREs expanding the known experimentally validated RAREs repertoire associated to a series of new genes involved in cell signaling, development, and tumor suppression. Finally, the present work provides a valuable knowledge base for the analysis of a wider range of RAtarget genes in different species.

Retinoic acid $(RA)^4$ is an active derivative of vitamin A that influences a range of essential biological processes such as

- ^S The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables S1–S5.
- ¹ Supported by the Association pour la Recherche sur le Cancer (ARC). Present address: Laboratory for Prenatal Medicine, Dept. of Biomedicine, University Hospital Basel, 4031 Basel, Switzerland.
- 2 Supported by the Ministère de l'Enseignement Supérieur et de la Recherche.
- ³ To whom correspondence should be addressed: 1 rue Laurent Fries, BP 10142, 67404 Illkirch Cedex, France. Tel.: 33-3-88-65-34-59; Fax: 33-3-88-65-32-01; E-mail: cegly@igbmc.fr.
- ⁴ The abbreviations used are: RA, retinoic acid; RAR, RA receptor; RXR, retinoid X receptor; DR, direct repeat; RARE, RA response element; TSS, transcription start site; qPCR, quantitative PCR; seq, sequencing.

development and homeostasis (1-4). RA exerts its action through nuclear RA receptors (RARs), which are typical liganddependent regulators of transcription with a central DNA binding domain linked to a ligand binding domain (for review, see Refs. 5 and 6). In response to RA signaling, RARs heterodimerize with retinoid X receptors (RXRs) and occupy characteristic RA response elements (RAREs) located in the promoter of target genes involved in cell proliferation and differentiation. RXR/RAR heterodimer occupancy at cognate response elements is commonly a determinant of transcriptional responsiveness. Within a given cell type, binding of RXR/ RAR heterodimers to RAREs can either up- or down-regulate transcription in a gene-specific manner. RAREs are composed of two direct repeats of a core hexameric motif (A/G)G(G/T)TCA. The classical RARE is a 5-bp-spaced direct repeat (referred to as a DR5), but RXR/RAR heterodimers can also bind to direct repeats separated by 2 bp (DR2) or 1 bp (DR1) (6, 7).

The development of high throughput technologies such as DNA microarrays revealed that within a given cell type or tissue, the RA response is composed of a huge and complex network of responsive genes (8-10). However, such techniques could not discriminate between direct primary and secondary target genes (which are modulated by the product of a primary target gene rather than by RXR/RAR heterodimers), and only a few of the RA target genes contained identified RAREs. More recently, chromatin immunoprecipitation coupled with array hybridization (ChIP-chip) allowed the identification of new RAR binding loci (11, 12). However, whether such loci bind RARs directly or indirectly through other bound factors could not be easily discriminated. Moreover, the identified loci do not correspond to the full repertoire, as the arrays do not represent all possible regions in a genome. The nascent genome-wide ChIP-seq (chromatin immunoprecipitation coupled with deep sequencing) technology should expand the repertoire of potential high affinity response elements (13, 14). Nevertheless, although powerful, such ChIP-based approaches are highly cell context-specific.

^{*} This work was supported by funds from CNRS, INSERM, the Association pour la Recherche sur le Cancer (ARC 3169), the Agence Nationale pour la Recherche (ANR-05-BLAN-0390-02 and ANR-09-BLAN-0127-01), the Fondation pour la Recherche Médicale (DEQ20090515423), and the Institut National du Cancer (INCa-PLO7-96099 and PL09-194).

Now, with the availability of an increasing number of genome sequences, *in silico* analysis of RAREs can be also performed. The advantage of computational techniques is that it overcomes the chromatin structure and, thus, the cellular context and provides a direct glance on the whole repertoire of possible RAREs.

Here we conducted a genome-wide *in silico* study of RA response elements. Although RXR-RAR heterodimers can bind to DR5, DR2, or DR1 response elements, the significance and the specificity of the DR2 and DR1 is still unclear. Therefore, we focused on DR5 RAREs. Computational techniques were developed for the genome-wide identification of DR5 RAREs and for the characterization of their genomic and phylogenetic context. In this way we amassed a collection of DR5 RAREs that is conserved across vertebrate species and that was validated for its occupancy and functionally analyzed for the RA-responsiveness of the associated genes. Such a strategy allowed us to characterize a new set of high confidence conserved DR5 RAREs associated to a series of new potential RA-target genes, thus providing a wider knowledge base for the analysis of the RA response in different species.

EXPERIMENTAL PROCEDURES

Bioinformatics—In silico analyses were performed using the Genomic Context data base (GeCo).⁵ This data warehouse, which was already exploited in a genome-wide study of the Staf transcription factor binding sites (15), aggregates genomic, phylogenetic, and epigenomic data from different sources, allowing the high-throughput contextual characterization of a given set of genetic elements. The underlying data base of annotated genes was built by computing refGene (proteins), rnaGene (snRNA, snoRNA, tRNA, rRNA, scaRNA), kgXref tables from the University of Santa Cruz California, mirna, mirna_literature_references, mirna_mature, and literature_references tables from the Sanger Institute and piRNA file from the piRNA Database. The data base also includes sequence conservation data extracted from the University of Santa Cruz California blastZ alignments. The data base is implemented in high speed DB2 architecture called Biological Integration and Retrieval of Data (BIRD), which can quickly address the whole set of sequences, genomic features, and alignments (15). RGK-TSA DR5 motifs were searched in the human (NCBI build 36, hg18)- and mouse (NCBI build 37, mm9)-masked genomes (RepeatMasker) using an in-house tool dedicated to the automatic search of short motifs and implemented in the GeCo system. The obtained motifs were subsequently characterized. For each motif, we retrieved the nearest gene and its localization as well as, if applicable, the position of the motif regarding the gene elements, exon, intron, transcription start site (TSS), and gene end (end of the last exon). Motif conservation was then analyzed on the basis of University of Santa Cruz California blastZ alignments between the human or mouse and 13 other vertebrate genomic sequences. The considered species were selected for the confidence of their sequencing, the quality of their annotation, and for their repartition through the vertebrate phylogenetic tree: zebrafish (danRer5), fugu (fr2), xeno-



pus (xenTro2), lizard (anoCar1), chicken (galGal3), platypus (ornAna1), opossum (monDom4), dog (canFam2), horse (equCab1), cow (bosTau4), rat (rn4), rhesus (rheMac2), and chimpanzee (panTro2). We considered a motif as conserved in a given species if the region encompassing the motif in human or mouse is aligned with a genomic region of the species also containing a RGKTSA DR5 motif.

Cell Culture, RNA Extraction, and qRT-PCR—F9 and P19 mouse embryocarcinoma cells, human MCF7 cells, and zebrafish PAC2 cells were cultured according to standard conditions as previously described (16–19). RNAs were extracted and subjected to qRT-PCR as previously described (20). Transcripts were normalized according to the ribosomal protein gene *RPLP0*. All mouse primers are listed in supplemental Table S1. The others are available upon request.

RNA Sequencing-After isolation of total RNA, a library of template molecules suitable for high throughput DNA sequencing was created according to the instructions of Illumina. Briefly, the poly(A)-containing mRNAs were isolated from total RNA (4 mg) by two runs of purification on Sera-Mag Oligo-dT Beads (Thermoscientific) and fragmented using divalent cations and heat-catalyzed hydrolysis. Fragmented mRNAs were used as a template to synthesize single-stranded cDNA with Superscript II reverse transcriptase and random primers. After second-strand synthesis, the cDNAs went through end-repair and ligation reactions using paired-end adapter oligos from Illumina and were electrophoresed on an agarose gel. A slice containing fragments in the 300-bp range was excised, and after elution and purification, the library was amplified with 15 cycles of PCR with Illumina sequencing primers and purified using Agencourt AMPure XP beads from Beckman.

The library was then used to build clusters on the Illumina flow cell according to protocol. Image analysis and base calling was performed using the Illumina pipeline. Reads were then mapped onto the mm9/NCBI37 assembly of the mouse genome using Tophat (21). Quantification of gene expression was done using Cufflinks (22) and annotations from Ensembl release 57. For each transcript the number of FPKM (fragments/kb of transcript/million fragments mapped) was converted into raw read counts, which were added for each gene locus by using an R script that we implemented. Then data normalization and identification of significantly differentially expressed genes were performed with the method proposed by Anders and Huber (23) and implemented into the DESeq Bioconductor package. The final *p* values were adjusted for multiple testing according to the method proposed by Benjamini and Hochberg (24), and a cutoff p value of 0.05 was applied for finding significant responsive genes.

RESULTS

Bioinformatic Genome-wide Research of DR5 RAREs Corresponding to the RGKTSA Motif—Only a few RAREs have been identified to date and associated to RA-target genes. Most of them are represented by two direct repeats of the hexameric motif (A/G)G(G/T)TCA, separated by five nucleotides (DR5) (25, 26). Such DR5s have been found in the promoters of human and mouse genes involved in RA metabolism (*Cyp26A1*) (27), in

⁵ Y. N. Anno, O. Poch, and O. Lecompte, manuscript in preparation.

RA signaling ($RAR\alpha 2$, $RAR\beta 2$, $RAR\gamma 2$) (28–30), or in development (*Hoxa1*, *Hoxa4*, *Hoxb1*) (31–33). Alignment of these RAREs (Fig. 1) clearly delineates a recurrent motif RGKTSA (coding is according to the IUPAC convention: R = AG; K = GT; S = CG), which differs from the classical consensus motif RGKTCA at position 5, with a G instead of a C (in $RAR\gamma 2$ and *Hoxa4*). Therefore, with the aim of identifying novel RA-driven primary target genes, we screened the masked human and mouse genomes for DR5 corresponding to two direct repeats of the RGKTSA motif at the genome-wide scale (see "Experimental Procedures"). Such *in silico* screens have the potential of identifying target genes independently of their tissue of expression. We identified 15,925 DR5s corresponding to two direct repeats of the RGKTSA motif in the mouse genome and 14,571 in the human genome (supplemental Tables S2 and S3).

Conservation of the DR5 RAREs during Evolution—A way to assess the potential relevance of response elements is to determine whether they are conserved between species (phylogenetic footprinting). Indeed, highly *in vivo* relevant DR5 RAREs are expected to be conserved and, thus, to be under an ancient strong selective constraint. Therefore, to delineate functional RAREs, we analyzed the conservation of the human and mouse RAREs across 13 additional vertebrate organisms (see "Experi-



DR5 RGKTSÄNNNNRGKTSÄ

FIGURE 1. Alignments of known DR5 RARE motifs in the promoters of the Cyp26A1, RAR α 2, RAR β 2, RAR γ 2, Hoxa1, Hoxa4, and Hoxb1 genes and definition of a RGKTSA motif.

mental Procedures") by using the BlastZ alignment of the University of Santa Cruz California genome browser. Due to the shortness and the divergence of the RGKTSA sequence, the criterion of conservation was deduced from the presence/absence of the complete DR5 motif RGKTSANNNNNRGKTSA in all considered genomes. We considered that a motif was conserved in a given species if the region encompassing the motif in human or mouse is aligned with a genomic region of the species also containing a RGKTSA DR5 motif (see "Experimental Procedures").

In a phylogeny of vertebrates, we visualized the number of human or mouse RAREs that are conserved in each studied species (Fig. 2). We also calculated for each relevant clade of vertebrates the number of RAREs that are conserved in all the members of these clades. Although these data can be influenced by the coverage of the studied genome (34), this analysis raised three interesting conclusions. (i) Overall, human RAREs are less conserved in rodents than in other mammals. As an example, about 900 human RAREs are conserved in the mouse genome, whereas more than 1500 are conserved in the cow genome. This is in accordance with the known increased evolutionary rates in rodents (35) but questions the use of mouse as a unique in vivo experimental system for studying RA signaling in mammals. (ii) There is a striking difference between the number of RAREs conserved in placental mammals (Eutherians) and in all mammals or in eutherians + marsupials. Indeed around 309 human RAREs (or 319 mouse RAREs) are conserved in placental mammals, whereas only half (162 human RAREs, 170 mouse RAREs) is conserved in eutherians + marsupials. The decrease is even higher if we consider the number of RAREs conserved in all mammals (101 from human and 107 from mouse genomes). This suggests that a specific elaboration of the RA regulatory network occurred in eutherians and highlights the importance of studying the corresponding RA target genes. (iii) Only six RAREs are conserved in all jawed verte-



FIGURE 2. **Phylogenetic tree of jawed vertebrates showing the phylogenetic conservation of the DR5 RAREs.** On the *right*, the number of human RAREs (*blue*) or mouse RAREs (*orange*) conserved in each species is indicated. For example, in chimpanzee, there are 11,992 RAREs conserved from the 14,571 found in human. In contrast there are 919 RAREs conserved from the 15,925 found in the mouse genome. At each relevant node of the tree, the number of RAREs conserved in the species of the relevant node is indicated in *red*. The *blue numbers* represent the human RAREs, and the *orange numbers* represent the mouse RAREs. For example, in the rodent primate clade we found 562 human RAREs and 597 mouse RAREs conserved in the 5 relevant species (mouse, rat, rhesus, chimpanzee, and human). *NA*, not applicable.



FIGURE 3. Genome-wide distribution of the identified human (A) and mouse (B) DR5 RAREs. The distance between the RARE and the corresponding gene was calculated by identifying its proximity to both boundaries of the genes, the TSS, and the end (end of the last exon). Following this rule for a RARE located upstream of the TSS, the distance was calculated from the TSS and was negative. For a RARE located downstream of a gene, the distance was calculated from the end of the gene, both distances to the TSS and to the end of the gene were calculated, and the ending distance in absolute value, called dTSS*, was considered (*cf.* Fig. 5). The genome-wide mapping of the RARE *versus* a canonical gene was calculated by cumulating the number of hits present in a 1-kb sliding window from each side from the TSS or from the end of the gene and 100 kb inside of the gene.

brates (gnathostomes). Three of these RAREs are associated to developmental genes, *Dach1* (Daschung homolog 1) (36), *Meis1* (Meis homeobox 1) (37), and *TSHZ3* (Teashirt 3) (38). The three others are associated to the *Gria2* (glutamate receptor 2) (39), *Lphn2* (latrophilin 2) (40), and *Paqr3* (an adiponectin receptor) (41) genes. It is interesting to note that, except Meis1 (8, 10, 37), these gene are not known RAR target genes. Nevertheless, they are likely to be RA-regulated in virtually any vertebrate species and thus might be considered as new interesting models.

Genome-wide Analysis of the Location of the Identified DR5 RAREs—The identified mouse and human RAREs were also annotated by analyzing genome-wide their locations using the GeCo system (see "Experimental Procedures"), which allows users to retrieve the genes in the neighborhood of factor binding sites with respect to annotated Refseq genes.⁵ Then, in both the human and mouse genomes, the RAREs were localized relative to the nearest matched gene boundary: upstream and downstream distance from the TSS and from the end of genes. As shown in Fig. 3, *A* and *B*, the regions flanking TSSs and the ends of genes depict the highest concentration of RAREs compared with the further regions (\pm 500 kb). This suggests that the RAREs located in the vicinity of TSSs and gene ends would be more relevant than the others, as described for most nuclear receptors and transcription factors (14, 42–46). Therefore, we selected the RAREs located between -10 and +10 kb, *i.e.* the RAREs \pm 10 kb from the TSSs and \pm 10 kb from gene ends. According to this criterion, 3862 RAREs were selected in the mouse genome and 3429 in the human one (supplemental Tables S2 and S3).

Selection of a List of RAREs Located ± 10 kb from Gene Limits and Conserved in Six Organisms or More—Considering the low number (6) of highly conserved RAREs and the overall reparti-

SEPTEMBER 23, 2011 · VOLUME 286 · NUMBER 38





FIGURE 4. Selection of 138 RAREs located \pm 10 kb from TSSs and gene ends and conserved in more than 6 organisms. *A* and *B*, shown is conservation of the human and mouse RAREs among the 15 organisms tested. *C* and *D*, for both the human and mouse genomes, the RAREs conserved in more than 6 organisms were crossed with those located \pm 10 kb from gene boundaries. Crossing the resulting mouse and human RAREs led to a list of 138 RAREs with highly confident conservation and located \pm 10 kb. *E*, the differences between mouse and human dTSS* were calculated and plotted into cumulative hits. *F*, conservation of the RARE positions (intron, exon, and promoter) between mouse and human is shown.

tion of the 15 organisms among the vertebrate tree, we arbitrary selected a criterion of conservation in 6 organisms. Only 7% of the human RAREs (1049 sites) (Fig. 4A) and 5% of the mouse RAREs (766 sites) (Fig. 4B) were found to be conserved in 6 organisms and more.

Then these human and mouse RAREs conserved across six and more organisms were further analyzed for their localization relative to the matched gene annotations. Among these RAREs, 238 human RAREs and 181 mouse RAREs were found to be located at the proximity of genes in the ± 10 kb regions that we defined above (Fig. 4, *C* and *D*, and supplemental Tables 2 and 3). By using these two criteria of restriction, we obtained a list of 138 RAREs that are common to both mouse and human organisms and that are reliable in terms of genome annotation, name of the corresponding genes, localization, and conservation in more than 6 organisms (Fig. 5).

The orientation and localization of each conserved RARE listed in Fig. 5 were compared. Most interestingly, 100% of these RAREs showed the same orientation in the human and mouse genomes, and \sim 70% showed less than a 500-bp difference in their distances to the nearest associated genes (Fig. 4*E*), suggesting that these RAREs are good candidates for being func-



			MOUSE									Mouse-Human
Official gene name	Position	BS strand	Gene strand	Phylogenetic Conservation	dTSS*	RARE Sequence	Position	BS strand	Gene strand	Phylogenetic Conservation	dTSS*	dTSS* difference
2310007D09Rik	Exon	+	+	8	-3436	GGTTCACGAAAAGTTCA	Exon	+	+	8	-5116	1680
4121402D02Rik	Intron		+	8	-5908	AGGTGAGTAAAAGTTGA	Intron		+	8	-5567	-341
4632412N22Rik			+	6	1259	GGGTCACAATCAGGTGA			+	6	1282	-23
4930452B06Rik	Exon	+		7	-459	AGTTCAGACAAGGGTCA	Exon	+	-	7	-495	36
4930562D19Rik	Exon			8	8578	GGGTGACCCCGAGTTCA	Exon	+	+	8	9552	-974
6430527G18Rik			100 A	8	-3426	GGGTGACGCGAGGTTCA				8	-4278	852
Abhd2	Exon		+	7	-1367	AGGTCAGATCCAGGTCA	Exon	-	+	7	-6849	5482
Aco2	Intron	+	+	9	404	AGGTGAGGACAAGGTGA	Intron	+	+	9	383	21
Actn4	Exon			6	-1391	GGGTGAGGCCGAGTTCA	Exon	+	+	6	-1535	144
Agap1	Intron	+	+	9	-9441	GGGTCACTCTGGGGTCA	Intron	+	+	8	-7485	-1956
Aipl1	Intron			6	-2338	GGGTCAAGGTCAGGTCA	Intron	-	-	6	-3811	1473
Alk	Intron			8	5054	AGGTGAGATGTGGGTCA	Intron		-	8	5420	-366
Ankrd50	Exon	+	100 A	9	-1989	AGTTGAAGCTAGGGTCA	Exon	+	-	9	-4700	2711
Ar	Exon		+	8	-1258	AGTTCATTCGAAGTTCA	Exon	-	+	8	-1413	155
Arpp21	Intron	-		7	-6279	GGTTCATTCATAGGTCA	Intron	+	+	7	-5225	-1054
Atxn2		+	+	6	-561	GGTTCAAATGAAGGTCA		-	-	6	-495	-66
BC052040	Exon		+	12	-1662	AGTTCAGCTGAAGTTCA	Exon		+	12	-1701	39
Bhlhbe40			+	12	1556	AGGTCAGCGCTGGGTGA		-	+	12	1775	-219
Bmf	Exon	+		8	-864	AGGTCACTGGGAGGTGA	Exon	+	-	8	-876	12
Bmp7	Intron	+		12	4301	AGTTCAAAGCTGGGTCA	Intron	+	-	12	5173	-872
Brs3	Exon		+	6	115	AGTTCAAATGCAGTTGA	Exon		+	6	95	20
Cacna1g	Intron	+	-	9	6105	GGGTCAGCCAGGGGTCA	Intron	-	+	9	6019	86
Camk2b	Exon	+		6	-3030	AGGTCAGCCCTGGGTCA	Exon	+	-	6	-3536	506
Ccdc132	Exon		+	8	-8558	AGTTCACTGGCGGTTGA	Exon	-	+	8	-9145	587
Ccno			+	7	-5719	AGTTCACTAAAAGTTCA		+	-	6	-8211	2492
Cdcp2		+	+	7	663	GGGTGACACACAGTTCA		-	-	7	1092	-429
Cited2		+	+	8	574	GGTTGACTCAAAGGTCA		-	-	8	632	-58
Clpb	Intron	100 B	+	7	5382	AGGTGATTCCAGGGTCA	Intron	+	-	7	5499	-117
Col24a1	Exon	-	+	7	292	AGTTCACTTGCGGTTCA		+	-	7	-296	588
Crygn		+		6	3244	GGGTGAGGGGGGAGGTCA		+	-	6	1061	2183
Ctsk	Exon		+	7	2154	GGTTCATAGCCAGTTCA	Exon	+	-	7	2220	-66
Cugbp1	Exon	100 B	+	11	-379	AGGTCAACACAAGGTCA	Exon	+	-	11	-375	-4
Cugbp2	Exon	+		14	-5337	AGTTCAGCAAAAGGTGA	Exon	-	+	14	-5629	292
Cyp26a1		-	+	14	-31	AGTTCACCCAAAGTTCA	Exon	-	+	14	329	-360
Cyp26a1		+	+	8	-1898	GGGTCACAGGCGGGTCA		+	+	8	-1557	-341
Diap3	Exon	-		6	-229	AGTTGACAGAAGGGTCA	Exon		-	6	-1227	998
E130309F12Rik	Intron	+	+	9	745	GGGTGAGGCTCGGGTCA	Intron	+	+	9	958	-213
Ebf3		-		11	-7615	AGGTGAAACCGAGTTCA				11	-8385	770
Elf4	Exon	+		7	-7145	GGGTCAGTGACAGGTGA	Exon	+	-	7	-6455	-690
Enpp2		+		7	2163	AGTTCAACCTGAGGTGA		+	-	7	2866	-703
Ephb3	Intron	-	+	8	-6596	GGGTCACCTGGAGTTCA	Intron	-	+	8	-7178	582
Esrrg	Exon		+	10	-3113	GGTTCACAATAAGTTCA	Exon	+	-	10	-3061	-52
Ext2	Intron			9	-7409	AGTTCAAGCCCAGGTCA	Intron	+	+	9	-8911	1502
Fbxo30	Exon	+	+	6	-5674	AGGTCATTTGGAGTTCA	Exon		-	6	-6502	828
Fit4	Intron	+	+	8	3918	AGGTCAGCAGGGGTTCA	Intron	-	-	8	4073	-155
Foxa2		+		9	8392	AGGTCAGGGGGAGGTCA		+		9	8252	140
Frap1	Exon		+	9	-7491	AGTTCAGCAAGGGGTCA	Exon	+	-	9	-7802	311
Gabarapl2	Intron	-	+	7	522	AGGTGACAGCGAGGTGA	Intron	-	+	7	579	-57
Gpr149	Intron	+	100 B	6	-8572	AGGTGAATCCAGGGTGA	Intron	+	-	6	-8030	-542
Gría2	Intron	-	1.1.1	15	-6381	AGGTGATTCCAAGGTCA	Intron	+	+	15	-4962	-1419
Gzmk		+	100 A.	8	-2144	GGGTGAGATTTAGTTCA		-	+	7	-2337	193
Hey1	Exon	+	100 A	7	-49	AGGTCAAACCCAGTTCA	Exon	+	-	7	-49	0
Hic1	Intron		1.1.1	9	1444	GGTTCACGGCGGGGTCA	Intron	+	+	9	322	1122
Hin1r	Exon	+	+	7	-2808	AGGTGAACGAGAGGTGA	Exon	+	+	8	-3142	334

FIGURE 5. List of the 138 conserved RAREs located ±10 kb from TSSs and genes ends. For the 138 conserved RAREs located ±10 kb, gene names and orthology were analyzed manually in both the human and mouse genomes. The sequence, localization, and dTSS* of each RARE are shown as well as the name of the associated gene. *BS*, binding sequence.

tional. Moreover, 43 RAREs located in exons, 49 in introns, and 30 in promoters correlated well between the two genomes (Fig. 4*F*). The other RAREs, although associated with a same gene in both genomes, depicted different localizations, most probably due to differences in genes annotations between the two genomes. Note that three RAREs associated with Hox genes differed between both genomes, most probably due to the complex organization and evolution of the Hox clusters.

RAR Binding to the Selected DR5 RAREs and Analysis of the Associated Genes—Then the key question to address was whether the DR5 RAREs that we selected *in silico* reflect biological significance *in vivo*; in other words, whether they are able to bind RAR/RXR heterodimers. To address this, we first crossed the list of 15,925 DR5 RAREs found in the mouse genome with the RAR and RXR binding sites mapped in ChIPseq experiments⁶ performed with a mouse embryocarcinoma cell line (F9 cell lines), which is well known to respond to RA (1). In these cells, 4% of these RAREs were occupied by RAR/RXR heterodimers in the absence of RA (Table 1). This percentage increased up to 9% after 48 h of RA treatment. In fact, taking into account that some sites become occupied whereas oth-



ers dissociate from RAR/RXR heterodimers in response to RA, 11% of the RAREs were found to be able to bind RAR/ RXR heterodimers (Table 1). As a control, a random list of 15,925 17-bp sequences extracted from the mouse-masked genome (supplemental Table S4) was crossed with the same binding sites (Table 1). Most interestingly, the percentage of occupied RAREs increased up to 42% when the same crossing was applied to the list of 181 conserved mouse RAREs and to our final in silico list of 138 RAREs (Table 1), thus validating our strategy.

Among these RAREs (Table 2), 39 were occupied in the absence of RA, among which 17 depicted an important increase in their occupancy in response to RA. In addition, 19 RAREs, although unoccupied in the absence of RA, became occupied after RA addition, raising to 58 the number of RAREs that can be occupied in F9 cells. Note that the increase in occupancy started rapidly (within 2 h) or later (24-48 h) depending on the RARE.

Some of these RAREs have been already reported to be direct RAR targets in EMSA, ChIP, or ChIP-chip experiments. It is the case for the canonical RAREs associated to the well known RA target genes involved in transcription regulation such as *RAR* β 2, *Hoxa1*, *Hoxb1*, and *Wt1* (19, 32, 47–50) or in RA metabolism (*Cyp26A1*, *Rbp1*). Most inter-

⁶ A. Chatagnon and G. Benoit, manuscript in preparation.

				9	48	AGGIGAICACAGGGIGA	Exon		+	9	-133	181
Hoxa1		-	-	12	1940	GGTTCACCGAAAGTTCA				12	1749	191
Hewe?	Introp			14	6472	ACCTCAACTTCACCTCA				14	2124	2248
CDXON	incron		-	14	-0472	AGGTGAACTTCAGGTCA		-		14	-5124	-3346
Hoxa3	Intron		-	14	6064	GGTTCAAGAAGAGTTCA	Intron			14	6071	-7
Hoxb1		+	+	13	5100	GGTTCATAGAGAGTTCA				14	7170	-2070
11	lution and			14	6604	COTTOALCAACACTTCA	later an			14	6561	100
HUXUS	incron	Ŧ	Ŧ	14	-0094	GGTTCAAGAAGAGTTCA	Incron	-		14	-0301	-135
Hoxc4		+	+	14	-5034	AGGTGAAATGCAGGTCA	Intron	+	+	14	-6699	1665
Hoxd3	Intron	+	+	13	-6731	GGTTCAAGCAGAGTTCA	Intron	+	+	12	-7185	454
Houdd	Introp			14	2770	AGGTGAAATGCAGGTCA				14	1240	4019
HOX04	incidit	т	т	14	2//0	AGGIGAAAIGCAGGICA		Ŧ	Ŧ	14	-1240	4018
Hoxd4	Intron	+	+	11	2696	AGTTGAGCGGGAGGTGA		+	+	11	-1322	4018
lyns1ahn			+	9	-187	AGGTCAGCGCAAGGTGA		+		9	-139	-48
to the	la transmissione de la constante de la constan			7	1001		Lotter and				1077	76
Jmjas	Intron	-	-	/	4801	GGG CALA CGGGG CA	Intron	+	+	/	48/7	-/6
Klk13	Intron	+	+	6	435	AGTTCACAGTGGGTTCA				6	-778	1213
Look 2	Introp			7	204	AGGTCAAGGTGAGGTCA	Introp			7	222	72
Lgaisz	incron			,	=354	AGGICAAGGIGAGGICA	IIILIOII			,	*322	-12
Lman2I	Intron	+	-	6	6032	AGGTCAAATCAGGTTGA	Intron	+	100 B	6	6093	-61
Lrrc29	Exon	+		6	-3339	GGTTGAAGCTGAGGTCA	Intron	+		6	-3496	157
Main3	Introp			14	7609	CETTCACTCAAACCTCA	Intern			14	7710	12
IVICISZ	incron	Ŧ		14	-7038	GGITCACTCAAAGGICA	Incron	Ŧ		14	-7710	12
Meis2	Intron	+	-	12	-6268	GGGTCATTCAGAGGTCA	Intron	+		12	-6210	-58
Meis2			-	8	-6129	AGGTCAAGAATAGTTCA				9	-4855	-1274
MIII	Introp			7	1910	AGTICAGTICAGGTICA	Introp	+	+	7	1972	62
IVIIIT	incron		-	,	1910	AGTIGAGTICAGGTICA	Incron	Ŧ	Ŧ	,	10/3	-03
Mmp24	Exon	+	+	6	-1581	AGGTGAACTAGAGGTGA	Exon	+	+	6	-1658	77
mmu-let-7c-2		+	+	8	-3954	AGGTGACTTCGGGGTGA		+	+	8	-5394	1440
manau mais 10a				12	2566	AGGTGAACCGCAGGTCA				14	2622	E7
mmu-mil-10a		т	т	15	-2300	AGGIGAACCGCAGGICA				14	-2025	37
Myf6	Intron	+		6	680	GGGTCACTGGGGGGTTCA	Intron	-	+	6	705	-25
Myo3b	Intron	-	+	6	-7559	GGTTCAACTAGAGTTCA	Intron		+	6	-6568	-991
NikI	Even			C C	200	AGGTGAAATTCCCTTC	Even			~	634	220
ткарі	EXON			0	-360	AGGTGAMATTGGGTTGA	EXON	+	+	0	-024	235
Onecut2	Intron	+	+	12	9911	AGTTCAGCTATGGTTCA	Intron	+	+	12	9964	-53
Osr1		+	+	8	-988	GGGTCAGCCGGAGGTCA				8	-1004	16
0512				7	415	ACCTCACCCCACCCCTCA				6	401	14
Отр		+	+	/	-415	AGGTLALGLLAGGGTLA				6	-401	-14
Parp8		-		8	9175	AGTTCATTTACAGGTCA		+	+	7	8462	713
Pchn2	Intron		+	9	-6347	GGTTGATGCTGAGGTGA	Intron		+	10	-6242	-105
Del 1	Interor			6	2200	ACCTCACCTCCACCTC	Interest				2/20	140
РСБр4	Intron	+	+	8	3290	AGGIGAGCIGGAGGICA	Intron			8	3438	-148
Pld2	Intron	-	+	7	-7034	AGGTGACAGCAAGGTCA	Intron	-	+	7	-9119	2085
Polb	Exon	-	-	7	-8806	GGGTGAGACAAAGTTCA	Exon	+	+	7	-9115	309
Dev 2f1				C.	2257	ACCTCATCCTCCCTTCA				£.	4020	(93
FOUSIT			Ŧ	0	3337	AGGIGAIGCIGGGIICA		Ŧ		0	4035	-082
Prnpip1		+	+	10	-3868	AGTTCAGATTCAGGTCA			100 B	10	-3986	118
Prss27		+	+	8	-148	GGTTCAGGGCTAGGTCA	Exon	-	-	8	195	-343
Ptch1	Introp	+		11	6913	GGGTCACAGGCAGGTCA	Introp	+		9	7203	-290
Probable	incron		-	11	4005	ACCTCL CTCL CCCTCL	finition f				7205	-2.50
Ptchal	Exon	-	-	11	-4895	AGGICAGIGAAGGGICA	Exon	+	+	11	-3721	-11/4
Ptprj	Exon	-	-	7	-917	AGGTCACACGAAGGTCA	Exon	+	+	8	-963	46
Ptpri	Evon	+		6	-7803	AGTIGATGAGCAGGTCA	Evon		+	6	-7307	-496
1 (p)	LAUII				7005	AdmontoAtocAdottcA	LAUT				7307	450
QK		+		14	608	AGGICAICAIIGGGIGA	Exon		+	14	-3124	3732
Rab11fip2	Exon	-	-	8	5995	GGGTGAGATAAAGGTCA	Exon			8	6099	-104
Pab20b	Evon			7	/10	GGGTCAAGAGAGGTTCA	Evon			7	/02	87
100550	LAUIT			,	410		LADIT	-		,	452	02
карпі	Intron	-	-	/	-8272	GGGTCAAGTAAGGTCA	Intron			/	-1126	-7146
Rarb			-	12	-303	GGTTCACCGAAAGTTCA		+	+	13	-52	-251
Barb	Exon	+		9	-4293	AGGTGAACACAAGGTCA	Exon		+	9	-4316	23
Dhhu7	Even			0	1390	CECTEATAACCAECTEA	Even			0	1202	3
Kuup/	EX()			0	1200	GOOICATAACCAGOICA	EX(1)	Ŧ	-	9	1/03	-5
	EROTI	-					EXOT				1205	
Rbm35b	Exon	+	-	8	-943	AGTTCAGGTAGAGTTGA	Exon	+	-	8	-2314	1371
Rbm35b Rbn1	Exon	+	- +	8	-943 7034	AGTTCAGGTAGAGTTGA	Exon	+		8	-2314	1371 1036
Rbm35b Rbp1	Exon Intron	+	-+	8	-943 7034	AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTCA	Exon Intron	+ +		8	-2314 5998	1371 1036
Rbm35b Rbp1 Rnf10	Exon Intron Exon	- + - +	- + -	8	-943 7034 -2326	AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTCA GGGTGACAGAGGGGGTCA	Exon Intron Exon	+ + -	- - +	8 8 8	-2314 5998 -6394	1371 1036 4068
Rbm35b Rbp1 Rnf10 Rnf214	Exon Intron Exon Exon	- + - +	+	8 8 6	-943 7034 -2326 6752	AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTCA GGGTGACAGAGGGGTCA GGTTCATCAGCAGGTGA	Exon Intron Exon Exon	+ + + + + + + + + + + + + + + + + + + +	- - + +	8 8 6	-2314 5998 -6394 5981	1371 1036 4068 771
Rbm35b Rbp1 Rnf10 Rnf214 Ror1	Exon Intron Exon Exon	- + - + -	- + - -	8 8 8 6 12	-943 7034 -2326 6752 -6573	AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTCA GGGTGACAGAGGGGGTCA GGTTCATCAGCAGGGGGGA AGGTCAAGCAGAGGGTCA	Exon Intron Exon Exon	+ - + + + + + +	- - + +	8 8 6 12	-2314 5998 -6394 5981 -6197	1371 1036 4068 771 -376
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Pyyb	Exon Intron Exon Exon Intron	- + - + - +	- + - - +	8 8 6 12	-943 7034 -2326 6752 -6573 -1559	AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTCA GGGTGACAGAGGGGTCA GGTTCATCAGCAGGGTGA AGGTCAAGCAGAGGTCA	Exon Intron Exon Exon Intron	+ + + + + + +	- + +	8 8 6 12	-2314 5998 -6394 5981 -6197	1371 1036 4068 771 -376 325
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Rxrb	Exon Intron Exon Exon Intron Intron	- + - + - + -	- + - + +	8 8 6 12 8	-943 7034 -2326 6752 -6573 -1559	AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTCA GGGTGACAGAGGGGCCA GGTTCATCAGCAGGGGCA AGGTCAAGCAGAGGTCA AGGTCACTCAAAGGTCA	Exon Intron Exon Exon Intron Intron	+ + + + + +	- + + +	8 8 6 12 8	-2314 5998 -6394 5981 -6197 -1884	1371 1036 4068 771 -376 325
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Rxrb Sat1	Exon Intron Exon Exon Intron Intron Exon	- + - + - + - + -	- + - + + + + -	8 8 6 12 8 8	-943 7034 -2326 6752 -6573 -1559 -460	AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTCA GGGTGACAGAGGGGTCA GGTTCATCAGCAGGAGGTGA AGGTCACTCAGCAGGGGCA AGGTCACTCAAAGGTCA AGTTGATGGATGGTTCA	Exon Intron Exon Exon Intron Intron Exon	+ + + + + + + + + +	- + + + +	8 8 6 12 8 8	-2314 5998 -6394 5981 -6197 -1884 -471	1371 1036 4068 771 -376 325 11
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Rxrb Sat1 Sema3e	Exon Intron Exon Exon Intron Exon Exon	- + - + - + - - + - - + -	- + - + + + +	8 8 6 12 8 8 8 8 8	-943 7034 -2326 6752 -6573 -1559 -460 63	AGTTCAGGTAGAGTTGA GGGTGATCCTAAGTTCA GGGTGACAGAGGGGTCA GGTTCATCAGCAGGTGA AGGTCATCAGCAGGGTCA AGGTCATCGATGGATGGTCA AGTTCATGGATGGTTCA AGTTCAGGCAGGGTTGA	Exon Intron Exon Exon Intron Intron Exon	+ + + + + + +	- + + + - +	8 8 6 12 8 8 8 8	-2314 5998 -6394 5981 -6197 -1884 -471 -117	1371 1036 4068 771 -376 325 11 180
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Ror1 Sat1 Sem3e Set2	Exon Intron Exon Intron Intron Exon Exon Exon	- + - + - - + - - -	- + - + + + + +	8 8 6 12 8 8 8 8 8 8	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190	AGTTCAGGTAGAGTIGA GGGTCATCCTAAGTTCA GGGTGACCGAGGGGTCA GGTTCATCAGCAGGGGTCA AGGTCAGCCAGAGGTCA AGGTCAGCCAGAGGTCA AGTTCAGGCAGGGTTCA AGGTCAGCCAGGGTTCA	Exon Intron Exon Exon Intron Intron Exon	+ + + + + + + +	- + + + - +	8 8 6 12 8 8 8 8 8	-2314 5998 -6394 5981 -6197 -1884 -471 -117 -755	1371 1036 4068 771 -376 325 11 180 -435
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Rurb Sat1 Sem38 Sgk2 Shurb2	Exon Intron Exon Exon Intron Intron Exon Exon Exon	- + - + - - + - -	- + - + + + + + + +	8 8 6 12 8 8 8 8 8 8 8 7	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190 630	AGTTCAGGTAGAGTIGA GGGTCATCCTAAGTTCA GGGTGACCGAGGGGTCA GGTTCATCAGCAGGGGCA AGGTCAAGCGAGAGGTCA AGGTCAAGCGAGAGGTCA AGGTCAGCGAGGGTCA AGGTCAGCAGGGTGA GGGTGAACTCTGGGTCA	Exon Intron Exon Exon Intron Intron Exon Exon	+ + + + + + - +	- + + + - + -	8 8 6 12 8 8 8 8 8 8 7	-2314 5998 -6394 5981 -6197 -1884 -471 -117 -755 2522	1371 1036 4068 771 -376 325 11 180 -435
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Rxb Sat1 Sem3e Sgk2 Shank3	Exon Exon Exon Exon Intron Exon Exon Exon Exon	- + - - + - - - - - -	· + - + + + + + + +	8 8 6 12 8 8 8 8 8 8 8 7 7	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620	AGTICAGGTAGAGTIGA GGGTCATCCTAAGTICA GGGTCATCAGAGGGGTCA GGTTCATCAGCAGGGGTGA AGGTCAAGCAGAGGTCA AGGTCAACCACAAGGTCA AGTTCAGGCAGGGTTGA GGGTGAACCTCGGGTCA GGGTCAGCCCAGGTGA	Exon Intron Exon Exon Intron Exon Exon Exon	+ + + + + + + + + + - + - + +	- + + + + + + + + + + +	8 8 6 12 8 8 8 8 8 8 7 7	-2314 5998 -6394 5981 -6197 -1884 -471 -117 -755 -2523 -2523	1371 1036 4068 771 -376 325 11 11 180 435 1903
Rbm35b Rbp1 Rnf10 Rnf224 Ror1 Rnrb Sat1 Sema3e Sgk2 Shank3 Shf	Exon Intron Exon Intron Intron Exon Exon Exon Intron Intron	- + - + - - + - - - - -	- + - + + + + + + + + + +	8 8 6 12 8 8 8 8 7 11	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -6710	AGTTCAGGTGAAGTTGA GGGTGAACAGAGGGGTCA GGGTGAACAGAGGGGTCA GGTTCATCAGCAGGGTGA AGGTCAAGCAGAGGTCA AGGTCAGCAGAGGTCA AGTTCATGCAGGGTTGA GGGTGAACTCTGGGTCA GGGTGAACTCTGGGTCA GGGTCAACGTCAAGGTGA AGTTGAACATTAAGGTCA	Exon Intron Exon Exon Intron Exon Exon Intron Exon	+ + + + + + + + - +	- + + - + - + - + + - + -	8 8 6 12 8 8 8 8 8 7 11	-2314 5998 -6394 5981 -6197 -1884 -471 -117 -755 -2523 -6489	1371 1036 4068 771 -376 325 11 180 -435 1903 -221
Rbm35b Rbp1 Rnf10 Rnf1 Rnf1 Stat Sat1 Sema3e Sgk2 Shank3 Shf Sk225	Exon Intron Exon Intron Intron Exon Exon Exon Intron Intron Intron	- + - - - - - - - - - -	- + - + + + + + + + + +	8 8 6 12 8 8 8 8 8 7 7 11 9	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -63 -1190 -620 -6710 -7623	AGTICAGGTAGAGTIGA GGGTCATCLAAGTIGA GGGTCATCLAAGTIGA GGTTCATCAGCAGGGTCA GGTTCATCAGCAGGGTCA AGGTCAACGAGGGTCA AGTGCATGCAGAGGTCA AGTTCAGGCAGGGTGA GGGTGAACTCGGGTCA GGGTCAACCCCAGGTGA AGTTGACATTAAGGTCA GGGTGAACTCAGGTCA	Exon Intron Exon Intron Intron Exon Exon Intron Exon Intron	+ + + + + + + + - + + + + + + + + + + +	- + + - + - + + + + + + +	8 8 6 12 8 8 8 8 8 7 7 11 8	-2314 5998 -6394 5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207	1371 1036 4068 771 -376 325 11 180 -435 1903 -221 -416
Rbm35b Rbp1 Rnf10 Rnf224 Ror1 Rnrb Sat1 Sem32e Spak2 Shank3 Shf Sk22a5 Sk22a5	Exon Intron Exon Intron Intron Exon Exon Exon Intron Intron Intron	- - - - - - - - - - - - -	- + - + + + + + + + + -	8 8 6 12 8 8 8 8 7 7 11 9 8	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -6710 -7623 -7256	AGTTCAGGTGAAGTTGA GGGTGAACAGAGGGGTCA GGTTGATCAGCAGGGGTCA GGTTGATCAGCAGGGGTGA AGGTCAGCAGAGGTCA AGGTCAGCAGAGGGTCA AGTTGATGGATGAGGGTGA GGGTGAACTCTGGGTCA GGGTGAACTCTAGGTCA GGGTGAAGCTCAGGGTCA AGTTGAACACTAAGGTCA	Exon Intron Exon Intron Intron Exon Exon Intron Exon Intron Intron	+ + + + + + + + + + + + + + + + + + + +	- + + - + + + + + + + + +	8 8 6 12 8 8 8 8 7 111 8 7	-2314 5998 -6394 5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343	1371 1036 4068 771 -376 325 11 180 -435 1903 -221 -413
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Stat1 Sem3e Sgk2 Shank3 Shf Sk22a5 Sk22a5 Sk2ca5	Exon Exon Exon Exon Intron Exon Exon Exon Intron Intron Intron Intron	- - - - - - - - - - - - - - - - - - -	· + - + + + + + + + - + - -	8 8 6 12 8 8 8 8 8 7 11 9 9 8 8 0	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -6710 -7623 -7756 137	AGTICAGGTAGAGTIGA GGGTGACAGAGGGGGCA GGTTCATCAGCAGGGGCA GGTTCATCAGCAGGGGCA AGGTCAACGCAGAGGTCA AGGTCAACGCAGAGGTCA AGTICAGGCAGGGGGCA AGTICAGGCAGGGGGCA GGGTGAACTCIGGGTCA GGGTGAACCCCAGGTGA AGTICACACAGAGGCCA AGTICAACAAAAGGTCA AGTICAACAAAAGGTCA	Exon Intron Exon Intron Intron Exon Exon Intron Intron Intron	+ + + + + + + + + + + + + + + + + + + +	· · · · · · · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 8 7 11 8 7 7 7 7	-2314 5998 -6394 5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343	1371 1036 4068 771 -376 325 11 180 -435 1903 -221 -416 -413 -45
Rbm35b Rbp1 Rnf10 Rnf244 Ror1 Rnrb Sat1 Sem32e Spak2 Shark3 Shf Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5	Exon Intron Exon Exon Intron Exon Exon Exon Intron Intron Intron Intron Exon	- - - - - - - - - - - - - - - - - - -	- + - + + + + + + + + + - - - -	8 8 6 12 8 8 8 8 7 11 9 9 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -6710 -7623 -7756 -1767	AGTTCAGGTAGAGTTGA GGGTGACACAGAGGGGTCA GGTTGATCAGCAGGGGTCA AGGTCAGCAGAGGGTGA AGGTCAGCAGAGGTCA AGGTCAGCAGAGGTGA AGTTGATGGCAGGGTTGA GGGTGACTCTGGGGTCA GGGTGACGTCAGGGTCA AGTTGACACTCAGGTCA GGGTCACGTCCAGGTCA	Exon Intron Exon Intron Intron Exon Exon Intron Exon Intron Intron Exon	+ + + + + + + + + + + + + + + + + + + +	- + + + + + + + + + + + + +	8 8 6 12 8 8 8 8 7 11 8 7 7 11 8 7 8 8	-2314 -5998 -6394 -5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912	1371 1036 4068 771 325 11 180 435 1903 -221 416 413 145
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Sat1 Sem32e Spk2 Shank3 Shf Sk22a5 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2	Exon Intron Exon Intron Exon Exon Exon Exon Intron Intron Intron Intron Exon Intron	- - - - - - - - - - - - - - - - - - -	· + - + + + + + - - - - - - -	8 8 6 12 8 8 8 8 8 7 11 9 9 8 8 8 7	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -6710 -7623 -7756 -1767 682	AGTTCAGGTAGAGTTGA GGGTGACAGAGGGGTCA GGGTGACAGAGGGGTCA GGTTCATCAGCAGGTGA AGGTCAACGCAGAGGTCA AGGTCAACGCAGAGGTCA AGTTCATGAGCAGGGTGA AGTTCAGGCAGGGTGA GGGTGAACTCTGGGTCA GGGTGAACCTCAGGTCA GGGTGAACCTCAGGTCA AGTTGACATAAAGGTCA AGTTGACAGAAGGTCA GGGTCAGCTCCGGGTGA AGGTCGACGAGAGGGTCA	Exon Intron Exon Intron Exon Intron Exon Exon Intron Intron Intron Exon Intron	+ + + + + + + + + + + + + + + + + + + +	· · · · · · · · · · · · · · · · · · ·	8 8 8 6 12 8 8 8 7 7 11 8 7 8 7 8 8 8 8 8 8	-2314 -5998 -6334 -5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931	1371 1036 4068 771 375 325 11 180 435 1903 -221 416 413 145 249
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Start Sem32e Spk2 Shank3 Shf Slc22a5 Slc22a5 Slc22a5 Slc22a5 Slc2a3 Slc9a3	Exon Intron Exon Exon Intron Intron Exon Exon Intron Intron Intron Intron Intron Intron Exon Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Exon Intron Exon Intron Exon Exon Exon Intron Exon	- - - - - - - - - - - - - - - - - - -	· + - - + + + + + - - - - - - - - + + -	8 8 6 12 8 8 8 8 8 7 11 9 8 8 8 7 7 7	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -6710 -7623 -7756 -1767 682 -1309	AGTTCAGGTACAGTTCA GGGTCATCCTAAGTTCA GGGTGACACAGAGGGGTCA GGTTCATCAGCAGGGGTCA AGGTCACCAGCAGGGTGA AGGTCACTCAAAGGTCA AGTTCATGCAGAGGTTCA AGTTCAGCAGGGTTGA GGGTCACCCCAGGTGA AGTTGACATTAAGGTCA GGGTCACCGCCAGGTCA AGTTCACACAAAGGTCA AGGTCACTGCGGTCA AGGTCACTGCGGTCA	Exon Intron Exon Exon Intron Intron Exon Intron Intron Intron Exon Intron Exon	+ + + + + + +	· · · · · · · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 8 7 11 8 7 8 7 7	-2314 -5998 -6394 -5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1380	1371 1036 4066 771 375 325 11 180 435 4903 -221 416 413 145 -249 71
Rbm35b Rbp1 Rnf214 Rnf214 Rnf2 Snf2 Stat1 Sema3e Sgk2 Shank3 Shf Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk23a3 Sk23a3 Sk25	Exon intron Exon Exon Intron intron Exon Exon Intron Intron Intron Intron Exon Exon Exon Exon Exon Exon Intron	- - - - - - - - - - - - - - - - - - -	· + - - + + + + - - - - - - - - - + + -	8 8 6 12 8 8 8 8 7 11 9 8 8 7 7 7 9	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -6710 -7623 -7756 -1767 682 -1309 -82	AGTICAGGTAGAGTIGA GGGTGACAGAGGGGTCA GGGTGACAGAGGGGTCA GGTTCATCAGCAGGTGA AGGTCAACGCAGAGGTCA AGGTCAACGCAGAGGTCA AGTTGAGGCAGGGTGA AGTTGAGGCAGGGTGA GGGTGAACTCTGGGTCA GGGTGAACCCCAGGTGA AGTTGACATTAAGGTCA GGGTGAGCCCAGGTGA AGTTGACATTAAGGTCA GGGTGAGCTCAGGTCA GGGTGACGTCCAGGTGA AGGTGCACGTCCAGGTGA AGGTGCACGTCCAGGTGA AGGTCACGTCAGGTCA GGGTCACGTCAGGTCA	Exon Intron Exon Exon Intron Exon Exon Intron Exon Intron Intron Exon Intron Exon Exon	+ + + + + - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 7 7 111 8 7 7 8 8 7 7 8 8 7 7 8 8 7 7	-2314 -5998 -6334 -5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1380	1371 1036 4068 771 376 325 325 11 180 435 1903 -221 416 413 145 -249 71 555
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Srm3e Sem3e Spat Shank3 Shi Sk22a5 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2	Exon Intron Exon Exon Intron Intron Exon Exon Intron Intron Intron Intron Exon Exon Exon Exon Exon Intron Exon	- - - - - - - - - - - - - - - - - - -	- + - - + + + - - - - - - - - - - - - -	8 8 6 12 8 8 8 8 8 8 7 11 9 9 8 8 7 7 7 8 8	943 7034 -2326 6752 -6573 -1559 -460 63 -1150 -6710 -6710 -7563 -7756 -1767 682 -1767 -384	AGTTCAGGTACAGTTGA GGGTGACACAGAGGGGTCA GGTTGATCAGCAGGGGTCA GGTTCATCAGCAGGGGTCA AGGTCACTCACAGAGGTGA AGGTCACTCAAAGGTCA AGTTCAGCAGAGGTTGA GGGTGACTCTCGGGTCA GGGTGACGCCCAGGGTGA AGTTGACGCCAGGGTCA AGTTGACAGCAGGGTCA AGTTGACAGGACGGTCA AGGTCACTGGGTCA GGGTGACGGCCCAGGTCA AGGTGACATGGGTCA	Exon Intron Exon Exon Intron Intron Exon Intron Intron Intron Exon Intron Exon Exon Exon	+ + + + - - - - - - - - - - - - - - -	- + + + + + + + + + + + + + + + + + + +	8 8 6 12 8 8 8 8 7 11 8 7 8 8 7 8 8 7 8 8 8 7 8 8 8 8	-2314 -5998 -6394 -5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1380 -972	1371 1036 4066 771 376 325 11 180 435 1903 -221 416 413 145 -249 71 588
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Rnb Sat1 Sem38 Sgk2 Shank3 Shf Sk22a5 Sk25	Exon Intron Exon Intron Intron Exon Exon Intron Intron Intron Intron Intron Intron Exon	- - - - - - - - - - - - - - - - - - -	· + - - - + + - - - - - - - - - - - - -	8 8 6 12 8 8 8 8 7 7 11 9 8 8 8 7 7 7 8 9 9	943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -620 -6710 -7623 -7756 -1767 -7623 -7756 -1767 -384 -633	AGTICAGGTAGAGTIGA GGGTGACAGAGGGGTCA GGGTGACAGAGGGGTCA GGTTGATCAGCAGGGTGA AGGTCAATCAGCAGGTGA AGGTCAATCAGAAGGTCA AGTTGACAGCAGGGTTCA AGTTGACATCGGGGTGA GGGTGAACCTCTGGGTCA GGGTGAACCTCAGGTCA GGGTCAGCCACAGGTCA AGTTGACATTAAGGTCA GGGTCACGCCAGGTGA AGTTCACACAAAGGTCA AGGTCACGGAAGGTCA AGGTGCACGCAGGTGA AGGTGCACCCCAGGTCA GGGTGACCCCAGGTCA GGGTGACCCCAGGTCA	Exon Intron Exon Exon Intron Exon Exon Intron Exon Intron Exon Intron Exon Exon Exon Exon Exon	+ + + + + - - - - - - - - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 6 12 8 8 8 7 7 111 8 7 7 8 8 7 8 8 7 8 8 7 8 9	-2314 -2314 5998 -6394 -5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1380 -972 -692	1371 1036 4068 771 376 325 11 180 -435 1903 -221 -416 -413 145 -249 71 588 59
Rbm35b Rbp1 Rnf10 Rnf10 Rnf1 Rnf1 Rnf1 St1 Sema3e Spt2 Shank3 Shf Sk22a5 Sk2 Sk2 Sk3 Sk3 Sk3 Sk3 Sk5 Sk5 Sk5 Sk5 Sk5 Sk5 Sk5 Sk5	Exon Exon Exon Exon Intron Intron Exon Exon Intron Intron Intron Intron Exon	- - - - - - - - - - - - - - - - - - -	- - - - - - - - - - - - - - - - - - -	8 8 6 12 8 8 8 8 8 8 7 11 9 8 8 8 7 7 7 8 8 9 8	-943 7034 -2326 6752 -6573 -1559 -460 63 -1150 -6710 -7623 -7756 -1767 -682 -1767 -682 -1309 -384 -633 3592	AGTTCAGGTAGAGTTGA GGGTCATCTAAGTTGA GGGTCATCTAAGTGA GGTTCATCAGCAGGGTCA GGTTCATCAGCAGGGTCA AGGTCAACGAGAGGTCA AGGTCAACGAGAGGTCA AGTTCAGGCAGGGTGA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA AGTTGACATTAAGGTCA GGGTCAGCTCAGGTCA AGGTGACAGCTCAGGTCA AGGTGACAGCAAGGTCA GGGTCACGTCAGGTCA GGGTCACGTCAGGTCA GGGTCACGTCAGGTCA GGTTGAGGCCCAGGTCA	Exon Exon Exon Exon Intron Intron Exon Intron Intron Exon Intron Exon Exon Exon Exon	+ + + + + + - - - - - - - - - - - - - -	· + + + + + +	8 8 6 12 8 8 8 8 8 7 11 8 7 8 8 7 8 8 7 8 8 9 8 8	-2314 -5998 -6394 -5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1380 -972 -692 -5189	1371 1336 4068 771 376 325 11 180 435 1903 -221 416 413 165 -249 71 588 59 -1597
Rbm35b Rbp1 Rnf20 Rnf214 Ror1 Rnb Sat1 Sem38 Sgk2 Shank3 Shf Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk23a3 Smyd5 Sp7 Sp68 Spn Trf72	Exon Intron Exon Exon Intron Intron Exon Exon Intron Intron Intron Intron Intron Exon Exon Exon Exon Intron Exon Exon Exon Exon Exon Intron Intron Intron Intron Intron Intron Intron Exon Exon Exon Exon Intron Intron Intron Intron Intron Intron Intron Intron Intron Intron Intron	- - - - - - - - - - - - - - - - - - -	· + · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 7 7 11 9 8 8 7 7 7 8 8 9 9 8 9 9	943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -620 -6710 -7623 -7756 -1767 -682 -1309 -384 -633 35592 -2855	AGTICAGGTAGAGTIGA GGGTGACAGAGGGGTCA GGGTGACAGAGGGGTCA GGGTGACAGCAGGGGTGA AGGTCAATCAGCAGGTGA AGGTCAATCAGCAGGGTGA AGTTCAGCAGGGGTGA AGTTCAGGCAGGGTTGA GGGTGAACCTCTGGGTCA GGGTGAACCCCAGGTGA AGTTCACACAAAGGTCA AGTTCACACAAAGGTCA AGGTCACGCAGGAGGTCA AGGTCACGTCCAGGTGA AGGTCACGCCAGGTGA AGGTGACCCCAGGGTCA GGGTGACCCCAGGGTCA GGGTGACCCCAGGGTCA AGTTCATCTATAGGTCA AGGTCACATCAGGTCA	Exon Intron Exon Exon Intron Exon Exon Intron Exon Intron Intron Exon Exon Exon Exon Exon Exon	+ + + + + + + + + + + + + + + + + + + +	· · · · · · · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 7 7 111 8 7 8 8 7 8 8 7 8 8 7 8 8 9 9 9	-2314 -2314 5998 -6394 -5991 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1380 -972 -692 5189 -1478	1371 1036 4068 771 376 325 11 180 435 1903 -221 416 413 145 -249 71 588 59 -1597 -1374
Rbm35b Rbp1 Rnf10 Rnf10 Rnf10 Rnf1 Rnf1 Sm32 Sm32 Sm32 Shank3 Shf Sk22a5 Sk2 Sk3 Sk3 Sk3 Sk3 Sk3 Sk3 Sk3 Sk3	Exon Exon Exon Exon Intron Intron Exon Exon Intron Intron Intron Intron Exon Exon Exon Exon Intron Exon Exon Intron Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Intron Exon Intron Intron Exon Intron Intron Intron Exon Intron Intron Exon Intron	- - - - - - - - - - - - - - - - - - -	· · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 8 8 7 11 9 9 8 8 7 7 7 8 8 9 8 8 9 9 8 9 9 8	943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -6710 -7623 -7763 -7763 -1767 682 -1767 682 -1309 -384 -633 3592 -2852 -2852 -2852 -2852	AGTICAGGTAGAGTIGA GGGTCATCLAGTIGA GGGTCATCLAGCAGGGTCA GGTTCATCAGCAGGGTCA AGGTCAACGAGAGGTCA AGGTCAACGAGAGGTCA AGTGCATCGAGAGGTCA AGTTCAGGCAGGGTGA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA AGTTCACCCCAGGTGA AGGTCACACCCCAGGTCA AGGTCACACCCAGGTCA AGGTCACGTCAGGTCA GGGTCAGGTCA	Exon Exon Exon Exon Intron Intron Exon Intron Exon Intron Exon Exon Exon Exon Exon Exon Exon	+ + + + + + + + + + + + + + + + + + + +	· · · · · · · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 7 11 8 7 8 8 7 8 8 7 8 8 7 8 8 9 8 8 9 9 8 8 9 9 8		1371 1336 4068 771 376 325 11 180 435 1903 -221 416 413 165 -249 71 588 59 -1597 -1374
Rbm35b Rbp1 Rnf20 Rnf214 Ror1 Rnb Sat1 Sem38 Spk2 Shank3 Shf Sk22a5 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2	Exon Intron Exon Exon Intron Intron Exon Exon Intron Intron Intron Intron Intron Intron Exon Exon Exon Intron	- - - - - - - - - - - - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 7 7 11 9 8 8 7 7 7 8 8 7 7 7 8 9 9 8 8 9 9 6 6		AGTICAGGTAGAGTIGA GGGTGACAGAGGGGTCA GGGTGACAGAGGGGTCA GGTTGATCAGCAGGGTGA AGGTCAATCAGCAGGTGA AGGTCAATCAGCAGGGTGA AGTTGACGCAGGGTGA AGTTGACGCAGGGTGA GGGTGAACCTCTGGGTCA GGGTGAACCCCAGGTGA AGTTGACATTAAGGTCA GGGTCAGCGCAGGGTGA AGTTGACATTAAGGTCA AGGTCACGGAGAGGTCA AGGTCACGCCAGGGTGA AGGTCACGCCAGGTCA GGGTGACCCCAGGGTCA GGGTGACCCCAGGGTCA GGGTGACCCCAGGGTCA AGTTCATCTATCAGGTCA AGGTCACATCAGGTCA AGGTCACATCAGGTCA	Exon Intron Exon Exon Intron Exon Exon Intron Exon Intron Intron Exon Exon Exon Exon Exon Exon Exon	+ + - + + + + - - - - - - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 7 7 111 8 7 7 8 8 7 8 8 7 8 8 7 8 8 9 9 8 9 9 6 6	-2314 -2314 5998 -6394 -5991 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1380 -972 -692 5189 -1478 -3016 -1478 	1371 1036 4068 771 376 325 11 180 435 1903 -221 416 413 145 -249 71 588 59 -1597 -374 -340
Rbm35b Rbp1 Rnf10 Rnf10 Rnf1 Rnf1 Rnf1 Rnf1 Str1 Sem3e Sgk2 Shank3 Shf Sgk2 Shank3 Shf Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk2 Shr Sk2 Sk2 Shr Sk2 Shr Sk2 Shr Sk2 Shr Sk2 Shr Sk2 Shr Sk2 Shr Sk2 Shr Sk2 Shr Sk2 Shr Sk2 Shr Sk2 Shr Sk2 Shr Shr Sk2 Shr Shr Shr Shr Shr Shr Shr Shr	Exon Exon Exon Exon Intron Intron Exon Exon Intron Intron Intron Exon Exon Exon Exon Exon Intron	- - - - - - - - - - - - - - - - - - -	· · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 7 11 9 9 8 8 7 7 7 8 8 8 7 7 9 9 8 9 9 9 9 9	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -620 -620 -7623 -7756 -1767 -682 -1309 -384 -633 -3592 -2852 2676 -1690	AGTTCAGGTAGAGTTGA GGGTCATCTAAGTTGA GGGTCATCTAAGTGA GGTTCATCAGCAGGGTCA GGTTCATCAGCAGGGTGA AGGTCAACGCAGAGGTCA AGGTCACCGAGGGTCA AGTTGATGATGATGATCA GGGTGACGTCGGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCACGCCCAGGTCA AGTTCACCAGGAGGTCA GGGTCACGCCCAGGTCA AGGTGACAGGAAGGTCA AGGTCACAGGAGGTCA GGGTCACGGTCAGGTCA	Exon Exon Exon Exon Intron Intron Exon Intron Intron Exon Exon Exon Exon Exon Exon Exon Ex	+ + - - + + + - - - - - - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 8 6 12 8 8 8 8 7 11 8 7 11 8 7 8 8 7 9 8 8 9 9 8 9 9 6 9	-2314 -5998 -6394 -5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -931 -931 -972 -692 -589 -1478 3016 -1666	1371 1336 4068 771 -376 325 11 180 -435 1903 -221 -416 -413 145 -249 71 588 59 -1597 -1374 -340 -24
Rbm35b Rbp1 Rrf120 Rrf214 Ror1 Rrb Sat1 Sem38e Sgk2 Shank3 Shf Sk22a5 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2	Exon Infron Exon Exon Infron Infron Exon Exon Infron Infron Infron Infron Infron Infron Exon Infron	- - - - - - - - - - - - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 7 11 9 8 8 7 7 8 8 7 7 8 9 9 8 9 9 6 6 9 6	943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -620 -620 -6710 -7623 -7756 -1767 -7623 -7756 -1767 -384 -633 3592 -384 -633 3592 -2852 2676 -1690	AGTICAGGTAGAGTIGA GGGTCAICCIAAGTIGA GGGTCAICCIAAGTIGA GGGTCAICAGCAGGGGTCA GGTTCAICAGCAGGGGTGA AGGTCAAGCAGAGGGTCA AGTICAGCAGGGGTGA AGTICAGGCAGGGTTCA AGTICAGGCAGGGTCA GGGTCAICCICGGGTCA GGGTCAICCICAGGTCA AGTIGACAITIAAGGTCA AGTICAACAAAAGGTCA AGGTCACGGAAGGTCA AGGTCACGTCAGGGTCA GGGTCAICGGAAGGTCA AGGTCACGTCCAGGTCA AGGTCACGACCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGGGTGGGTCA AGGTCACGGGTGGGTCA AGGTCACGGGGGGGTGA GGTTCAGGGTGAGGGTGA	Exon Exon Exon Exon Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Intron Exon Intron	+ + + + + + + - - - - - - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 6 12 8 8 8 7 7 111 8 7 7 8 7 8 7 8 7 8 8 7 8 8 7 8 8 9 9 6 6	2314 5998 -6394 5998 -6394 -6394 -6394 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1380 -972 -692 -5189 -1478 -3016 -1666 -2816	1371 1036 4068 771 376 325 11 180 435 1903 -221 -416 413 145 -249 71 588 59 -1597 -340 -24 124
Rbm35b Rbp1 Rnf10 Rnf10 Rnf11 Rnf1 Rnf1 Snf1 Sem3e Sgk2 Shank3 Shf Sk22a5 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2	Exon intron Exon Exon intron intron Exon Exon intron intron intron intron Exon intron intron intron intron intron intron intron intron intron intron intron intron intron intron intron intron intron intron Exon intron intron intron intron intron intron intron intron Exon intron intron intron Exon intron Exon intron Exon intron Exon intron Exon intron Exon intron Exon intron Exon Intron	- - - - - - - - - - - - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 7 7 11 9 8 8 7 7 8 8 7 7 8 8 9 9 8 9 9 8 9 9 6 9 9 6 8	-943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -620 -620 -7623 -7756 -1767 -682 -1309 -384 -633 -384 -633 -384 -633 -2852 -2852 -2852 -2676 -1690 -2692 -1145	AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTGA GGGTCATCCTAAGTTGA GGGTCATCCAGCAGGGTCA GGTTCATCAGCAGGGTGA AGGTCAACGCAGAGGTCA AGTGCATCGCAGAGGTCA AGTTGATGATGATGATGA GGGTGAACTCTGGGTGA GGGTGAACTCTGGGTGA GGGTGAACTCCAGGTCA GGGTCAGCCCAGGGTCA AGTTCAACAAGAGGTCA GGGTCAGCTCAGGTGA GGGTCAGCTCAGGTGA GGGTCAGGTCCAGGTGA GGTTCAAGGTGAGTCA GGGTCAGGTGGGTGA GGTCCAGGTGGGTGA GGTCCAGGTGGGTGA	Exon Intron Exon Exon Intron Intron Exon Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Intron Intron Intron Intron Intron	+ + - - + + + - - - - - - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 8 6 12 8 8 8 8 8 7 11 8 8 7 7 8 8 7 7 8 8 9 9 8 9 9 8 9 9 6 9 9	2314 5998 -6394 5998 -6394 -6394 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1380 -977 -692 5189 -1478 -3016 -16666 -2816	1371 1336 4068 771 -376 325 11 180 -435 1903 -221 -416 -413 145 -249 71 588 59 -1597 -1374 -340 -24 124 28
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Rorb Sat1 Sem32e Sgk2 Shank3 Shf Sic22a5 Sic22a5 Sic22a5 Sic22a5 Sic22a5 Sic22a5 Sic23a3 Sic23a3 Sic23a3 Sic25 Sic25a3 Sic25a5 Sic	Exon intron Exon Exon intron intron intron Exon Exon intron intron intron intron intron intron Exon intron Exon Exon Exon Exon Exon Exon Exon Exon Exon intron Exon intron Exon intron Exon intron Exon Exon intron Exon intron Exon intron Exon intron Exon intron Exon intron Exon intron Exon intron Exon intron Exon intron Exon intron intron Exon intron Exon intron Exon intron	- - - - - - - - - - - - - - - - - - -	· · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 7 11 9 9 8 8 7 7 8 8 7 7 8 9 9 8 8 9 9 6 6 9 9 6 6 8 8	943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -6710 -7623 -7756 -1767 -7623 -7756 -1767 -882 -1309 -384 -633 -3592 -2852 -2652 -1145 -2652 -1145	AGTTCAGGTAGAGTTGA GGGTCALCCAGAGGGGTCA GGGTCALCCAGAGGGGTCA GGGTCACCAGCAGGGTGA AGGTCAACGCAGAGGTCA AGGTCAACGCAGAGGTCA AGTTCAGCAGGGGTGA GGGTGAACCTCTGGGTCA GGGTGAACCTCGGGTGA GGGTCAGCCCAGGTGA AGTTCACACAAAGGTCA AGTTCACACAAAGGTCA AGGTCACGTCCAGGTGA AGGTCACGTCCAGGTGA AGGTCACGTCCAGGTGA AGGTCACGTCCAGGTCA AGGTCACGTCCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGTCAGGTCA AGGTCACGGTGGGTGA AGGTCACGTCAGGTCAG	Exon Exon Exon Exon Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron	+ + + + + + + + + + + + + + + + + + + +	· · · · · · · · · · · · · · · · · · ·	8 8 6 12 8 8 8 8 7 7 11 8 7 8 7 8 7 8 7 8 7 8 8 7 8 8 7 8 8 7 8 8 9 9 6 6 9 9 6	-2314 -2314 -5998 -6394 -5998 -6394 -6197 -11884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1380 -972 -692 -5189 -1478 -3016 -1666 -2816 -907	1371 1036 4068 771 376 325 11 180 435 1903 -221 -416 413 145 -249 71 -588 59 -1597 -374 -340 -24 124 238 -20
Rbm35b Rbp1 Rnf10 Rnf10 Rnf1 Rnf1 Rnf1 Sm32 Sm32 Sk2 Shank3 Shf Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk25a23 Sk25	Exon intron Exon Exon Intron Intron Exon Exon Intron Intron Intron Intron Exon Exon Intron Intron Intron Intron Intron Intron Intron Intron Intron Intron Intron Intron Exon Intron Intron Intron Intron Exon Intron Intron Intron Intron Intron Intron Intron Intron Intron Exon Intron Intron Intron Intron Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Intron Exon Intron	- - - - - - - - - - - - - - - - - - -		8 8 6 12 8 8 8 8 7 7 11 9 8 8 7 7 8 8 7 7 8 8 9 9 8 9 9 6 9 9 6 6 8 8 6	943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -620 -620 -620 -7623 -7756 -1767 -682 -1309 -384 -633 3592 -2852 2676 -1690 -2692 2692 -145 304	AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTGA GGGTCATCCTAAGTTGA GGGTCACTCAAGTGGA AGGTCAAGCGAGAGGTCA AGGTCAACGCAGAGGTCA AGTGCATCGAGAGGTGA AGTTGAGGAGGTGGA GGGTGAACTCTGGGTCA GGGTGAACTCGGGTGA GGGTGAACTCCAGGTGA AGTTGACATTAAGGTCA AGGTCACACCCAGGTGA AGGTCACATCCAGGTGA AGGTCACAGCCAGGTCA AGGTCACAGCCAGGTCA AGGTCACAGCCAGGTCA AGGTCACAGGCAGGTCA AGGTCACAGGCAGGTCA AGGTCACAGGCAGGTCA AGGTCACGTCCAGGTGCA AGGTCACGTCCAGGTGCA AGGTCACGTCGGGGGGTGA GGTCACGGCGGGGTGA AGGTCACGTCGGGGGGTGA AGGTCACGGCAGGTCA	Exon Intron Exon Exon Intron Intron Exon Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Intron Intron Exon Intron Exon	+ + - - + + + - - - - - - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 8 6 12 8 8 8 8 7 7 11 8 8 7 7 8 8 8 7 7 8 8 8 9 9 8 9 9 6 9 9 6 9 9 6	2314 5998 -6394 5998 -6394 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 931 -1380 -972 -692 5189 -1478 3016 -2817 -2927 -7243 -743 -743 -743 -743 -743 -743 -743 -743 -743 -747 -743 -747 -743 -747 -747 -748 -747 -	1371 1336 4068 771 -376 325 11 180 -435 1903 -221 -416 -413 145 -249 71 588 59 -1597 -1374 -340 -24 124 238 1529
Rbm35b Rbp1 Rnf120 Rnf214 Ror1 Rorb Sat1 Sem38e Sgk2 Shank3 Shf Sk22a5 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2	Exon Infron Exon Exon Infron Infron Exon Exon Exon Infron Infron Infron Exon Infron Exon Infron Infron Exon Infron Infron Infron Exon Infron Infron Infron Exon Infron Infron Exon Infron Exon Infron Exon Infron Exon Infron Infron Infron Infron Infron Infron Infron Infron Infron Infron Infron Infron Infron Infron Infron Exon Infron Infron Infron Infron Exon	- - - - - - - - - - - - - - - - - - -		8 8 6 12 8 8 8 8 7 11 9 8 8 7 7 8 9 7 8 9 9 8 9 9 6 6 9 6 8 8 8 9 6 8 8 8 9 8 8 9 8 8 9 8 8 9 9 8 8 8 8		AGTTCAGGTAGAGTTGA GGGTCALCCAGAGGGGTCA GGGTCALCAGAGGGGTCA GGGTCALCAGCAGGGGTCA AGGTCAACGCAGAGGTCA AGGTCAACGCAGAGGTCA AGTTCAGCAGGGTGA AGTTCAGGCAGGGTTCA AGTTCAGGCAGGGTTGA AGTTCAGCCAGGGTCA GGGTCACGCCCAGGTCA AGTTCACACAAAGGTCA AGGTCACGCACGGAGGTCA AGGTCACGTCCAGGTGA AGGTCACGTCCAGGTGCA AGGTCACGTCCAGGTCA AGGTCACGTCCAGGTCA AGGTCACGTCCAGGTCA AGGTCACGTCCAGGTCA AGGTCACGTCAGGGTCA AGGTCACGAGGGTCA AGGTCACGTCAGGGTCA AGGTCACGAGGGTCA	Exon Exon Exon Exon Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Intron Exon Intron	+ + + + + + + + + + + + + + + + + + + +	· · · · · · · · · · · · · · · · · · ·	8 8 8 6 12 8 8 8 8 7 11 8 7 11 8 7 7 8 7 8 7 8 9 6 9 6 9 6 9 6 8 8	-2314 -5998 -6394 -6394 -6394 -6394 -6394 -6394 -6394 -7384 -117 -755 -2523 -6489 -7207 -7207 -7343 -1912 -931 -1380 -972 -692 -5189 -1478 -3016 -1666 -2816 -907 -1225 -4339 	1371 1036 4068 771 376 325 11 180 435 1903 -221 -416 413 145 -249 -1597 -1374 -340 -24 228 1529 -844
Rbm35b Rbp1 Rnf10 Rnf10 Rnf1 Rnf1 Rnf1 Sm32 Sm32 Sgk2 Shank3 Shf Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk25a3 Sk27a5 S	Exon intron Exon Exon Intron Intron Exon Exon Intron Intron Intron Intron Exon Exon Intron Intron Intron Intron Intron Intron Exon Exon Intron Intron Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Intron Intron Intron Intron Intron Intron Exon Intron Intron Intron Intron Intron Intron Intron Intron Exon Intron Intron Intron Exon Intron			8 8 6 12 8 8 8 8 7 7 11 9 8 8 7 7 7 8 8 9 9 8 9 9 6 9 9 6 6 8 8 9 9 6 8 7		AGTTCAGGTAGAGTTGA GGGTCATCLAGTAGT GGGTCATCLAGCAGGGTCA GGTTCATCAGCAGGGTGA AGGTCAACGCAGAGGTCA AGGTCAACGCAGAGGTCA AGTTCATGCAGAGGTCA AGTTCAGGCAGGGTGA GGGTGAACTCTGGGTCA GGGTCAGCCCAGGTGA AGTTCACGCCAGGTGA AGTTCACGCAGGAGGTCA AGTTCACACAAAAGGTCA AGGTCACATCCAGGTGA AGGTCACATCCAGGTGA GGGTCACGTCCAGGTCA GGGTCACGTCCAGGTCA GGGTCACGTCCAGGTCA GGGTCACGTCCAGGTCA GGGTCACGTCCAGGTCA GGTCACGTCCAGGTCA GGTCACGTCCAGGTCA AGTTCACATCAGGTGA GGTCACGTCGAGGTGA AGGTCCAGGTGAGTCA AGGTCCAGGTGAGTCA AGGTCCAGGTGAGTCA AGGTCCAGGTGAGTCA AGGTCCACGTGAGGTCA AGGTCCACGTCAGGTCA GGTCACGTCAGGTCAG	Exon Intron Exon Exon Intron Intron Exon Exon Intron Intron Intron Exon Intron Exon Intron Intron Exon Intron Intron Intron Exon Intron Intron Intron Intron Intron Intron Exon Intron Intron Intron Intron Intron Exon Intron	+ + + + + + + - - - - - - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 8 8 8 8 8 8 7 11 11 8 8 7 8 8 8 7 8 8 9 9 8 9 9 6 9 9 6 9 9 6 9 9 6 6 9 6 6	2314 5998 -6394 5998 -6394 -6394 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1380 -972 -692 5189 -1478 3016 -2816 -2917 -29777 -2977 -297	1371 1336 4068 771 376 325 11 180 -435 1903 -221 -416 -413 145 -249 71 588 59 -1597 -1374 -340 -24 124 238 1529 -844 68
Rbm35b Rbp1 Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Ror1 Srm3e Spf2 Shm3e Shm43 Shn43 Shf Sk22a5	Exon infron Exon Exon infron Exon Exon Exon Exon Exon Exon Exon Exon infron	- - - - - - - - - - - - - - - - - - -		8 8 6 12 8 8 8 8 7 11 9 8 8 7 7 8 8 9 6 8 9 6 6 9 6 6 8 8 6 8 8 7 7		AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTGA GGGTCACCAGAGGGGTCA GGGTCACCAGCAGGGGTCA AGGTCAACGCAGAGGGTCA AGGTCAACGCAGAGGGTCA AGTTCAGGCAGGGTGA AGTTCAGGCAGGGTTGA AGTTCAGGCAGGGTGA AGTTGACGCAGGGTGA AGTTGACACAAAGGTCA AGTTCACACAAAGGTCA AGGTCACGCCAGGTGA AGGTCACGTCCAGGTGA AGGTCACGTCCAGGTGCA AGGTCACGTCAGGTGCA AGGTCACGTCAGGTGCA AGGTCACGTCAGGTGCA AGGTCCACGGTGGGTCA GGGTCACGTCAGGTCA AGGTCCACGGTGGGTCA GGTTCATGGGTGGGTCA AGGTCCACGGGTGGTCA AGTTCCACGAGGTCA AGTTCCACGAGGGTCA AGTTCCACGAGGGTCA AGTTCCACGAGGGTCA AGTTCCACGAGGGTCA AGTTCCACGAGGGTCA AGTTCCACGAGGGTCA AGTTCCACGAGGGTCA AGTTCCACGAGGTCA AGTTCCACGAGGTCA AGTTCCACGAGGTCA AGTTCCACGAGGTCA AGTTCCCCAGGGTCA AGTTCCCCAGGGTCA AGTTCCCCAGGGTCA AGTTCCCCAGGGTCA AGTTCCCCAGGGTCA	Exon Exon Exon Exon Exon Intron Intron Exon Intron	+ + + + + + + + + + + + + + + + + + + +		8 8 8 6 12 8 8 8 8 7 11 8 7 7 8 8 7 8 8 7 8 8 7 8 8 7 8 8 7 8 8 9 9 6 6 9 6 9 6 6 9 9 6 8 8 9 9 6	-2314 -5998 -6394 -5998 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -72343 -1912 -931 -1380 -972 -692 -5189 -1478 -3016 -1666 -2816 -907 -1225 -4339 -85 -85 -85 -85 -1226 -1266 -1	1371 1036 4068 771 376 325 11 180 435 1903 -221 4.16 4.13 145 -249 -1597 -1597 -340 -24 238 1529 -844 68 264
Rbm35b Rbp1 Rnf10 Rnf10 Rnf11 Rnf1 Snf1 Sem3e Sgk2 Shank3 Shf Slc22a5 Slc22a5 Slc22a5 Slc22a5 Slc2a3 S	Exon intron Exon Exon Intron intron Exon Exon Intron Intron Intron Exon Exon Intron Intron Exon Exon Intron Exon Intron Intron Intron Exon Intron			8 8 8 12 8 8 8 7 7 11 9 8 8 7 7 9 8 8 9 9 6 9 6 9 6 6 9 6 8 8 9 6 7 7 10		AGTTCAGGTAGAGTTGA GGGTCATCLAGTGAG GGGTCATCLAGCAGGGTCA GGTTCATCAGCAGGGTGA AGGTCAACGCAGAGGTCA AGGTCAACGCAGAGGTCA AGTTCATGCAGAGGTCA AGTTCAGGCAGGGTGA GGGTCAACCCCAGGTGA GGGTCAACCCCAGGTGA AGTTCAGCCCAGGTGA AGTTCACACAAAAGGTCA AGTTCAACAAAAGGTCA GGGTCAACCCCAGGTGA AGGTCCAGCTCAGGTCA GGGTCAACCCAGGGTCA GGGTCACGTCCAGGTCA GGGTCACGTCCAGGTCA GGGTCACGTCCAGGTCA GGGTCACGTCCAGGTCA GGTCACGTCCAGGTCA GGTCACGTCAGGTCA GGTCACGTCAGGTCA GGTCACGTCGGGGGGTA GGTCACGTCGAGGTCA AGTTCCAGCTCAGGTCA AGTTCCAGCAGGTCA AGTTCCAGCTCAGGTCA AGTTCCAGCTCAGGTCA AGTTCCAGCTCAGGTCA AGTTCCAGCTCAGGTCA AGTTCCAGCTCAGGTCA AGTTCCAGCTCAGGTCA AGTTCCAGCTCAGGTCA AGTTCCAGCGTCAGGTCA AGTTCCAGCTCAGGTCA AGTTCCAGCGTCAGGTCA AGTTCCAGCGTCAGGTCA AGTTCCAGCGTCAGGTCA AGTTCCAGCGTCAGGTCA AGTTCCAGGTCAGTCA	Exon Intron Exon Exon Intron Intron Exon Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Intron Exon Intron Intron Exon Intron Exon Intron Intron Exon Intron Intron Intron Intron Exon Intron	+ + + + + + + + - - - - - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 8 8 8 8 8 7 7 111 8 8 7 7 8 8 8 7 7 8 8 8 9 9 6 9 9 6 9 9 6 6 9 6 6 9 9 6 6 9 9 6	2314 5998 -6394 5998 -6394 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 931 -1380 -972 -692 5189 -1478 3016 -2816 -2816 -2816 -2816 -2816 -285 -28	1371 1036 4068 771 376 325 11 180 435 1903 -221 416 413 145 -249 71 59 -1597 -1374 -340 -24 124 28 1529 -844 68 254
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Srm32e Spm	Exon intron Exon Exon intron intron intron Exon Exon intron intron intron intron intron intron Exon intron	- - - - - - - - - - - - - - - - - - -		8 8 8 6 12 8 8 8 7 11 9 8 8 7 7 8 8 9 6 8 9 6 6 9 9 6 6 8 8 9 6 8 8 7 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTCA GGGTCATCCTAAGTTCA GGGTCATCCAGCAGGGTCA GGTTCATCAGCAGGGTGA AGGTCAACGAGAGGTCA AGTGCATCGCAGAGGTCA AGTTCAGGCAGAGGTGA GGGTCAGCTCAGGTGA GGGTCAGCCCAGGTGA GGGTCAGCTCAGGTCA AGTTCAACAGGGTCA GGGTCAGCTCAGGTCA AGGTCAGTCAGGTCA GGGTCAGCTCAGGTCA AGTTCAACAGGTCAGGT	Exon Exon Exon Exon Exon Intron Exon Exon Intron In	+ + + + + + + - - - - - - - - - - - - -	· · · · · · · · · · · · · · · · · · ·	8 8 8 6 12 8 8 8 8 7 11 8 7 7 8 8 7 8 8 7 8 8 7 8 8 9 9 6 6 9 6 9 6 6 9 9 6 8 8 9 9 13	2314 5998 -6394 5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1380 -972 -692 5189 -1478 -3016 -1666 -2816 -907 -1225 -4339 -85 -1239 -4441	1371 1336 4068 771 -376 325 11 180 -435 1903 -221 -416 -413 145 -249 71 588 59 -1597 -1374 -340 -24 124 238 1529 -844 68 254 112
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Rrb Sat1 Sem3e Sgk2 Shank3 Shf Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk23a3 Sk23a3 Smyd5 Sp7 Srp68 Sspn Tcf7l2 Tcfap2c Tfg Tnks1bp1 Top2b Trpc1 Tfc27 Wm1 Wm15a Wt1	Exon intron Exon Exon Intron intron Exon Exon Intron Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Intron Intron Exon Intron	- - - - - - - - - - - - - - - - - - -		8 8 8 6 12 8 8 8 7 7 11 9 8 8 7 7 9 8 8 9 9 6 9 6 9 6 9 6 8 9 9 6 8 9 9 6 7 7 7 10 10		AGTTCAGGTAGAGTTGA GGGTCATCLAGTAGT GGGTCATCLAGCAGGGTCA GGTTCATCAGCAGGGTGA AGGTCAACGCAGAGGTCA AGGTCAACGCAGAGGTCA AGTTCATGCAGAGGTCA AGTTCAGGCAGGGTGA GGGTGAACTCTGGGTCA GGGTCAGCCCAGGTGA AGTTGACATTAAGGTCA GGGTCAGCCCAGGTGA AGTTCACCAGAAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGTCAGGTGGAGGTCA AGTTCCACCAGGGTCA AGTTCCAGCAGGGGGA GGTCAGTCAGGTGGGTGA GGGTCATCGGGGGGGTCA AGTTCCACCAGGGTCA AGTTCCACCAGGGTCA AGTTCCAGCAGGTCA GGGTCATCGGGTGGGTCA GGGTCATCGGTGGGTGA GGGTCATCGGGTGGGTCA AGTTCCACCAGGGTCA GGGTCATCGGGGGCA GGGTCATCGGGGCA GGGCACTCCAGGGTCA AGTTCCACCAGGGTCA GGGTCATCGGGGGCA CAGGTCAGGT	Exon Intron Exon Exon Exon Intron Intron Exon Exon Exon Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron	+ + + + + + + + - - - - - - - - - - - -		8 8 8 8 8 8 8 7 7 111 8 8 7 7 8 8 8 7 7 8 8 9 9 6 9 6 9 6 9 6 9 6 6 9 6 6 9 9 6 6 9 9 6 6 9 9 6 6 9 9 13 10	-2314 -2314 5998 -6394 -6394 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1112 -931 -1380 -972 -692 5189 -1478 -3016 -2816 -907 -1225 -4339 -4441 -855 -1239 -4441 -8508 	1371 1036 4068 771 376 325 11 180 -435 1903 -221 416 413 145 -240 71 58 59 -1597 -1374 -340 -24 238 1529 -844 68 254 112 490
Rbm35b Rbp1 Rnf10 Rnf10 Rnf10 Rnf1 Rnf1 Rnf1 Sm32 Sm32 Sm32 Shank3 Shf Sk22a5 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2	Exon intron Exon Exon intron intron intron Exon Exon Exon intron intron intron Exon intron intron Exon intron intron Exon intron intron intron Exon intron			8 8 8 6 12 8 8 8 7 7 11 9 8 8 8 7 7 8 8 9 6 8 9 6 6 8 9 6 6 8 8 9 6 6 8 8 7 7 10 13 10 10 8		AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTGA GGGTCATCCTAAGTTGA GGGTCATCCAGCAGGGTGA AGGTCAACGAGAGGTGA AGGTCAACGAGAGGTGA AGTTGATCGAGAGGTGA AGTTGATCGAGAGGTGA GGGTCAGCTGAGTGA GGGTCAGCTCAGGTGA GGGTCAGCTCAGGTGA AGTTGACATTAAGGTCA GGGTCAGCTCAGGTGA AGGTGACAGGAAGGTCA AGGTGACAGGAAGGTCA GGGTCAGTCAGGTGA GGGTCAGTGAGGTGA GGGTCAGTGAGGTGA GGGTCAGTGAGGTGA GGGTCACTCAGGTGA AGTTCACACAGGGGTGA AGTTCACACAGGGGTGA AGGTCACTAGGGGGG GGTCCAGGGGGGG GGGTCACTCAGGGGA GGGTCACTCAGGGGCA GGGTCACTCAGGGGCA GGGTCACTCAGGGGCA GGGTCACTCAGGGGCA GGGTCACTCAGGGGCA GGGTCACTCAGGGGCA GGGTCACTCAGGGGCA GGGTCACTCAGGGGCA GGGTCACTCAGGGGCA GGGTCACTCAGGGGCA GGGTCACTCAGGGGCA GGGTCACTCAGGGGCA GGGTCACTCGAGGGCCA GGGTCACTCGAGGGCCA GGGTCACTCGAGGGCCA GGGTCACTCGAGGGCCA GGGTCACTCGGGGCCA GGGTCACTCGGGGCCA GGGTCACTCGGGGCCA GGGTCACTCGGGGCCA GGGTCACTCGGGGCCA GGGTCACTCGGGGCCA GGGTCACTCGGGGCCA GGGTCACTCGGGGCCA GGGTCACTCGGGGCCA GGGTCACTCGGGGCCA	Exon Exon Exon Exon Exon Intron Intron Exon Intron Intro	+ + + + + + + + - - - - - - - - - - - -		8 8 8 6 12 8 8 8 8 7 11 8 8 7 7 8 8 7 8 8 9 9 8 9 9 6 9 9 6 9 9 6 9 9 6 9 9 6 9 9 6 9 9 6 9 9 8 8 8 9 9 8 8 8 9 9 8 8 8 8	2314 5998 -6394 5981 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1912 -931 -1912 -931 -1912 -692 5189 -1478 -3016 -1666 -2816 -907 -1225 -4339 -85 -1239 -4441 -8308 -3781	1371 1336 4068 771 376 325 11 180 435 1903 -221 416 413 165 -249 71 588 59 -1597 -1374 -340 -24 124 128 1529 -844 68 254 112 -490 419
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Ror1 Sat1 Sem32e Sgk2 Shank3 Shf Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk25a23 Sk25a	Exon intron Exon Exon Intron intron Exon Exon intron intron intron intron Exon Exon intron Exon intron Exon intron Intron Intron Exon Intron Intron Intron Exon Intron Intron Exon Intron Intron Exon Intron Intron Exon Intron Intron Exon Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon			8 8 8 6 12 8 8 8 7 7 11 9 8 8 7 7 7 8 8 9 9 6 8 9 9 6 6 9 6 6 8 8 9 6 6 9 6 6 8 8 7 7 10 11 9 8 8 9 9 10 12 12 12 12 12 12 12 12 12 12 12 12 12		AGTTCAGGTAGAGTTGA GGGTCATCCAAGTGA GGGTCACCAGAGGGTCA GGTTCATCAGCAGGGTGA AGGTCAACGCAGAGGTCA AGGTCAACGCAGAGGTCA AGTTCAGGCAGAGGTCA AGTTCAGGCAGGGTGA GGGTCACCCCAGGTGA AGTTCAGCCCAGGTGA AGTTCAACGCCAGGTGA AGTTCAACGCCAGGTGA AGTTCAACGCAGGTCA GGGTCACGTCCAGGTCA GGGTCACGTCCAGGTCA GGGTCACGTCCAGGTCA GGTCAGGGAAGGTCA GGTCAGGTCA	Exon Intron Exon Exon Exon Exon Intron Exon Exon Intron			8 8 8 6 12 8 8 8 8 7 7 8 8 7 7 8 8 7 7 8 8 7 7 8 8 9 9 6 9 6 9 6 9 6 9 6 9 6 9 6 9 6 9	-2314 -2314 5998 -6394 -6394 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -11912 -931 -1380 -972 -692 5189 -1478 -3016 -2816 -907 -1225 -4339 -85 -1239 -4441 -85 -1239 -4441 -85 -1239 -4441 -85 -1239 -4441 -85 -1239 -4441 -85 -125 -1	1371 1036 4068 771 376 325 11 180 435 1903 -221 416 413 145 -244 71 588 59 -1574 -340 -24 28 1529 -844 68 254 112 419 -70'
Rbm35b Rbp1 Rnf10 Rnf10 Rnf11 Rnf1 Rnf1 Sm32 Sm32 Sm32 Shank3 Shank3 Shf Sk22a5 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2	Exon intron Exon Exon intron intron intron Exon Exon intron intron intron intron Exon intron intron Exon intron intron Exon intron intron intron Exon intron intron intron intron Exon intron intron Exon intron Exon intron intron intron intron intron intron intron intron intron intron intron Exon intron			8 8 8 6 12 8 8 8 8 7 7 11 9 8 8 8 7 7 7 8 8 9 9 8 8 9 9 6 6 9 9 6 6 8 9 9 6 8 8 9 9 6 8 8 9 9 10 10 10 10 10 10 10 10 10 10 10 10 10	943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -620 -620 -7623 -7756 -1767 -682 -1309 -384 -633 -384 -633 -3559 -2852 -2852 -2852 -2676 -1650 -2669 -1650 -2669 -1145 -145 -145 -145 -145 -145 -145 -14	AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTGA GGGTCATCCTAAGTTGA GGGTCATCCAAGGGGTGA AGGTCAACGAGAGGTCA AGGTCAACGAGAGGTCA AGTTCATCGAGAGGTCA AGTTGATCGAGATGGTTCA GGGTCAGCTGAGGTGA GGGTCAGCTGAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA AGTTCCATCTAGGTCA GGGTCAGCGGGGGGGG GGTCCAGGGGCA GGGTCACGAGGGTCA AGTTCCACTCAGGGTCA AGTTCCACTCAGGGTCA AGTTCCACTCAGGGTCA AGTTCCACTCAGGGTCA AGTTCCACTCAGGGTCA AGTTCCACTCAGGGTCA AGTTCCACTCGAGGGTCA AGTTCCACTCGAGGGTCA AGTTCCACTCGAGGGTCA AGTTCCACTCGAGGGTCA AGTTCCACTCGAGGGTCA AGTTCCACTCGAGGGTCA AGTTCCACTCGAGGGTCA AGTTCCACTCGAGGGTCA AGTTCCACTCGGGGCCA GGGTCCTCGGGGCCA GGGTCCTCGGGGGCTCA AGTTCCACTCGGGGTCA AGTTCCACTCGGGGGCA GGGTCCTCGGGGGCCA GGGTCCTCGGGGGCCA GGGTCCTCGGGGGTCA AGTTCCACTCCGGGGTCA AGTTCCACTCCGGGGTCA AGTTCCACTCCGGGGTCA	Exon Exon Exon Exon Intron Intron Exon Exon Exon Intron Exon Exon Exon Intron Exon Exon Exon Intron Intron Intron Intron Intron Intron Intron Intron Intron Intron Intron Intron Intron Intron Intron Exon			8 8 8 6 6 12 8 8 8 7 11 8 7 11 8 7 7 8 8 7 11 8 7 7 8 9 9 6 9 9 6 9 6 9 6 9 6 6 8 6 9 9 13 10 8 8 9 9 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	2314 5998 -6394 5998 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -72343 -1912 -931 -1912 -931 -1912 -931 -1380 -972 -692 -5189 -1478 -3016 -1666 -2816 -907 -1225 -4339 -85 -1239 -4441 -8308 -3781 -1239 -4441 -8308 -3781 -1891 -	1371 1336 4068 771 376 325 11 180 435 1903 -221 416 413 145 -249 71 58 59 -1577 -340 -24 124 1529 844 68 254 112 490 419 -701
Rbm35b Rbp1 Rnf10 Rnf214 Ror1 Rnf1 Sard S	Exon intron Exon Exon Intron Exon Exon Exon Intron Intron Intron Intron Exon Exon Intron Intron Exon Intron Intron Intron Exon Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon Exon Intron Exon Exon Exon Intron Exon	- - - - - - - - - - - - - - - - - - -		8 8 8 6 12 8 8 8 7 7 11 9 8 8 7 7 7 8 8 9 9 8 9 9 6 6 9 9 6 8 9 9 6 8 9 9 6 8 9 9 6 8 9 9 10 10 10 10 10 10 10 10 10 10 10 10 10		AGTTCAGGTAGAGTTGA GGGTCATCCAAGTGA GGGTCACCAGAGGGTCA GGTTCATCAGCAGGTGA AGGTCAACGCAGAGGTCA AGGTCAACGCAGAGGTCA AGTTCAGGCAGAGGTCA AGTTCAGGCAGGGTGA GGGTCACCCCAGGTGA AGTTGACATCAGAGGTCA GGGTCAGCCCAGGTGA AGTTCACCAGAGGAGGTCA GGGTCAGCCCAGGTGA AGTTCACCAGGAGGTCA GGGTCACGTCCAGGTCA GGGTCAGCTCAGGTCA GGGTCACGTCCAGGTCA GGGTCACGTCCAGGTCA AGTTCACACTAGGTGA GGTCACGTCAGGTCA GGGTCACGTCAGGTCA GGGTCACGTCAGGTCA GGGTCACGTCAGGTCA GGGTCACGTCAGGTCA GGGTCACGTCAGGTCA GGGTCACGTCAGGTCA GGGTCACGTCAGGTCA GGGTCACGTCAGGTCA GGGTCACGTCAGGTCA AGTTCAGCAGGTGA GGGTCACGGCAGGTCA AGTTCAGCCAGGTCA GGGTCACTCGGGGGCA GGGTCACTCGAGGTCA GGGTCACTCGAGGTCA GGGTCACTCGAGGTCA GGGTCACTCGAGGTCA GGGTCACTCGAGGTCA AGTTCAGCCAAGGTCA AGTTCAGCCCAAGGTCA AGTTCAGCCCAAGGTCA AGTTCACCCAAGGTCA AGTTCACACCGAGGTCA AGTTCACACCGAGGTCA AGTTCACACCGAGGTCA AGTTCACACCGGGTCA	Exon Exon Exon Exon Exon Exon Intron Exon Exon Intron Exon Intron Exon Intron Exon Intron Exon Exon Intron Exon Intron Exon Exon Exon Exon Exon Exon Exon Ex			8 8 8 8 8 8 8 7 7 111 8 8 7 7 8 8 7 7 8 8 7 7 8 8 9 9 8 9 9 6 9 6 9 6 9 6 9 6 6 9 9 6 6 9 9 6 8 8 9 9 8 9 9 8 8 9 9 9 8 8 9 9 9 8	-2314 -2314 5998 -6394 -6394 -6394 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1380 -972 -692 -5189 -1478 -3016 -2816 -907 -1225 -4339 -85 -1239 -4120 	1371 1036 4068 771 376 325 31 180 435 1903 -221 416 413 145 -244 71 588 59 -1574 -340 -24 28 1529 -844 68 254 112 490 419 -701 307
Rbm35b Rbp1 Rnf10 Rnf10 Rnf11 Rnf1 Rnf1 Sm32 Sm32 Sm32 Shank3 Shank3 Shf Sk22a5 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2 Sk2	Exon intron Exon Exon intron intron intron Exon Exon intron intron intron intron intron intron intron Exon Exon intron intron intron intron intron Exon intron Exon intron Exon			8 8 8 6 12 8 8 8 8 7 11 9 8 8 7 11 9 8 8 7 7 11 9 8 8 7 7 9 8 9 6 9 9 6 9 6 8 6 8 6 8 7 7 10 13 10 8 9 9 7 10 13 10 8 9 9 7 7	943 7034 -2326 6752 -6573 -1559 -460 63 -1190 -620 -620 -620 -621 -7756 -1767 -682 -1309 -384 -633 -384 -633 -3592 -2852 2676 -1650 -2692 -1650 -2692 -1145 -304 -3495 -153 -1453 -1453 -1453 -1453 -1455	AGTTCAGGTAGAGTTGA GGGTCATCCAAGTGA GGGTCATCCAAGTGA AGGTCAACCAGAGGGTCA AGGTCAACCAGAGGTCA AGGTCAACCAGAGGTCA AGGTCACCCAAGGGTCA AGTTGATGAGTGGTCA AGTTGACGATGGTGA GGGTCAGCTGGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCACGCCCAGGGTCA AGTTCACCAGGGTCA GGGTCACGCCCAGGGTCA AGGTCACAGGAGGTCA AGGTCACAGGAGGTCA AGGTCACAGGGGTCA AGGTCACGAGGGTCA AGGTCACGAGGGTCA AGGTCACGAGGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACGAGGTCA AGGTCACTCCAGGGTCA AGGTCACTCCAGGGTCA AGGTCACTCCAGGGTCA AGGTCACTCCAGGGTCA AGGTCACTCCAGGGTCA AGGTCACTCCAGGGTCA AGGTCACTCCAGGGTCA AGGTCACTCCAGGGTCA AGGTCACTCCAGGGTCA AGGTCACTCCAGGGTCA AGGTCACTCCAGGGGTCA AGGTCACGGGGGTCA	Exon Intron Exon Exon Intron Intron Intron Intron Intron Intron Intron Exon Intron Intron Exon Intron Intron Intron Intron Intron Intron Intron Intron Exon Intron Intron Exon Intron Intron Exon Intron Intron Exon Intron Intron Exon Intron Exon Exon			8 8 8 8 6 12 8 8 8 8 7 11 8 8 7 11 8 7 7 8 8 7 11 8 7 7 8 8 9 9 6 9 9 6 9 6 9 13 10 8 9 9 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	2314 5998 -6394 5998 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -72343 -1912 -931 -1912 -931 -1912 -931 -1380 -972 -692 -5189 -1478 -3016 -1666 -2816 -907 -1225 -4339 -85 -1239 -4441 -8308 -3781 -1891 -4120 -948	1371 1336 4068 771 376 325 11 180 435 1903 -221 416 413 145 -249 71 58 59 -1577 -340 -24 124 128 1529 -844 68 254 112 -490 -419 -701 307 -272
Rbm35b Rbp1 Rnf10 Rnf11 Rnf1 Rnf1 Rnf1 Sem3e Sgk2 Shank3 Shf Sk22a5 Sk22a5 Sk22a5 Sk22a5 Sk25a23	Exon Intron Exon Exon Intron Intron Exon Exon Intron Intron Intron Intron Exon Exon Exon Intron Intron Intron Exon Intron Intron Exon Intron Intron Exon Intron Exon Intron Exon Intron Exon Intron Exon			8 8 8 6 12 8 8 8 8 7 7 11 9 8 8 7 7 7 8 8 9 9 8 8 9 9 6 6 9 6 6 8 8 6 9 6 8 8 7 7 10 13 10 10 8 8 9 9 7 13 10 12		AGTTCAGGTAGAGTTGA GGGTCATCCAAGTGA GGGTCACCAGAGGGGCA GGTTCATCAGCAGGGTGA AGGTCAACGAGAGGTCA AGGTCAACGAGAGGTCA AGTTCAGCAGAGGTCA AGTTCAGGCAGGGTGA GGGTGAACCTCTGGGTCA GGGTCACGCCAGGTGA AGTTGACATTAAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA AGTTGACATTAAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCCAGGTCA AGGTCAGCTCCAGGTCA AGGTCAGCTCCAGGTCA AGGTCAGCTCCAGGTCA GGGTCAGCTCCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA GGGTCAGCTCAGGTCA AGTTCAGCTAGGTGGAGTCA AGTTCAGCTAGGTGGAGTCA AGTTCAGCTGAGGTGAG GGGTCACGTCAGGTCA AGTTCAGCAGGGCA GGGTCAGCTCAGGTCA AGTTCAGCTCAGGTCA AGTTCAGCTCAGGTCA AGTTCAGCAGGTCA AGTTCAGCAGGTCA AGTTCAGCCAAGGTCA AGTTCAGCCCAAGGTCA AGTTCAGCCCAAGGTCA AGTTCACCTCGGGTCA GGGTCACCTCGGGCCA GGGTCACCTCGGGCCA GGGTCACCTCGGGCCA GGGTCACCTCGGGCCA AGTTCCACCTCAGGTCA AGTTCACCCAAGGTCA AGTTCACCCAAGGTCA AGTTCACCCCAGGTCA AGTTCACCCAGGTCA AGTTCACCCGGGTCA AGTTCACGCGGCCA	Exon Exon Exon Exon Exon Exon Intron Exon Exon Intron Exon Intron Exon Intron Exon Intron Exon Intron Intro			8 8 8 6 12 8 8 8 8 7 7 111 8 8 7 7 8 8 8 7 7 8 8 9 9 8 9 9 6 9 6 9 6 9 6 6 9 9 6 6 9 9 6 6 9 9 6 8 8 9 9 7 12 8 8 8 9 9 9 8 8 9 9 9 8 8 8 9 9 9 9 8 8 8 9 9 9 8	-2314 -2314 5998 -6394 -6394 -6394 -6197 -1884 -471 -117 -755 -2523 -6489 -7207 -7343 -1380 -972 -692 -5189 -1478 -3016 -1666 -2816 -907 -1225 -4339 -85 -1239 -4120 -948 -311 -4120 -948 -311	1371 1036 4068 771 376 325 31 180 435 1903 -221 416 413 145 -249 71 588 59 -1574 -340 -24 124 288 1529 -844 68 254 112 -419 -701 307 -272 1736
Rbm35b Rbp1 Rnf10 Rnf10 Rnf11 Rnf1 Rnf1 St11 Sema3e Spt2 Shank3 Sh	Exon intron Exon Exon intron intron intron Exon Exon intron intron intron intron intron intron intron intron intron Exon intron intron intron intron intron Exon intron intron Exon intron Exon intron Exon intron Exon intron Exon intron Exon intron Exon intron Exon			8 8 8 6 12 8 8 8 8 7 7 11 9 9 8 8 7 7 8 8 9 9 8 8 9 9 6 9 9 6 8 8 9 9 6 8 8 9 9 6 8 8 9 9 6 8 8 9 9 9 6 9 9 9 9		AGTTCAGGTAGAGTTGA GGGTCATCCTAAGTTGA GGGTCATCCTAAGTTGA GGGTCATCCAGCAGGGTGA AGGTCACCGAGAGGTCA AGGTCACCGAGAGGTCA AGTGCATCGAGAGGTCA AGTTGATGCAGATGGTTGA GGGTGACGTCGGGTCA GGGTCAGCTCAGGTGA GGGTCAGCTCAGGTGA GGGTCAGCTCAGGTGA GGGTCAGCTCAGGTGA GGGTCACGCCCAGGGTCA AGTTCACCAGGGTGA GGGTCACGCCCAGGGTCA AGTTCACCAGGGTGA GGGTCACGCCCAGGGTCA AGTTCACCAGGGTGA GGTCACGGGGGGGTGA GGTCACGGGGGGGTGA GGTCCACGGGGGGTGA AGTTCCACTAGGGGGGGTGA GGGTCACCGAGGGTGA AGTTCCACTCAGGGGG GGTCCCGGGGGGTCA AGTTCCACTCAGGGGG GGTCCCGGGGGGTCA AGTTCCACCAGGGTGA AGTTCCACTCAGGGGGG AGTCCAGGGGGGGTCA AGTTCCACCGGGGGGTGA AGTTCCACTCGGGGGGGTCA AGTTCCACGGGGGGGTCA AGTTCCACGGGGGGCA AGTTCCACGGGGGGCA AGTTCCACGGGGGGCA AGTTCCACGGGGGGCA AGTTCCACGGGGGGCA AGTTCCACGGGGGGCA AGTTCCACGGGGGGCA AGTTCCACGGGGGGCA	Exon Intron Exon Exon Exon Intron Int			8 8 8 8 6 12 8 8 8 8 8 7 11 8 8 7 11 8 7 7 8 8 7 11 8 9 9 6 9 6 9 6 9 6 9 6 9 13 10 8 9 9 7 12 6 6	2314 5998 -2314 5998 -6394 -6394 -6394 -117 -755 -2523 -6489 -7207 -7343 -1912 -7207 -7343 -1912 -931 -1380 -972 -692 5189 -1478 -3016 -251	1371 1336 4068 771 376 325 11 180 435 1903 -221 416 413 445 -249 71 58 59 -1577 -374 -24 124 238 1529 -844 68 254 112 -490 419 -701 307 -2727 1736

FIGURE 5—continued

estingly, this analysis revealed that two RAREs are associated to the $RAR\beta 2$, Wt1, and Cyp26A1 genes. However, only one of the RAREs associated to the Wt1 and Cyp26A1 genes was occupied in F9 cells, whereas both RAREs associated to the $RAR\beta 2$ gene were occupied. Of note is that for $RAR\beta 2$, the RARE located in the promoter was more efficiently occupied than the other one, located in an exon, increasing the complexity of the transcriptional regulation of this gene in F9 cells.

Other occupied RAREs were associated to genes that are already known as RA-responsive genes but for which no RAREs had been identified yet. Among these genes are the "stimulated by RA" (*Stra*) genes such as *Bhlhb40* (*Stra13*) (51), *Tcfap2C* (*Stra2*) (52), and *Meis2* (*Stra10*) (53), and zinc finger proteins (*Zfp598* and *Zfp503*) (8, 10). Note that three RAREs are associated to the *Meis2* gene but that only two were occupied by RAR/RXR heterodimers in F9 cells. The analysis also revealed occupied RAREs associated to gene regulatory regions, which were recently found to be occupied by RARs in ChIP-chip and ChIP-qPCR experiments performed with other cell lines but without any indication whether this occupancy was direct or indirect through other bound factors (11). This is exemplified by the RAREs associated to the *Atxn2*, *Top2b*, *Wnt1*, and *Wnt5* genes.



TABLE 1

Numbers of occupied mouse DR5 RAREs in the initial list (15,925 predicted sites), the list of conserved 181 sites located \pm 10 kb from TSSs, and in the final list of 138 RAREs

For finding significant occupied sites, a cutoff p value of 0,00001 was applied. As a control, the same strategy was applied to a random list of 15,925 sequences of 17 kb.

	RAREs							
F9 cells	Random list	Predicted (159,25)	Conserved (181)	Final list (138)				
Untreated	32 (0.2%)	691 (4%)	48 (26%)	39 (28%)				
RA 2 h	62	1,109	49	45				
RA 24 h	39	964	46	33				
RA 48 h	85	1,523 (9%)	71	55				
 Total	112	1,791 (11%)	76 (42%)	58 (42%)				

Most interestingly, a new repertoire of occupied RAREs was found to be associated to new potential RA target genes encoding transcription regulators (RXRB, Jmjd3, Foxa2), several Homeobox genes belonging to clusters (Hoxa3, Hoxb3, Hoxd3), galectins (Lgals2), membrane-associated proteins (Sema3e, Abhd2, Crygn), RNA-binding proteins (Cugbp1, Qk, Srp68, Pcbp4), ATPases (Clpb), and proteins involved in cell death (Sspn), neuronal functions (Agap1), developmental processes (Otp), cell signaling (Raph1, Arpp21, Zdhhc3, Cacna1g, Camk2b, Ephb3, and Pld2), and cytoskeleton organization (Ivns1abp). RAREs were also found associated to the tumor suppressor HIC1 gene, the kallikrein-related peptidase 13 (KLK13) gene, the Myf6 gene, which belongs to the family of muscle regulatory factors, and the Prss27 gene, which encodes a membrane-anchored protease. Note that in the two latter cases, the occupation of the sites decreased after RA addition.

The other RAREs of the bioinformatics list were not occupied by RARs in F9 cells either in the absence or presence of RA. However, as RAR binding relies on the cellular and/or physiological context, one cannot exclude that these RAREs would be occupied in other cell lines or tissues or in other species.

RA Regulation of the Genes Associated to the Selected RAREs— Next we assessed whether the genes associated to the selected RAREs are RA-regulated. Our *in silico* screen identified 138 DR5 RAREs, but taking into account that several RAREs were found associated to a same gene, there are 129 potential RAregulated genes. First, the set of RA-regulated genes was analyzed by high throughput qPCR sequencing (RNA-seq) using F9 cells for which we already had a list of 58 occupied DR5 RAREs. A list of 167 genes that were either induced or repressed after a 4-h treatment with RA was generated after data normalization and identification of the significant differentially expressed genes (supplemental Table S5). This list was finally reduced to 164 distinct genes after removal of the duplicated genes.

Then this list of 164 RA-regulated genes was crossed with the list of 129 RARE-associated genes raised *in silico*, resulting in the selection of 9 RA-responsive genes common to the two lists (Fig. 6). This list includes indeed the canonical RAR target genes (*Cyp26A1*, *RARβ2*, *Rbp1*, *Hoxa1*, and *Hoxb1*). It also includes two new *Hox* genes (*Hoxa3*, *Hoxb3*) as well as two new *Stra* genes, *Tcfap2C* and *BHBLH4*. In F9 cells, for all these nine RA-responsive genes, the associated DR5 RAREs were occupied by RAR/RXR heterodimers, and this occupancy was increased in response to RA (see Table 2).

Conserved and Functional DR5 RAREs

Note that several other genes that are not in our bioinformatics list are activated in RA-treated F9 cells. However, some rely on other DR elements (*cdx1*) and/or reflect the complexity of Hox clustering (*Hoxa5*, *Hoxb5*, *Hoxa4*, and *Hoxb2*) (54–57). Others are known RA-responsive genes (1, 58–60) with DR5 RAREs (see supplemental Table S2) but are not conserved in several species (*Cyp26B1*, *Stra6*, *Stra8*, *Foxa1*, *Gbx2/Stra7*) or are located out of the \pm 10-kb limits (*Gata6*).

The RA-responsiveness of the genes we selected *in silico* was also analyzed in qRT-PCR experiments performed with F9 cells after RA treatment for different times up to 8 h. According to the confidence of the quality of their annotation and sequencing, 49 genes among the 129 genes (supplemental Table 1) were analyzed. This approach confirmed the RA inducibility of the nine genes selected above (Fig. 7, A–D). Interestingly, it also revealed that the inducibility of these genes increases with time (Fig. 7, A–D). Moreover, some additional RAR-bound genes, such as *Meis2*, *KLK13*, and *HIC1*, can be also activated in response to RA but with a low efficiency and at later times (8 h) (Fig. 7, D and E), raising to 12 the list of RA-responsive genes controlled by conserved DR5 RAREs and located ±10 kb from TSSs in F9 cells (Fig. 6).

Given that the RA response of target genes is well known to be cell type-specific, the same qRT-PCR experiments were performed with another RA-responsive mouse embryocarcinoma cell line, the P19 cell line. As shown in Fig. 7, *F–J*, the same genes were activated in response to RA, although with different intensities and kinetics. As an example, the *Hoxa1* and *Meis2* genes were more efficiently activated in P19 cells than in F9 cells. Note that the *Myf6* gene, which was not activated in F9 cells, responded to RA in P19 cells (Fig. 7*I*), raising the number of RA-responsive genes to 13.

Finally, as the RAREs controlling these 13 genes are highly conserved between species (Fig. 5), we analyzed whether they also responded to RA in other cell lines from other species such as a human breast cancer cell line (MCF7 cells) (Fig. 8*A*) and a zebrafish cell line (PAC2) (Fig. 8*B*). The *Bhlhe40* gene was significantly activated in MCF7 cells but not in zebrafish PAC2 cells. In contrast, *Meis2* was strongly activated in PAC2 cells and not in MCF7 cells. These results are summarized in Fig. 9 and point out that the RA response of the new RARE-associated genes we identified may vary from one cell type to the other and from one species to the other.

DISCUSSION

Here we describe a genome-wide *in silico* analysis of consensus DR5 RAREs with recurrent RGKTSA motifs. The advantage of such a computational approach is that it overcomes the chromatin and cellular context and thus provides a direct glance on the whole repertoire of possible RAREs. Moreover, the choice of recurrent RGKTSA motifs was expected to expand this repertoire of RAREs.

This computational study revealed around 15,000 DR5 RAREs in the human and mouse genomes. Among these RAREs, 24% are concentrated in regions located ± 10 kb from the TSSs and the gene ends, and 5–7% are conserved in 6 organisms or more. It also revealed that the degree of conservation of the overall RAREs is not linear with time in the various verte-



TABLE 2

RAR/RXR occupancy of the selected RAREs in F9 cells treated or not with RA (10^{-7} M) for the indicated times

			Occupied	RAREs in F	9 cells
Official gene name	dTSS*	Control	RA 2h	RA 24h	RA 48h
2310007D09Rik	-3436				
4121402D02Rik	-5908				
4632412N22Rik	1259				
4930452B06Rik	-459				
4930562D19Rik	8578				
6430527G18Rik	-3426	+	+	+	+
Abhd2	-1367	+	+	+	+
Aco2	404				
Actn4	-1391	-	+	-	+
Agap1	-9441	-	-	++	++
Aipl1	-2338				
Alk	5054				
Ankrd50	-1989				
Δr	-1258				
Arnn21	-6279		+		+
Atyn2	-561	+	++	++	
BC052040	-1662				1
Rblbo40	1556				
Brof	-864				
Bren 7	4201				Ŧ
Brip/	4301				
BISS	115				
Cachaig	0105	+	+	++	++
Camk2b	-3030	+	+	+	+
Ccdc132	-8558				
Ccno	-5719				
Cdcp2	663	-	-	-	+
Cited2	574	-	+	-	+
Clpb	5382	+	+	+	+
Col24a1	292	-	-	-	+
Crygn	3244	++	++	++	++
Ctsk	2154				
Cugbp1	-379	+	+	-	+
Cugbp2	-5337				
Cyp26a1	-31				
Cyp26a1	-1898	++	++++	+++	+++
Diap3	-229				
E130309F12Rik	745				
Ebf3	-7615				
Elf4	-7145				
Enpp2	2163				
Ephb3	-6596	+	+	-	+
Esrrg	-3113				
Ext2	-7409				
Fbxo30	-5674				
Flt4	3918				
Foxa2	8392	-	+	-	+
Frap1	-7491				
Gabarapl2	522				
Gpr149	-8572				
Gria2	-6381				
Gzmk	-2144				
Hey1	-49	-	-		+
Hic1	1444	+	++	++	++
Hip1r	-2808				
Hmbox1	48				
Hoxa1	1940	+	+++	++++	++
НохаЗ	-6472	+	+	++	+
Hoxa3	6064	+	+++	+++	++
Hoxb1	5100	+	+	++	++
Hoxb3	-6694	+		++	++
Hoxc4	-5034				
Hoxd3	-6731			+	+
Hoyd4	2778				
Hoxd4	2696				
lyns1ahn	-187		4	4	4
Imida	4801				
KIP13	4301		4		4
L colo3	435				
rgaisz	-394	+	++	++	++

	Occupied RAREs in F9 cells				
Official gene name	dTSS*	Control	RA 2h	RA 24h	RA 48h
Lrrc29	-3339				
Meis2	-7698	++	++	++	++
Meis2	-6268	+	++	++	+
Meis2	-6129				
MII1	1810				
Mmn24	-1581				
mmu-let-7c-2	-3954				
mmu-mir-10a	-2566				
Muf6	680	+			
Myo3h	-7559	Ŧ			
Nkanl	-386				
Onecut2	9911				
Ocr1	088				
Otn	-415	+			
Parn 8	9175				
Palpo Behn2	6247				
Pcbp2 Bcbp4	3290				
PC0P4	7024	Ţ	Ţ	Ŧ	T.
Piaz	-7034	+	+		+
Poin	-66Ub				
Pousti	2000				
Prnpip1	-3868				
Prss27	-148	+	+	-	
Ptch1	6913				
Ptchd1	-4895				
Ptprj	-917				
Ptprj	-7803				
Qk	608	+	+++	++	++
Rab11fip2	5995				
Rab39b	410				
Raph1	-8272	+	+	-	+
Rarb	-303	+	++	-	+
Rarb	-4293	+++	++++	++++	++++
Rbbp7	1280	+	+	-	+
Rbm35b	-943				
Rbp1	7034	++	++	+++	++
Rnf10	-2326				
Rnf214	6752				
Ror1	-6573				
Rxrb	-1559	++	++	++	++
Sat1	-460				
Sema3e	63	+	+	+	+
Sgk2	-1190				
Shank3	-620				
Shf	-6710				
Slc22a5	-7623				
Sic22a5	-7756	-	-	-	+
Slc25a23	-1767				
Slc9a3	682				
Smyd5	-1309				
Sp7	-384				
Srp68	-633	+	+	-	+
Sspn	3592	-	+	+	+
Tcf7l2	-2852				
Tcfap2c	2676	-	-	-	+
Tfg	-1690				
Tnks1bp1	-2692				
Top2b	1145	-	+	++	++
Trpc1	304				
Ttc27	3495				
Wnt1	153	+	-	+	+
Wnt5a	1493	+	+		+
Wt1	-4329	-	+	+	
Wt1	7818				
Yes1	3362				
Ypel5	-2592				
Zbtb5	-3813				
Zdhhc3	676	-	+		+
Zfp503	1645	++	++	++	++
7fn598	6048	4			+







C Other genes from the bioinformatic list that are RA-regulated later and/or in other cell types

Hic1
KLK13
Meis2
Myf6

FIGURE 6. **RA-responsiveness of the conserved DR5 RARE-associated genes identified** *in silico* **as assessed by RNA-seq in F9 cells.** *A*, shown are Venn diagrams. *B* and *C*, shown is a summary of the RA-regulated genes.

brates and that the RA gene regulatory network is specifically elaborated in specific groups. Surprisingly, this occurred specifically in placental mammals (eutherians) *versus* all mammals. Indeed, 3-fold more RAREs are conserved in the former than in the latter. As no major events of genomic reorganization are known to have occurred at the base of placental mammals, this elaboration might be specific to RA signaling.

Finally it provided a list of 138 RAREs located \pm 10 kb from TSSs and gene ends and conserved in 6 organisms or more. This list includes the majority of known RAREs, validating the restrictive criteria of our analysis. It also includes RAREs associated to "stimulated by RA" (*Stra*) genes for which no RAREs had been identified yet. The interesting point is that it provided a newly expanded set of high confidence conserved DR5 RAREs associated to a series of new genes involved in transcription, cell signaling, development, neuronal functions, and tumor suppression. The other interesting point is that in some cases, two to three RAREs were found to be associated to a same gene (*e.g. Cyp26A1, RARβ2,* and *Meis2*), increasing the complexity of the transcriptional regulation of these genes.

However, *in silico* identification of RAREs does not assure their functionality. Therefore, we combined the present computational analysis to experimental biology to determine whether the selected RAREs can bind RARs (ChIP-seq) and respond to RA (RNA-seq and qRT-PCR). Such an integrated strategy performed with mouse embryocarcinoma cells (F9 cell line) revealed that 11% of the 15,925 mouse RAREs present in the starting list were occupied by RAR/RXR heterodimers. Interestingly, this percentage increased to 40% in the final list of



FIGURE 7. Real time RT-PCR analysis of the RA regulation of the genes associated to the conserved DR5 RAREs identified *in silico* in F9 (*A–E*) and P19 (*F–J*) mouse embryocarcinoma cells. The results correspond to the mean \pm S.D. of three independent experiments.

conserved RAREs located ± 10 kb from TSSs, validating our selection strategy.

Of note is that, in F9 cells, among the 58 occupied RAREs of our final list, only 12 of the corresponding genes were rapidly activated in response to RA. These genes include indeed the canonical RA target genes (*Cyp26A1*, *RARβ2*, *Rbp1*, *Hoxa1*, Hoxb1) as well as new Hox genes (Hoxa3 and Hoxb3), Stra genes (Tcfap2c, Bhlhe40, Meis2), HIC1, and KLK13. These 12 genes were also activated in another mouse embryocarcinoma cell line (P19). However, some of them (exemplified by the Bhlhe40 and Meis2 genes) did not respond to RA in human MCF7 cells or in zebrafish PAC2 cells. In contrast, another gene, Myf6, which was occupied but not RA-responsive in F9 cells, was significantly induced in P19 cells. This corroborates that the RA regulation of target genes differs from one cell type to the other (Fig. 9), most probably in line with their chromatin context and final feature (differentiation or proliferation). In fact, the majority of the genes associated to occupied RAREs were not RA-regulated in F9 cells. This lack of RA response may be due to the fact that the genes are already expressed (and thus cannot be further stimulated). However, one cannot exclude that RA regulation requires longer times as exemplified for





FIGURE 8. Real time RT-PCR analysis of the RA regulation of the genes associated to the conserved DR5 RAREs in human MCF7 (A) and zebrafish PAC2 (B) cells. The results correspond to the mean \pm S.D. of three independent experiments.

RA-target gene		qRT-PCR validation	in different cell lines		Associated RARE(s)				
	Mouse F9	Mouse P19	Human MCF7	Zebrafish PAC2	Number	Conservation	Occupancy in F9 cells		
Canonical DR5 RARE-associated genes									
Cum26a1				+	1	14	=		
Cypzoar	•	· · · ·	Ť		2	8	+++		
Hoxa1	+	+	+	ND	1	12	+++		
Hoxb1	+	+	+	ND	1	13	+		
Parb			±	Gene loss	1	12	++++		
Raib	, T	Ť			2	9	++		
Rbp1	+	+	ND	ND	1	8	++		
			New DR5 RAF	E-associated genes					
Hoxa3	+	•	ND	ND	1	14	+		
Hoxb3	+	+	ND	ND	1	14	+		
Bhlhbe40	· · · · · · · · · · · · · · · · · · ·	1000 + 1000 -	+		1	12	••••••••		
Tcfap2c	+	+	1		1	6	+		
				+	1	14	++		
Meis2	+	+	100 C		2	12	++		
					3	8	· · · · · · · · · · · · · · · · · · ·		
Hic1	+	· · · · · · · · · · · · · · · · · · ·	1000 (1000 (1000 (1000)))	ND	1	9	+		
Kik13	+		1. S. C. C. C. S.	Gene loss	1	6	· · · · · · · · · · · · · · · · · · ·		
Myf6	1000 - 10 1 000 - 1000	*****	100 C	1997 - 1 997 - 1997 -	1	6	±		

FIGURE 9. Recapitulation of the conserved DR5 RAREs that are RA-activated in mouse embryocarcinoma cells (F9 and P19 cell lines), human breast cancer cells (MCF7 cells), and a zebrafish cell line (PAC2 cells).

Zfp503 (10), specific RARE-mediated conformational changes of the bound RAR (61), and/or cross-talks with other signaling pathways (12, 62), emphasizing the complexity of the RARmediated regulation of gene expression.

Remarkably, the majority of the RAREs present in our *in silico* list were not occupied *in vivo* in F9 cells. This is not surprising, as RAR binding relies on the cellular and physiological context and/or may require other cell specific transcription factors (12). Thus, one can predict that the other RAREs present in the *in silico* list would be occupied in other appropriate cell types or tissues with the corresponding genes being RA regulated under specific conditions.

The final interesting point of this study is the identification of 6 RAREs that are conserved in all the 15 species studied. However, except the RARE associated to the *Gria2* gene, all these

RAREs are located out of the \pm 10-kb limits we defined. Moreover, none of the corresponding genes were RA-regulated in F9 cells, as assessed in RNA-seq experiments, except *Meis1*, which was activated 24 h after the RA addition to F9 cells (10). Nevertheless, these genes are mostly developmental genes (36–41) that are expressed in specific cell types and tissues and at specific developmental stages. Therefore, they might be new markers of the RA response, valid at specific times, in specific tissues from any jawed vertebrate species, opening new avenues for the study of RA signaling during development.

In conclusion, the novelty of the present study resides in an integrated strategy combining genome-wide biocomputing analysis and biological experiments for discovering and characterizing new RAR target genes and response elements. In addition to providing a wider valuable knowledge base for the anal-



ysis of robust RA-responsive genes, such a strategy also brought significant biological information. Indeed, it revealed (i) low conservation of RAREs between human and mouse (6%) and significant differences in the RA regulation of the highly conserved RAR target genes between species. Thus, it suggests that the RA response will differ from one species to the other as well as from one tissue to the other and under different situations. Finally, one can predict that the small set of conserved RAR direct target genes would act as key effectors of evolutionary steps.

Acknowledgments—We warmly acknowledge all members of the team and of the cell culture facilities for help. Special thanks to Regis Lutzing for qRT-PCR and to Bernard Jost and Celine Keime for RNA-seq.

REFERENCES

- Bour, G., Taneja, R., and Rochette-Egly, C. (2006) in *Nuclear Receptors in Development* (Taneja, R., ed.) pp. 211–253, Elsevier Science Publishing Co., Inc., New York
- 2. Duong, V., and Rochette-Egly, C. (2011) Biochim. Biophys. Acta 1812, 1023–1031
- Mark, M., Ghyselinck, N. B., and Chambon, P. (2009) Nucl. Recept. Signal. 7, e002
- 4. Samarut, E., and Rochette-Egly, C. (2011) Mol. Cell. Endocrinol., in press
- 5. Rochette-Egly, C., and Germain, P. (2009) *Nucl. Recept. Signal.* **7**, e005
- 6. Bastien, J., and Rochette-Egly, C. (2004) *Gene* **328**, 1–16
- Cotnoir-White, D., Laperrière, D., and Mader, S. (2011) Mol. Cell. Endocrinol. 334, 76–82
- Eifert, C., Sangster-Guity, N., Yu, L. M., Chittur, S. V., Perez, A. V., Tine, J. A., and McCormick, P. J. (2006) *Mol. Reprod. Dev.* 73, 796–824
- 9. Harris, T. M., and Childs, G. (2002) Funct. Integr. Genomics 2, 105-119
- 10. Su, D., and Gudas, L. J. (2008) Biochem. Pharmacol. 75, 1129-1160
- Delacroix, L., Moutier, E., Altobelli, G., Legras, S., Poch, O., Choukrallah, M. A., Bertin, I., Jost, B., and Davidson, I. (2010) *Mol. Cell. Biol.* 30, 231–244
- 12. Hua, S., Kittler, R., and White, K. P. (2009) Cell 137, 1259-1271
- 13. Hoffman, B. G., and Jones, S. J. (2009) J. Endocrinol. 201, 1-13
- 14. Reddy, T. E., Pauli, F., Sprouse, R. O., Neff, N. F., Newberry, K. M., Garabedian, M. J., and Myers, R. M. (2009) *Genome Res.* **19**, 2163–2171
- Anno, Y. N., Myslinski, E., Ngondo-Mbongo, R. P., Krol, A., Poch, O., Lecompte, O., and Carbon, P. (2011) *Nucleic Acids Res.* 39, 3116–3127
- Rochette-Egly, C., Gaub, M. P., Lutz, Y., Ali, S., Scheuer, I., and Chambon, P. (1992) *Mol. Endocrinol.* 6, 2197–2209
- 17. Taneja, R., Rochette-Egly, C., Plassat, J. L., Penna, L., Gaub, M. P., and Chambon, P. (1997) *EMBO J.* **16**, 6452–6465
- Samarut, E., Amal, I., Markov, G., Stote, R., Dejaegere, A., Laudet, V., and Rochette-Egly, C. (2011) *Mol. Biol. Evol.* 28, 2135–2137
- Bruck, N., Vitoux, D., Ferry, C., Duong, V., Bauer, A., de Thé, H., and Rochette-Egly, C. (2009) *EMBO J.* 28, 34–47
- Bour, G., Plassat, J. L., Bauer, A., Lalevée, S., and Rochette-Egly, C. (2005) J. Biol. Chem. 280, 17027–17037
- 21. Trapnell, C., Pachter, L., and Salzberg, S. L. (2009) *Bioinformatics* 25, 1105–1111
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J., and Pachter, L. (2010) *Nat. Biotechnol.* 28, 511–515
- 23. Anders, S., and Huber, W. (2010) Genome Biol. 11, R106
- 24. Benjamini, Y., and Hochberg, Y. (1995) J. R. Stat. Soc. Series B Stat. Methodol. 57, 289-300
- Umesono, K., Murakami, K. K., Thompson, C. C., and Evans, R. M. (1991) Cell 65, 1255–1266
- 26. Leid, M., Kastner, P., and Chambon, P. (1992) *Trends Biochem. Sci* 17, 427-433

 Loudig, O., Maclean, G. A., Dore, N. L., Luu, L., and Petkovich, M. (2005) Biochem. J. 392, 241–248

Conserved and Functional DR5 RAREs

- de Thé, H., Vivanco-Ruiz, M. M., Tiollais, P., Stunnenberg, H., and Dejean, A. (1990) *Nature* 343, 177–180
- Leroy, P., Nakshatri, H., and Chambon, P. (1991) Proc. Natl. Acad. Sci. U.S.A. 88, 10138-10142
- Lehmann, J. M., Zhang, X. K., and Pfahl, M. (1992) Mol. Cell. Biol. 12, 2976–2985
- 31. Langston, A. W., and Gudas, L. J. (1992) Mech. Dev. 38, 217-227
- Huang, D., Chen, S. W., Langston, A. W., and Gudas, L. J. (1998) *Development* 125, 3235–3246
- Doerksen, L. F., Bhattacharya, A., Kannan, P., Pratt, D., and Tainsky, M. A. (1996) Nucleic Acids Res. 24, 2849–2856
- Milinkovitch, M. C., Helaers, R., Depiereux, E., Tzika, A. C., and Gabaldón, T. (2010) *Genome Biol* 11, R16
- 35. Bromham, L., and Penny, D. (2003) Nat. Rev. Genet. 4, 216-224
- Jing, Y., Machon, O., Hampl, A., Dvorak, P., Xing, Y., and Krauss, S. (2011) Cell. Mol. Neurobiol. 31, 715–727
- Mercader, N., Leonardo, E., Piedra, M. E., Martínez-A, C., Ros, M. A., and Torres, M. (2000) *Development* 127, 3961–3970
- Faralli, H., Martin, E., Coré, N., Liu, Q. C., Filippi, P., Dilworth, F. J., Caubit, X., and Fasano, L. (2011) *J. Biol. Chem.* 286, 23498–23510
- 39. Mead, A. N., and Stephens, D. N. (2003) J. Neurosci. 23, 9500-9507
- Xing, Y., Nakamura, Y., and Rainey, W. E. (2009) Mol. Cell. Endocrinol. 300, 43–50
- Garitaonandia, I., Smith, J. L., Kupchak, B. R., and Lyons, T. J. (2009) J. Recept. Signal. Transduct. Res. 29, 67–73
- Lin, C. Y., Ström, A., Vega, V. B., Kong, S. L., Yeo, A. L., Thomsen, J. S., Chan, W. C., Doray, B., Bangarusamy, D. K., Ramasamy, A., Vergara, L. A., Tang, S., Chong, A., Bajic, V. B., Miller, L. D., Gustafsson, J. A., and Liu, E. T. (2004) *Genome Biol.* 5, R66
- Smeenk, L., van Heeringen, S. J., Koeppel, M., van Driel, M. A., Bartels, S. J., Akkers, R. C., Denissov, S., Stunnenberg, H. G., and Lohrum, M. (2008) *Nucleic Acids Res.* 36, 3639–3654
- Carroll, J. S., Meyer, C. A., Song, J., Li, W., Geistlinger, T. R., Eeckhoute, J., Brodsky, A. S., Keeton, E. K., Fertuck, K. C., Hall, G. F., Wang, Q., Bekiranov, S., Sementchenko, V., Fox, E. A., Silver, P. A., Gingeras, T. R., Liu, X. S., and Brown, M. (2006) *Nat. Genet.* 38, 1289–1297
- 45. Carroll, J. S., and Brown, M. (2006) Mol. Endocrinol. 20, 1707-1714
- Fullwood, M. J., Liu, M. H., Pan, Y. F., Liu, J., Xu, H., Mohamed, Y. B., Orlov, Y. L., Velkov, S., Ho, A., Mei, P. H., Chew, E. G., Huang, P. Y., Welboren, W. J., Han, Y., Ooi, H. S., Ariyaratne, P. N., Vega, V. B., Luo, Y., Tan, P. Y., Choy, P. Y., Wansa, K. D., Zhao, B., Lim, K. S., Leow, S. C., Yow, J. S., Joseph, R., Li, H., Desai, K. V., Thomsen, J. S., Lee, Y. K., Karuturi, R. K., Herve, T., Bourque, G., Stunnenberg, H. G., Ruan, X., Cacheux-Rataboul, V., Sung, W. K., Liu, E. T., Wei, C. L., Cheung, E., and Ruan, Y. (2009) *Nature* 462, 58–64
- 47. Bollig, F., Perner, B., Besenbeck, B., Köthe, S., Ebert, C., Taudien, S., and Englert, C. (2009) *Development* **136**, 2883–2892
- 48. Gillespie, R. F., and Gudas, L. J. (2007) J. Biol. Chem. 282, 33421-33434
- 49. Gillespie, R. F., and Gudas, L. J. (2007) J. Mol. Biol. 372, 298-316
- Lalevée, S., Bour, G., Quinternet, M., Samarut, E., Kessler, P., Vitorino, M., Bruck, N., Delsuc, M. A., Vonesch, J. L., Kieffer, B., and Rochette-Egly, C. (2010) *FASEB J.* 24, 4523–4534
- 51. Boudjelal, M., Taneja, R., Matsubara, S., Bouillet, P., Dolle, P., and Chambon, P. (1997) *Genes Dev.* **11**, 2052–2065
- Oulad-Abdelghani, M., Bouillet, P., Chazaud, C., Dollé, P., and Chambon, P. (1996) *Exp. Cell Res.* 225, 338–347
- Oulad-Abdelghani, M., Chazaud, C., Bouillet, P., Sapin, V., Chambon, P., and Dollé, P. (1997) *Dev. Dyn.* 210, 173–183
- 54. Lickert, H., and Kemler, R. (2002) Dev. Dyn. 225, 216–220
- Tabariès, S., Lapointe, J., Besch, T., Carter, M., Woollard, J., Tuggle, C. K., and Jeannotte, L. (2005) *Mol. Cell. Biol.* 25, 1389–1401
- Coulombe, Y., Lemieux, M., Moreau, J., Aubin, J., Joksimovic, M., Bérubé-Simard, F. A., Tabariès, S., Boucherat, O., Guillou, F., Larochelle, C., Tuggle, C. K., and Jeannotte, L. (2010) *PLoS One* 5, e10600
- Balmer, J. E., and Blomhoff, R. (2005) J. Steroid Biochem. Mol. Biol. 96, 347–354

Downloaded from http://www.jbc.org/ at INRA Institut National de la Recherche Agronomique on June 13, 2018



- Bouillet, P., Sapin, V., Chazaud, C., Messaddeq, N., Décimo, D., Dollé, P., and Chambon, P. (1997) *Mech. Dev.* 63, 173–186
- Bouillet, P., Chazaud, C., Oulad-Abdelghani, M., Dollé, P., and Chambon, P. (1995) *Dev. Dyn.* 204, 372–382
- 60. Oulad-Abdelghani, M., Bouillet, P., Décimo, D., Gansmuller, A., Heyberger, S., Dollé, P., Bronner, S., Lutz, Y., and Chambon, P. (1996) J. Cell

Biol. 135, 469–477

- 61. Meijsing, S. H., Pufall, M. A., So, A. Y., Bates, D. L., Chen, L., and Yamamoto, K. R. (2009) *Science* **324**, 407–410
- Ross-Innes, C. S., Stark, R., Holmes, K. A., Schmidt, D., Spyrou, C., Russell, R., Massie, C. E., Vowler, S. L., Eldridge, M., and Carroll, J. S. (2010) *Genes Dev.* 24, 171–182



Genome-wide *in Silico* Identification of New Conserved and Functional Retinoic Acid Receptor Response Elements (Direct Repeats Separated by 5 bp)

Sébastien Lalevée, Yannick N. Anno, Amandine Chatagnon, Eric Samarut, Olivier Poch, Vincent Laudet, Gerard Benoit, Odile Lecompte and Cécile Rochette-Egly

J. Biol. Chem. 2011, 286:33322-33334. doi: 10.1074/jbc.M111.263681 originally published online July 29, 2011

Access the most updated version of this article at doi: 10.1074/jbc.M111.263681

Alerts:

- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material: http://www.jbc.org/content/suppl/2011/07/29/M111.263681.DC1

This article cites 60 references, 19 of which can be accessed free at http://www.jbc.org/content/286/38/33322.full.html#ref-list-1