Supplemental Material

Figure 1: Organisation of the *Igf2* - *Mrpl23* region in human and chicken

A: Multiple alignments of genomic sequences between *Igf2* and *Mrpl23*. The genomic human sequence was taken as reference sequence and compared to the genomic mouse, and chicken sequences. Before alignment, repetitive elements had been masked using the REPEATMASKER software (Smit and Green, RepeatMasker at http://repeatmasker.org). Aligned regions are shown in green, highly conserved elements in read (>70% identity, >100 bp length). Positions of the human *IGF2*, *H19* and *MRPL23* exons, of the imprinting center IC1, and of the differentially methylated regions (DMRs) are indicated. The position of CpG islands in the human sequence is indicated by the CpG island plot above the multiple alignment. Position of enhancer elements downstream of H19 are also shown (Ishihara *et al.* (2000), Genome Res. 10: 664-671).The given scale bar is related to the human sequence. The CpG island plot shows CpG islands that fulfil the definition of a CpG island (length >200 bp, G+C content >50%, $\text{CpG}_{\text{observed}}/\text{CpG}_{\text{expected}} > 0.6$, http://www.ebi.ac.uk/emboss/cpgplot/). **B**: The genomic chicken sequence was taken as reference sequence in an aligment to the human sequences. Similar to the upper portion (A) of the figure positions of exons and CpG islands in the chicken sequence are indicated.

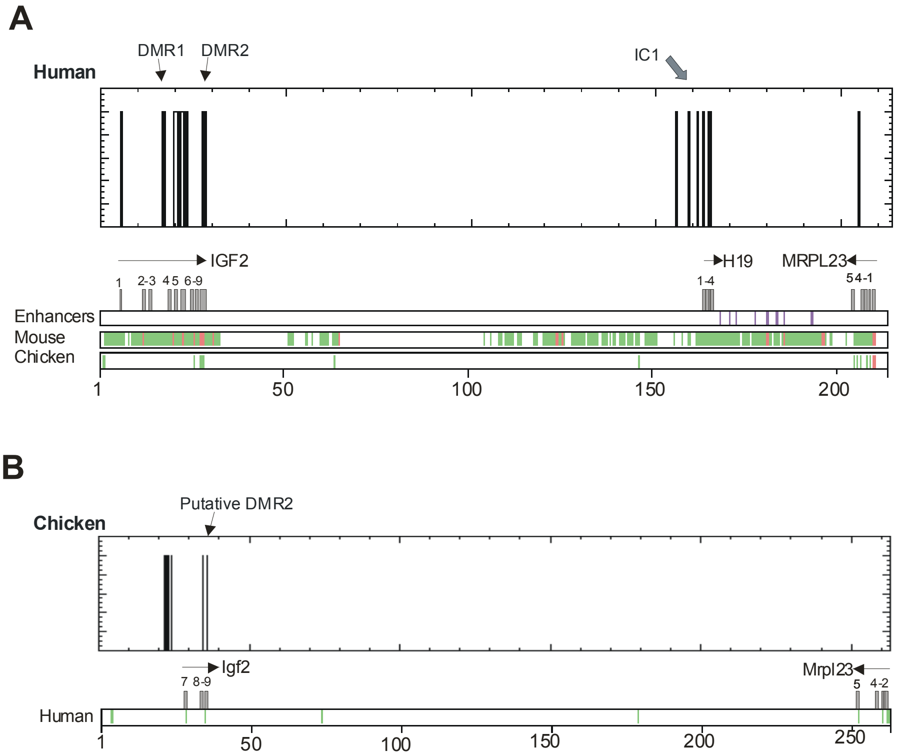

Table 1: Chromosomal positions of orthologous zebrafish genes

Listed are the genes that were included in the analyses. For most genes the human orthologues were taken for analyses. In some cases that are labelled by one asterisk* the mouse sequence was taken for the analyses. Corresponding GenBank accession no. are given. For the searches the protein sequences encoded by the human or mouse genes, respectively, were compared to the translated genomic zebrafish DNA sequences. The given positions reflect the distance to the top of the given chromosome (http://www.ensembl.org, zebrafish database: version 3, release 27th November 2003). Matches to chromosome 7 are highlighted by grey background. The numbers of listed positions per gene were limited to the four best hits. nd: no zebrafish sequences identified. UN: chromosomal position is unknown, instead the genomic contig is given.

Human		**zebrafish**							
gene	accession no	chr	position	chr	position	chr	position	chr	position
MUC2	NM_002457	5	23.8	4	18.8	18	0.4	7	11.9
STK29	NM_003957	16	3.1	22	22.1	25	7.3	18	18.1
HCCA2	NM_053005	25	15.4	5	19.3	8	11.3	11	7.9
DUSP8	NM_004420	3	14.2	21	7.7	14	8.3	22	30.1
CTSD	NM_001909	18	10.3	8	27.9				
*Syt8**	NM_018802	14	26.4	24	4.3	21	8.1	18	3.1
TNNI2	NM_003282	3	5.0	14	26.3	6	2.8	22	16.8
LSP1	NM_002339	7	28.6						
TNNT3	NM_006757	7	28.5	21	8.1	4	7.7	3	4.9
MRPL23	NM_021134	UN	NA54278						
IGF2	NM_000612	25	13.7	7	32.6	4	21.1	8	7.2
INS2	NM_000207	14	19.9	25	13.7	4	21.1		
TH	NM_000360	18	12.6	25	13.7				
ASCL2	NM_005170	7	31.0	8	19.1	22	14.9	1	30.0
PHEMX	NM_005705	nd							
CD81	NM_004356	25	6.3	4	10.6	21	7.5	17	32.7
TSSC4	NM_005706	nd							
TRPM5	NM_014555	7	32.4	3	14.3	13	17.6	12	1.6
KCNQ1	NM_000218	7	32.4	8	18.2	19	23.2	2	16.4
CDKN1C	NM_000076	7	32.4	8	25.4				
SLC22A1L	NM_002555	7	28.5	5	28.5	9	1.6		
PHLDA2	NM_003311	UN	NA32869						
NAP1L4	NM_005969	7	17.4	14	2.5				
CARS	NM_001751	7	17.4						
*Tnfrs22**	NM_023680	5	12.6	16	10.1	8	15.1		
*Tnfrs23**	NM_024290	5	12.5	16	10.1	8	15.1		
*Tnfrs26**	NM_175649	5	12.5	16	10.1	8	15.0		
OSBPL5	NM_020896	1	16.5	9	23.6	16	23.5	22	15.1
*Mrgg**	AY042212	nd							
*Mrge**	XM_146164	nd							

Table 2: Identified fugu scaffolds

Listed are the genes that were included in the analyses. For most genes the human orthologues were taken for analyses. In some cases that are labelled by asterisks the mouse sequence was taken for the analyses. Corresponding GenBank accession no. are given. For the searches the protein sequences encoded by the human ormouse genes, respectively, were compared to the translated genomic fugu DNA sequences. The numbering of the identified fugu sequence scaffolds corresponds to their numbers in the database (Release 3 March 2003, version 18.2.1, http://www.ensembl.org). The numbers of listed sequence scaffolds per gene were limited to the four best hits. nd: no fugu sequences identified. Matches to scaffold 9 are highlighted by grey background.

Human		**Fugu**			
gene	accession no	similar sequence scaffolds			
MUC2	NM_002457	981	3671	65	124
STK29	NM_003957	857	1306	9	1435
HCCA2	NM_053005	9	198	204	40
DUSP8	NM_004420	9	3433	629	316
CTSD	NM_001909	46	197	232	2471
*Syt8**	NM_018802	1010	88	871	65
TNNI2	NM_003282	321	2211	54	950
LSP1	NM_002339	4295			
TNNT3	NM_006757	253	1617	824	1711
MRPL23	NM_021134	9			
IGF2	NM_000612	9	253		
INS2	NM_000207	580	382		
TH	NM_000360	9	6636	981	253
ASCL2	NM_005170	902	253		
PHEMX	NM_005705	nd			
CD81	NM_004356	981	2965	1329	1735
TSSC4	NM_005706	nd			
TRPM5	NM_014555	71	1306	319	983
KCNQ1	NM_000218	350	3568	309	391
CDKN1C	NM_000076	723	71	350	127
SLC22A1L	NM_002555	6006			
PHLDA2	NM_003311	9	1970	54	
NAP1L4	NM_005969	1970	9		
CARS	NM_001751	1438	1849		
*Tnfrsf22**	NM_023680	173	760	87	1036
*Tnfrsf23**	NM_024290	1506	228	173	87
*Tnfrsf26**	NM_175649	87	1506	760	1036
OSBPL5	NM_020896	9	219	2477	1235
*Mrgg**	AY042212	nd			
*Mrge**	XM_146164	nd			

Table 3: Identified human paralogues

Listed are genes for which human paralogues have been identified (see Material and Methods). The given positions of the paralogues are determined by aligment of the cDNA sequences of identified genes to the human genomic sequences in the ENSEMBL data base (http://www.ensembl.org, Release 4 November 2003, version 18.34.1).

gene name	paralogue name	chr	position (Mb)	paralogue name	chr	position (Mb)	paralogue name	chr	position (Mb)	paralogue name	chr	position (Mb)
MUC2	*MUC5AC*	11	nd	*VWF*	12	5.5	*LOC340990*	11	17.5			
	MUC5B	11	1.2									
STK29	*KIAA1811*	19	60.5									
DUSP8	*DUSP16*	12	12.5	*LOC340852*	10	74.8						
SYT8	*SYT1*	12	78.1	*SYT2*	1	199.9						
TNNI2	*TNNI1*	1	198.7	*TNNI3*	19	60.4						
TNNT3	*TNNT1*	19	60.3	*TNNT2*	1	198.6						
IGF2	*IGF1*	12	101.3									
TH	*PAH*	12	101.7									
ASCL2	*ASCL1*	12	101.9									
CD81	*CD9*	12	6.2	*TSPAN-2*	1	114.9						
TRPM5	*TRPM4*	19	54.4	*TRPM2*	21	44.6	*TRPM8*	2	235.1	*TRPM3*	9	68.6
CDKN1C	*CDKN1B*	12	12.8									
PHLDA2	*PHLDA1*	12	74.7	*PHLDA3*	1	198.7						
NAP1L4	*NAP1L1*	12	74.7									
*Tnfrsf22**	*TNFRSF1A*	12	6.3									
*Tnfrsf23**	*TNFRSF1A*	12	6.3									
*Tnfrsf26**	*TNFRSF1A*	12	6.3									
OSBPL5	*OSBPL8*	12	75.2									
MRGG/ MRGE	*MRGX3*	11	18.1	*MAS1L1*	6	29.6	*MAS1*	6	160.2	*MRGF*	11	68.5
	MRGX4		18.3									
	MRGX1	11	18.9									
	MRGX2		19.0									

Table 4: Identified highly conserved NICE elements

Listed are the positions of NICE elements in human and mouse sequences of which GenBank accession numbers are given.

Human **Acc. no. AJ006345.1**	**position**
NICE1	263898-264305
NICE2	268312-268891
NICE3	292486-292668
NICE4	302048-302397
Mouse Acc. no: AJ271885.2	**position**
NICE1	131424-131026
NICE2	126472-126126
NICE3	108717-108532
NICE4	96434-96071

Table 5: Putative IC2 CpG islands in mammals and chicken.

Given are GenBank accession numbers of genomic sequences, and the positions of the analysed segments containing the putative IC2 therein.

	Accession no	**position (nt)**
human	AJ006345.1	255617-253983
galago	AC147392.2	172075-166876
cow	AC147396.2	39982-38098
mouse	AJ271885.2	140331-142078
bat	AC146964.2	41686-40306
armadillo	AC147402.3	49501-48242
chicken	BX649221.3	54949-53748

Table 6: Control groups of human and murine CpG islands.

Given are the GenBank accession numbers of the genomic sequences and the positions of the selected CpG islands therein.

	accession no.	**position (nt)**
human	AC007879.4	36652-37201
	AC009225.4	21911-22280
	AL359815.26	76361-77040
	AL049869.6	137251-137980
	AC022460.4	14754-15580
	AC010552.6	112311-113530
	AC098805.2	108111-109540
	AC011140.4	46721-47910
	AC105337.7	111421-112620
	AC090398.12	97251-97820
mouse	AC087062.25	47785-48424
	AC079775.6	64193-65102
	AC073803.1	16421-16800
	AC087134.1	102181-102670
	AC073711.2	14061-14530
	AC025053.3	54059-54758
	AC073732.1	119961-120700
	AC055818.9	126981-127580
	AC087063.20	35297-36166
	AC122248.5	88859-89838