3 Results

3.1 Embryonic expression profiles of *Dpf3* splice variants

The early embryonic expression pattern of *Dpf3* during mouse embryogenesis has been described previously (Kaynak, 2005). To characterize the splice variant specific expression levels of *Dpf3a* and *DPF3b* throughout all stages of heart development, mouse embryos were collected at stages E9.5 and onwards, hearts dissected from the rest of the body. Isolated hearts were subsequently pooled, RNA extracted and used for real-time PCR using splice variant specific primers (Figure 3.1).

In line with data obtained from *in situ* hybridization experiments (Figure 1.5), expression levels of *Dpf3a* were highest at E9.5, while *Dpf3b* was expressed at lower levels. The expression levels of both splice variants gradually dropped and remained low throughout the second half of gestation (E12.5-E16.5). Expression levels slightly increased directly after birth to levels observed at E11.5. In adult hearts, both splice variants were persistently expressed with *Dpf3b* having slightly higher expression levels.

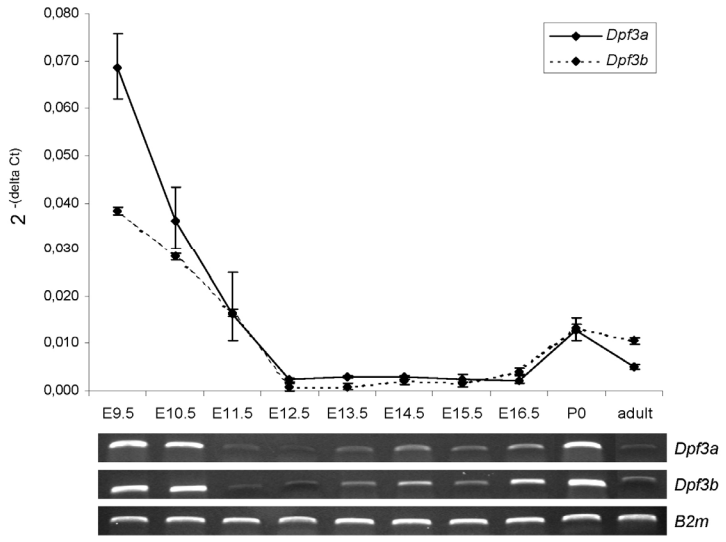

Figure 3.1: Temporal expression profiles of *Dpf3* splice variants (*Dpf3a*/*Dpf3b*) during mouse embryogenesis and adulthood in the heart. Real-time PCR was performed on samples obtained from isolated embryonic and adult hearts at indicated stages. Expression values were normalized to housekeeping gene *B2m*.

3.2 The role of *dpf3* in zebrafish embryonic development

To characterize the functional role of Dpf3 during development, the zebrafish system was used. Zebrafish embryos are widely used to study processes of development as they offer several advantages over other model organisms. They are easy to maintain and produce a large quantity of offspring. Furthermore, embryos develop externally, are transparent and thus easily manipulated. Several tools such as antisense-oligonucleotide-mediated gene knockdown, mRNA overexpression and ENU-mutagenesis are available. The fact that zebrafish embryos are able to take up oxygen from the surrounding water makes them less dependent on a functional circulatory system, enabling the study of genes that affect heart function for a longer period of time compared to other model systems such as the mouse. *In situ* hybridization and knockdown of *dpf3* in zebrafish were performed in collaboration with Dr. Salim Abdelilah-Seyfried at the Max Delbrück Center for Molecular Medicine (MDC) in Berlin and Dr. Wolfgang Rottbauer at the University of Heidelberg. A complete summary of these results can be found in the accompanied manuscript.

3.2.1 Cloning of full-length zebrafish *dpf3*

Sequence database searches revealed that a full-length transcript of zebrafish *dpf3* had not been cloned previously. A bioinformatically predicted, putative full-length clone (XM_686113) was found, which included additional intron sequence leading to a falsely translated C-terminus of the predicted protein. PCR primers were designed based on sequence similarity to mouse *Dpf3* and the full-length cDNA of zebrafish *dpf3* was amplified from adult zebrafish total RNA, cloned into pCR-II-TOPO and verified by sequencing. The sequence was submitted to NCBI and assigned as the new reference sequence NM_001111169. The zebrafish dpf3 protein has a total sequence similarity of 70% compared to human (Figure 3.2). Regions of high conservation are found within the N-terminal 2/3 domain, the nuclear localization signal (NLS), nuclear receptor interaction domain (NID), as well as in the two PHD fingers (PHD1, PHD2). The C2H2-type zinc finger is also highly conserved, expect for a 15 amino acid insertion in zebrafish.

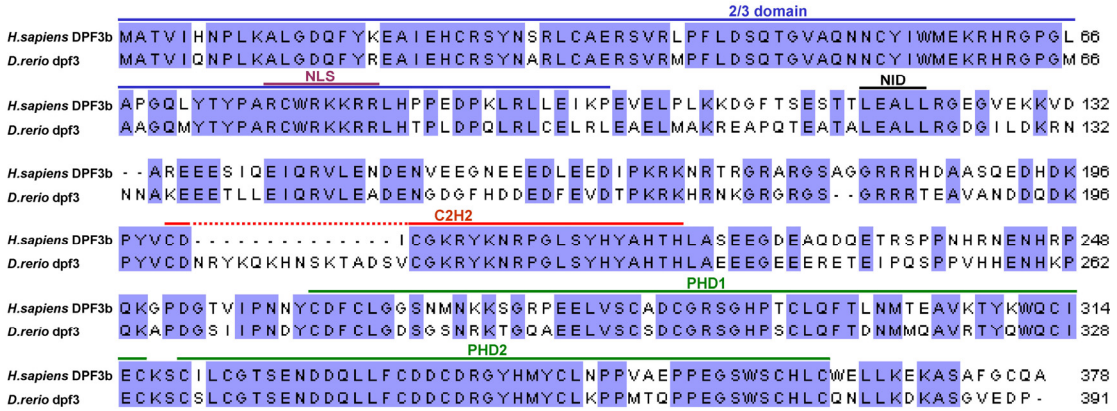

Figure 3.2: Comparison of DPF3 protein sequences from human (AAX20019.1) and zebrafish (NP_001104639). Nuclear localization signal, (NLS); nuclear receptor interaction domain, (NID); C2H2, zinc-finger domain; plant-homeodomain, (PHD).

3.2.2 *In situ* hybridization of *dpf3* in zebrafish embryos

It had previously been established that *Dpf3* transcripts are expressed early in the heart, skeletal muscle and the central nervous system of mouse and chicken (Kaynak, 2005). As an initial step in the characterization of dpf3 in zebrafish, RNA *in situ* hybridization was performed to test whether the expression pattern is also conserved in this species.

Embryos were collected at 36 and 72 hours post fertilization (hpf) and stained with a digoxygenin-labeled RNA antisense probe against *dpf3* (Figure 3.3). At 36 hpf, *dpf3* was strongly expressed within the developing brain and throughout somitic tissues along the entire length of the embryonic trunk and tail. Within the heart, *dpf3* was expressed in the ventricle (indicated by the asterisk), whereas atrial expression was undetectable. At 72 hpf, a similar pattern of expression was observed with additional expression in the neural tube of the embryo.

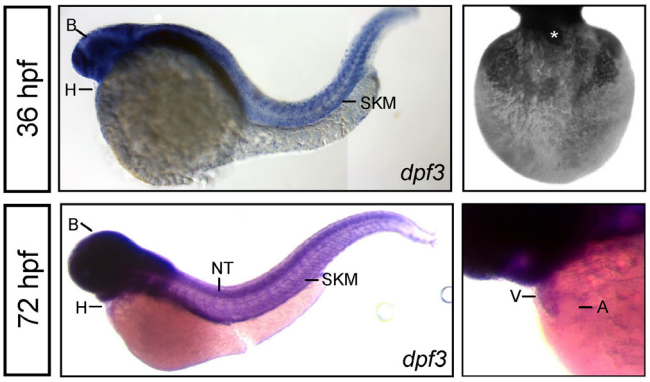

Figure 3.3: RNA *in situ* hybridization of *dpf3* mRNA expression in zebrafish embryos. Upper panel, 36 hpf; lower panel, 72 hpf. B, brain; H, heart; SKM, skeletal muscle; NT, neural tube; V, ventricle; A, atrium; *, ventricle.

3.2.3 Knockdown of *dpf3* by morpholino antisense-oligonucleotide injection of zebrafish embryos

Alteration of the expression level of a gene can be used to analyze its function during development and adulthood. The function of dpf3 during development of the zebrafish embryo was investigated using morpholino-modified antisense-oligonucleotide (MO)-mediated gene knockdown. Embryos derived from MO^{dpf3} injections will be referred to as *dpf3* morphants in the following text. The efficacy of the MO^{dpf3} was tested by RT-PCR using cDNA obtained from *dpf3* morphant embryos and controls, which showed that the majority of *dpf3* mRNA was incorrectly spliced leading to two truncated proteins (Figure 3.4A). MO^{dpf3} targeted the exon4-intron4 boundary of *dpf3* pre-mRNA and interfered with splicing by sterical blockage of the splice-donor site of exon 4. This led to either full integration of intron 4 into the mature transcript or partial intron 4 integration through usage of a cryptic splice-site. cDNA cloning and sequencing of the altered transcripts revealed changes in amino acid sequence at position 178 and premature stop codons at positions 235 and 251, respectively (Figure 3.4B). The 5´ 2/3-domain was unaffected by the morpholino injection, but the C2H2-type zinc-finger and the double plant-homeodomain at the 3´end were not present in *dpf3* morphant transcripts. Protein domain searches using SMART (http://smart.embl-heidelberg.de) and PFAM (http://pfam.sanger.ac.uk) showed that the newly translated sequences had no similarity to any known protein domains.

The specificity of the MO^{dpf3} was demonstrated by co-injection of synthetic and mature *dpf3* mRNA resulting in rescue of the MO^{dpf3} phenotypes (Figure 3.5).

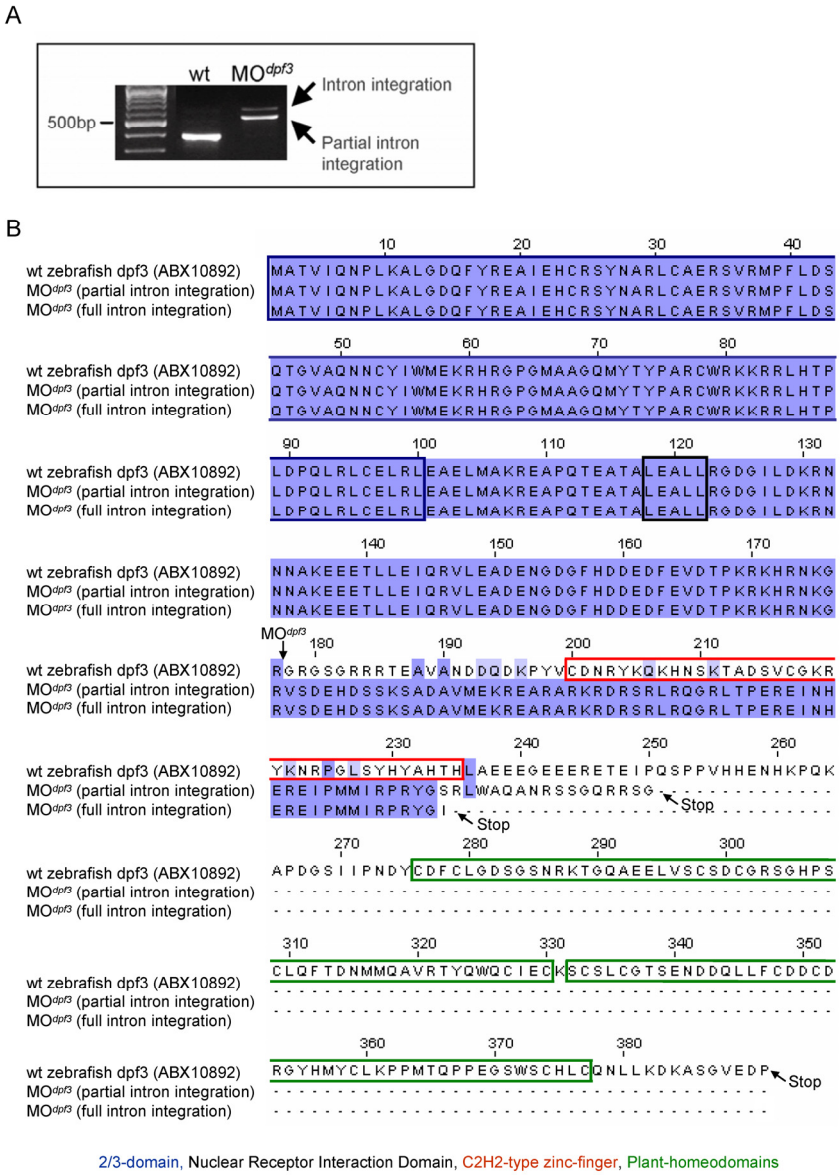

Figure 3.4: Effect of MO^{dpf3} injection on *dpf3* pre-mRNA splicing and corresponding protein sequence. A) PCR analysis of MO^{dpf3} efficacy. Embryos from MO^{dpf3}-injections and controls were harvested at 72 hpf. Total RNA from embryos was isolated, reverse transcribed to cDNA and tested by PCR. Primers were used to amplify a region of the *dpf3* transcript including the MO^{dpf3} targeting site of exon4-intron4. PCR products were cloned and analyzed by sequencing. B) dpf3 protein sequences of wildtype- and MO^{dpf3}-injected embryos inferred from cDNA sequencing. Premature stop-codons and wildtype stop-codon are indicated. 2/3-domain, blue; putative Nuclear Receptor Interaction Domain, black; C2H2-type zinc-finger, red; plant homeodomains, green.

The embryos obtained from MO^{dpf3}-injections were characterized phenotypically using a variety of methods. To assess cardiac morphogenesis and differentiation, the MO^{dpf3} was used in a transgenic line of zebrafish that expresses green fluorescent protein (GFP) under control of the *cardiac myosin light chain 2* (*cmlc2*) promoter region [Tg(*cmlc2*:GFP)], thus labeling all cells of the myocardium (Huang *et al.*, 2003). Injection of MO^{dpf3} at the 1-cell stage resulted in 73 % of embryos with a curved tail (n=255) and in 91 % of embryos with abnormal heart morphology (n= 101) at 36 hpf (Figure 3.5A, B). Consistent with strong somitic expression of

dpf3, MO^{dpf3}-injected embryos frequently displayed disturbed forward swimming movements indicating skeletal muscle defects. The heart phenotype was characterized by a thin and elongated heart tube, with both ventricular and atrial portions being affected. Moreover, looping of the heart was strongly reduced (Figure 3.5B). Co-injection of synthetic full-length *dpf3* mRNA produced a significant rescue effect, with the percentage of embryos with a *dpf3* morphant heart phenotype decreasing to 63 % (n= 145) and the percentage of embryos with abnormal body morphology decreasing to 46 % (n=321) (Figure 3.5A, B).

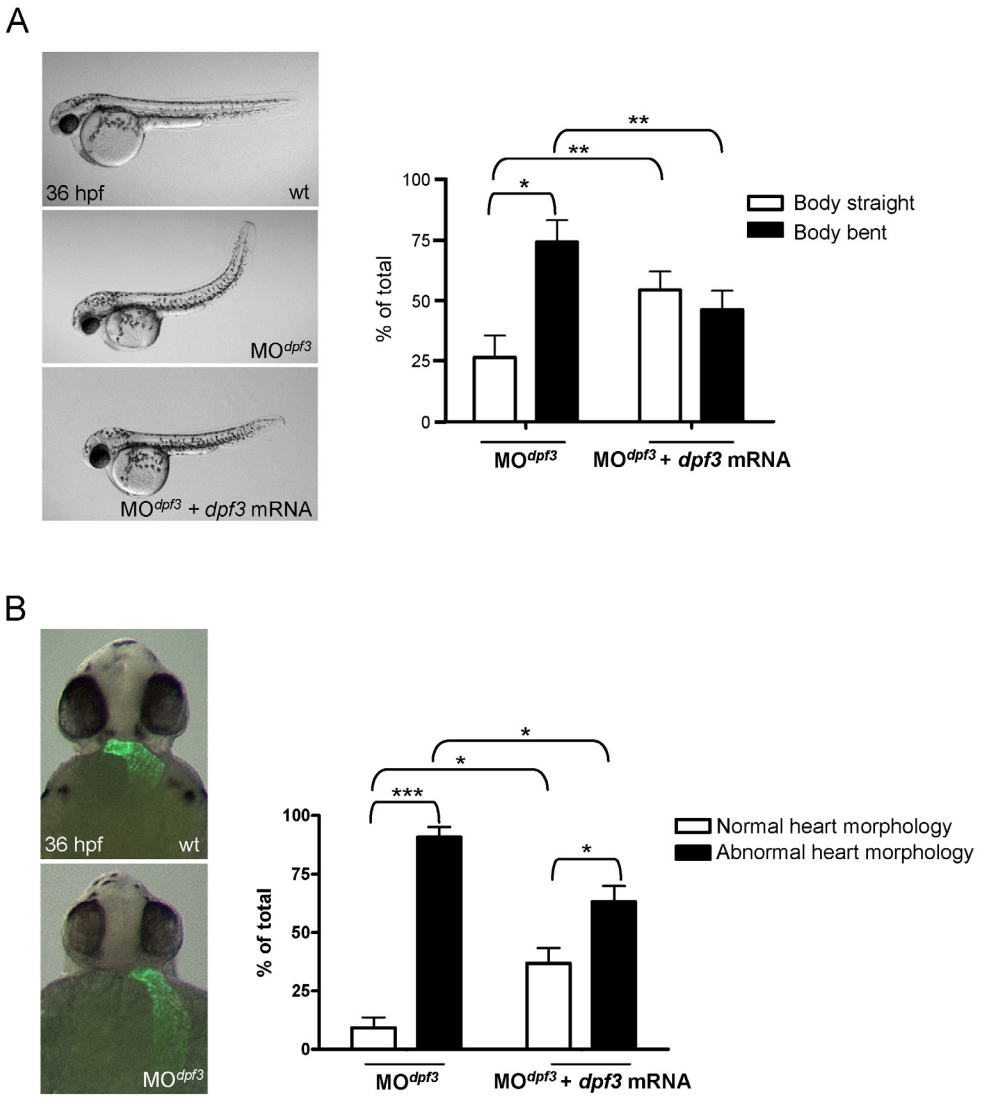

Figure 3.5: Phenotype of *dpf3* morphant zebrafish embryos. A) Lateral views of zebrafish embryos at 36 hpf. MO^{dpf3}-injection affected body posture of zebrafish embryos (n=255; p=0.05). Co-injection of mature *dpf3* mRNA significantly reduced the morphant phenotype (n= 321); MO^{dpf3} vs. MO^{dpf3} +*dpf3* mRNA, p<0.01). B) Dorsal GFP-fluorescence views of zebrafish embryos at 36 hpf. MO^{dpf3}-injection also affected heart morphology of zebrafish embryos (n=101; p<0.001). Co-injection of mature *dpf3* mRNA significantly reduced the morphant phenotype (n= 145; MO^{dpf3} vs. MO^{dpf3} +*dpf3* mRNA, p<0.05).

3.2.4 Microarray gene expression analysis of *dpf3* morphant embryos

To identify genes deregulated in *dpf3* morphants on a genome-wide level, gene expression analyses were performed using the Affymetrix GeneChip Zebrafish Genome Array. Embryos were injected with MO^{dpf3} or $MO^{control}$ at the one-cell stage, collected at 72 hpf and RNA from *dpf3* morphant embryos with severely reduced ventricular contractility and control-injected stage-matched embryos was isolated (n=30, 2 replicates). The RNA was labeled and hybridized on arrays at the German Resource Center for Genome Research (RZPD), Berlin.

The set of upregulated genes contained many genes essential for transcriptional regulation, nucleosome assembly and metabolic processes, whereas genes involved in ion and electron transport were overrepresented among downregulated genes. A subset of differentially expressed genes was confirmed by real-time PCR including genes directly involved in sarcomere assembly and muscle function (Table 3.1). Significantly increased expression of *cardiomyopathy associated 1 (cmya1*, fold change 2.9) and of *actin binding protein 280-like* (*flncb*, fold change 2.5) was observed. Furthermore, *thymosin beta* (fold change 0.3), and a novel protein (*zgc:101755*) similar to mouse *actin filament capping protein of muscle Z-lines* (fold change 0.5) were strongly downregulated.

Table 3.1: Real-time PCR confirmation of differentially expressed genes in *dpf3* morphant embryos at 72 hpf. FC, fold change; Adj. p-val. – p-value adjusted for multiple testing using Benjamini & Hochberg correction.

Gene name	Ensembl ID	Array		Real-time PCR	
		FC	Adj. p-val.	FC	p-val.
hsp70	ENSDARG00000055723	13.9	0.04	18.9	0.03
mmp9	ENSDARG00000042816	6.7	0.05	17.4	0.02
zgc:110715(troponin I)	ENSDARG00000035958	3.2	0.03	4.7	0.003
akt1s1	ENSDARG00000060779	3.1	0.01	4.1	0.01
fibronectin 1b	ENSDARG00000006526	3.0	0.05	6.7	0.01
cmya1	ENSDARG00000030722	2.5	0.02	2.9	0.01
flncb	ENSDARG00000018820	2.0	0.08	2.5	0.06
hand2	ENSDARG00000008305	0.6	0.08	0.5	0.04
zgc:101755 (CapZalpha1)	ENSDARG00000056090	0.6	0.08	0.5	0.05
calm3a	ENSDARG00000037014	0.5	0.04	0.6	0.07
pik3r3	ENSDARG00000034409	0.4	0.05	0.5	0.04
gamma-crystallin	ENSDARG00000057515	0.2	0.01	0.09	0.01
mybbp1a	ENSDARG00000028323	2.3	0.06	4.3	0.009
irx1b	ENSDARG00000056594	1.7	0.05	2.8	0.06
zgc136930	ENSDARG00000055192	0.1	0.004	0.05	0.004
irx4a	ENSDARG00000035648	0.6	0.07	0.3	0.02
zgc:64199	ENSDARG00000042027	0.6	0.06	0.5	0.0009

prmt1	ENSDARG00000010246	0.5	0.04	0.5	0.03
dlx4a	ENSDARG00000011956	0.4	0.03	0.6	0.11
thymosin-beta	ENSDARG00000054911	0.4	0.06	0.3	0.04
crabp1a	ENSDARG00000045926	0.3	0.01	0.07	0.01
tropomodulin 4	ENSDARG00000020890	0.5	0.08	0.6	0.07

3.2.5 Gene ontology analysis of differentially expressed genes from zebrafish genome-wide expression arrays

Genes differentially expressed in *dpf3* morphant embryos with an adjusted p-value <0.1 (1,210 of ~15,000 transcripts) were analyzed for overrepresentation of gene ontology (GO) terms in the ontologies biological process and cellular component (Table 3.2) (Falcon and Gentleman, 2007). GO analysis is a tool to classify a subset of genes according to their biological function or subcellular localization. The table shows odds-ratio, expected gene count (ExpCount) and actual gene count (Count) for each term tested along with the total number of transcripts associated with this term in the universe-of-transcripts (Size). The universe-of-transcripts was defined as all transcripts with an interquartile range greater or equal to 0.5 (1.760 transcripts).

Table 3.2: Gene ontology analysis of differentially expressed genes from microarray expression profiling of *dpf3* morphant embryos.

Ontology biological process of differentially upregulated genes					
GO	Odds Ratio	ExpCount	Count	Size	Term
GO:0006796	2.63	6.67	12	25	phosphate metabolic process
GO:0006544	Inf	0.80	3	3	glycine metabolic process
GO:0031564	2.07	11.47	18	43	transcription antitermination
GO:0031554	2.07	11.47	18	43	regulation of transcription termination
GO:0043624	2.07	11.47	18	43	cellular protein complex disassembly
GO:0006334	11.19	1.33	4	5	nucleosome assembly
GO:0006520	3.95	3.20	7	12	amino acid metabolic process
GO:0006082	2.84	5.34	10	20	organic acid metabolic process
GO:0006139	1.49	43.49	54	163	nucleic acid metabolic process
GO:0044275	3.23	4.00	8	15	cellular carbohydrate catabolic process
GO:0006006	3.23	4.00	8	15	glucose metabolic process
GO:0006468	3.23	4.00	8	15	protein amino acid phosphorylation
GO:0019320	3.28	3.47	7	13	hexose catabolic process
GO:0043283	1.45	39.76	49	149	biopolymer metabolic process
GO:0015980	2.82	4.27	8	16	oxidation of organic compounds
GO:0006807	2.54	5.07	9	19	nitrogen compound metabolic process
GO:0046164	2.81	3.74	7	14	alcohol catabolic process
GO:0043412	1.85	10.14	15	38	biopolymer modification
GO:0009628	8.35	1.07	3	4	response to abiotic stimulus

GO:0006333	3.50	2.40	5	9	chromatin assembly or disassembly
GO:0005975	1.80	9.61	14	36	carbohydrate metabolic process
GO:0006096	2.80	3.20	6	12	glycolysis
GO:0007001	2.80	3.20	6	12	chromosome organization (sensu Eukaryota)
GO:0009408	Inf	0.53	2	2	response to heat
GO:0006003	Inf	0.53	2	2	fructose 2,6-bisphosphate metabolic process
GO:0006541	Inf	0.53	2	2	glutamine metabolic process
GO:0019882	Inf	0.53	2	2	antigen processing and presentation
GO:0051091	Inf	0.53	2	2	positive regulation of transcription factor activity
GO:0051056	Inf	0.53	2	2	of small GTPase mediated signal transduction
GO:0019318	Inf	0.53	2	2	hexose metabolic process
GO:0048592	3.72	1.87	4	7	eye morphogenesis
GO:0050794	1.33	38.16	45	143	regulation of cellular process
GO:0009057	1.78	8.27	12	31	macromolecule catabolic process
GO:0007049	2.79	2.67	5	10	cell cycle
GO:0006396	2.79	2.67	5	10	RNA processing
GO:0006351	1.38	28.02	34	105	transcription, DNA-dependent

Ontology biological process of differentially downregulated genes

GO	Odds Ratio	ExpCount	Count	Size	Term
GO:0042440	5.09	3.71	7	9	pigment metabolic process
GO:0006783	5.09	3.71	7	9	heme biosynthetic process
GO:0006778	3.38	4.12	7	10	porphyrin metabolic process
GO:0006879	2.62	5.77	9	14	iron ion homeostasis
GO:0006826	2.62	5.77	9	14	iron ion transport
GO:0006118	2.62	5.77	9	14	electron transport
GO:0051188	2.29	7.41	11	18	cofactor biosynthetic process
GO:0019725	2.42	6.59	10	16	cell homeostasis

Ontology cellular component of differentially upregulated genes

GO	Odds Ratio	ExpCount	Count	Size	Term
GO:0005634	1.63	32.76	43	125	nucleus
GO:0000786	Inf	0.79	3	3	nucleosome
GO:0005960	Inf	0.52	2	2	glycine cleavage complex
GO:0031301	Inf	0.52	2	2	integral to organelle membrane
GO:0043227	1.38	40.36	48	154	membrane-bound organelle

Ontology cellular component of differentially downregulated genes

GO	Odds Ratio	ExpCount	Count	Size	Term
GO:0005829	12.40	4.28	9	10	cytosol
GO:0031967	3.64	4.70	8	11	organelle envelope
GO:0005739	2.74	6.41	10	15	mitochondrion
GO:0005737	1.47	34.20	41	80	cytoplasm
GO:0031966	3.18	4.28	7	10	mitochondrial membrane

3.3 DPF3 is a novel component of the BAF chromatin remodeling complex

The plant-homeodomains of DPF3 are frequently found in nuclear proteins whose substrate tend to be nucleosomes. To isolate potential nuclear binding partners of human DPF3a and DPF3b, tandem affinity purification (TAP) combined with mass spectrometry was performed in HEK293T cells in ccoperation with Martje Tönjes within the workgroup and Dr. Johan Gobom at the mass spectrometry unit of the Max Planck Institute for Molecular Genetics. Among others proteins, nearly all core components of the BAF chromatin remodeling complex were found to be associated with both isoforms of DPF3 (Table 3.3). The threshold set for the identification of interaction partners in the mass spectrometry analysis was a Mascot Score of > 50 and a minimum number of two Spectral counts. The simple calculation of protein abundance by counting the obtained spectra is biased by protein size, as larger proteins generate more peptides and thus more counts. To estimate the abundance of the individual identified proteins irrespective of protein size, and in relation to overexpressed DPF3a and DPF3b, the normalized spectral abundance factor (NSAF) was calculated. The NSAF is calculated by first dividing the spectral count by the length of the protein, resulting in the spectral abundance factor (SAF) (Powell *et al.*, 2004) The individual SAF values are then normalized by dividing by the sum of all SAFs for proteins in the complex, which results in the NSAF (Zybailov *et al.*, 2006). This value is useful for the comparison of protein abundance in independent experiments, such as those for DPF3a and DPF3b.

Table 3.3: Peptides associated with human DPF3a and DPF3b identified by a combination of tandem affinity purification (TAP) and mass spectrometry. MW, calculated molecular weight; NSAF, normalized spectral abundance factor.

Bait	Name	Accession number	Length (aa)	MW (kDa)	Mascot score	Spectral counts	Sequence coverage (%)	NSAF
DPF3a	BAF250A	O14497	2285	242.0	1844	47	24	0.038
	BAF250B	Q8NFD5	2236	236.1	1171	33	14	0.027
	BRG1	P51532	1647	184.6	2196	51	29	0.057
	BRM	P51531	1586	180.8	959	27	16	0.031
	BAF170	Q8TAQ2	1214	132.9	2160	48	30	0.073
	BAF155	Q92922	1105	122.8	2401	49	38	0.081
	BAF60C	Q6STE5	483	55.0	644	19	36	0.072
	BAF60A	Q96GM5	476	54.9	80	4	11	0.014
	BAF60B	Q92925	456	52.3	846	23	44	0.093
	BAF53	O96019	429	47.5	826	17	39	0.073
	BAF57	Q969G3	411	46.6	982	19	49	0.085
	BAF47	Q12824	385	44.1	659	14	39	0.067
	Beta-actin	P60709	375	41.7	697	18	38	0.088
	DPF3	Q92784	224	25.8	844	15	63	0.123
DPF3b	BAF250A	O14497	2285	242.0	1461	49	27	0.036
	BAF250B	Q8NFD5	2236	236.1	457	16	9	0.012
	BRG1	P51532	1647	184.6	1912	43	22	0.044
	BRM	P51531	1586	180.8	792	21	12	0.022
	BAF170	Q8TAQ2	1214	132.9	1676	42	28	0.058
	BAF155	Q92922	1105	122.8	1511	42	29	0.064
	BAF60C	Q6STE5	483	55.0	616	19	37	0.066
	BAF60A	Q96GM5	476	54.9	1062	26	45	0.091
	BAF60B	Q92925	456	52.3	957	21	43	0.077
	BAF53	O96019	429	47.5	699	15	35	0.059
	BAF57	Q969G3	411	46.6	969	19	44	0.077
	BAF47	Q12824	385	44.1	880	18	56	0.078
	Beta-actin	P60709	375	41.7	604	18	43	0.080
	DPF3	Q92784	224	25.8	909	16	63	0.120

3.3.1 Interaction of DPF3 proteins with components of the BAF complex

To assess the results of the affinity purification and mass spectrometry, protein-protein interactions were tested by GST-pulldown assays. Recombinant GST-DPF3a and GST-DPF3b fusion proteins were used together with nuclear protein extracts of HEK293T cells to test for the ability of DPF3 to pull down proteins of the BAF complex. Immunodetection of the complex was carried out using an antibody against SMARCA4 (BRG1), a core component of the BAF chromatin remodeling complex. Both isoforms of DPF3 ware able to pull down BRG1 from nuclear protein extracts (Figure 3.6A).

Subsequently, experiments were performed to test whether the interaction of DPF3a and DPF3b with the BAF complex is mediated through Baf60c, the heart and somite specific subunit of the complex. Again, recombinant GST-DPF3a or GST-DPF3b fusion proteins were used together with HEK293T nuclear protein extracts from cells previously transfected with Flag-BAf60c expression vector. Immunodetection using an anti-Flag antibody showed that both DPF3 isoforms directly interact with Baf60c (Figure 3.6B). To test the specificity of the interaction, reciprocal pulldown assays were performed using GST-Baf60c and HEK293T

nuclear protein extracts from cells previously transfected with Flag-DPF3a and DPF3b expression vector, respectively (Figure 3.6C). These experiments confirmed the interactions of DPF3a and DPF3b with Baf60c.

Thus, both DPF3 isoforms associate with the BAF chromatin remodeling complex through direct interaction with its subunit Smarcd3.

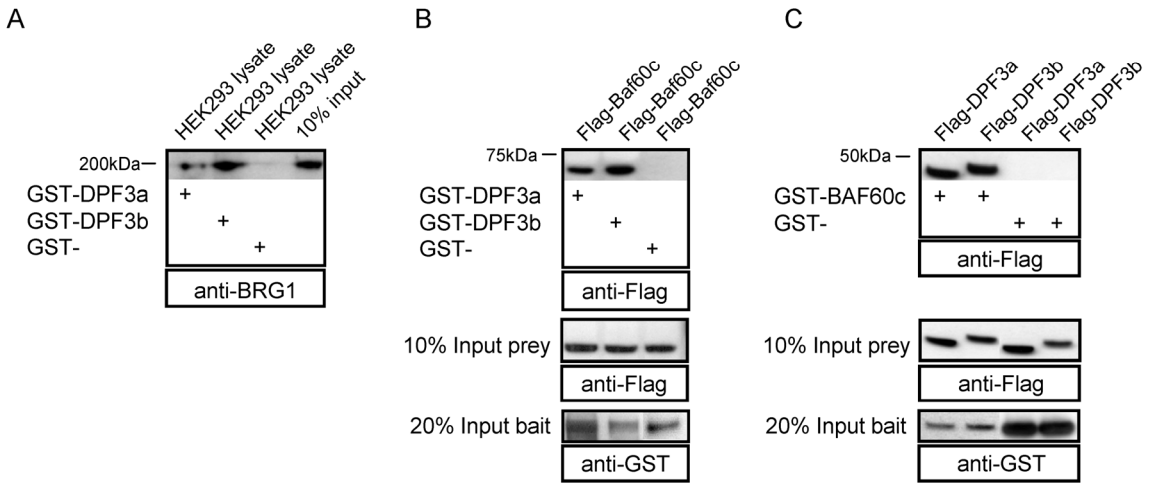

Figure 3.6: DPF3 associates with the BAF chromatin remodeling complex through direct interaction with SMARCD3. A) GST-pulldown assay using recombinant GST-DPF3a/b and nuclear protein extract from HEK293T cells followed by immunodetection of SMARCA4 (BRG1), a core component of the BAF complex. B) GST-pulldown assay using recombinant GST-DPF3a and DPF3b and nuclear protein extract of cells overexpressing Flag-Smarcd3 (Baf60c), followed by immunodetection of Flag-Smarcd3. C) GST-pulldown assay using recombinant GST-Baf60c and nuclear protein extract of cells overexpressing Flag-DPF3a and DPF3b, followed by immunodetection of Flag-Smarcd3.

3.4 The PHD fingers of DPF3b interact with modified histone tails

It has recently become evident that proteins involved in chromatin remodeling recognize specific modifications on histone tails. The recognition of the methylation state of lysine residues on histone 3 and histone 4 has been shown to be mediated, among others, by the plant-homeodomain. On the other hand, lysine acetylation is recognized by the bromodomain (see Chapter 1.4).

To address whether DPF3 binds to histones, a GST-pulldown system was used to test for the ability of recombinant full-length GST-DPF3a and GST-DPF3b to pull down histones from calf thymus extracts followed by western analyses using histone specific antibodies against H2A, H2B, H3 and H4.

DPF3b was able to pull down histones H3 and H4, but not histones H2A and H2B. DPF3a did not bind any histones in this assay (Figure 3.7A).

To analyze if DPF3b recognizes histones and moreover specific modifications through its PHD fingers, a broad panel of histone 3 and 4 peptides harboring specific modifications such

as methylation, acetylation or phosphorylation on different residues was tested in pulldown assays (Figure 3.7B).

The double PHD finger of DPF3b (GST-DPF3b-PHD1-PHD2) interacted specifically with mono- and di-methylated lysine 4 on histone 3 (H3K4me1, H3K4me2), whereas unmodified (H3, aa1-21) and tri-methylated lysine 4 (H3K4me3) were not detected. In addition to the methyllysine recognition, GST-DPF3b-PHD1-PHD2 also recognized acetylated lysines on histone 3 (H3K14ac, H3K9ac) and histone 4 (H4K5ac, H4K8ac, H4K12ac, H4K16ac).

Again, DPF3a, which only contains a truncated PHD finger, did not bind any of the studied peptides supporting the previous results using calf thymus histone extracts.

Since DPF3b contains a double PHD finger, the binding properties of the plant-homeodomain 1 and 2 alone were tested to see whether single PHD fingers are able to recognize histone modifications. Pulldown assays revealed that single DPF3 PHD fingers were sufficient for the interaction with lysine acetylations on histone 4 (H4K5ac, H4K8ac, H4K12ac, H4K16ac), whereas histone 3 acetylations and methylations were only recognized by the double PHD finger (Figure 3.7B).

It has been shown that interactions between the PHD finger and histone lysine methylations are mediated by the formation of a cage of aromatic amino acids within the PHD finger, into which the histone-methyllysine projects and becomes stabilized through interactions with acidic residues surrounding the cage across the length of the peptide (Ruthenburg *et al.*, 2007a). Moreover, the strucural integrity of the PHD finger is mediated by conserved cysteine residues which complex zinc ions in its center. The get to the structural basis of the DPF3-PHD-histone interactions, site-directed mutagenesis was performed on a conserved tryptophan residue within PHD1 (W311E) and a tandem of cysteines within PHD2 (C360R/C363R), respectively. The unpolar, hydorphobic tryptophan was replaced by a polar, acidic glutamic acid while the neutral cysteines were replaced by basic arginines. These mutant proteins were then tested for their histone peptide binding properties (Figure 3.7B).

GST-pulldown assays using the mutated form GST-DPF3b-PHD1-PHD2 (W311E) showed that histone 3 modifications were not recognized anymore, while histone 4 modifications were still detected (Figure 3.7B). This result was in good agreement with results comparing the tandem PHD finger binding properties to the single PHD fingers, which showed that both domains are sufficient and necessary for the histone 3 interaction. The same result was obtained for GST-DPF3b-PHD1-PHD2 (C360/C363R), again showing that an intact double PHD finger is necessary for histone 3 interactions. To analyze the effect of the point mutations on single DPF3-PHD fingers, the same mutations were introduced into the GST-

DPF3b-PHD1 and GST-DPF3b-PHD2 constructs. Pulldown assays revealed that histone 4 binding was completely abolished in the GST-DPF3b-PHD1 (W311E) and GST-DPF3b-PHD2 (C360R/C363R) mutants. This showed that single PHD fingers 1 and 2 are sufficient and necessary for histone 4 acetyl-recognition. In line with these data, introduction of all point mutations in the double PHD finger construct lead to complete abolishment of histone recognition.

To test the specificity of the binding reactions, comparative pulldown assays were performed using the PHD finger of BPTF and the bromodomain of BRG1, which have been shown the recognize H3K4me2/me3 and multiple acetlyated lysines, respectively (Li *et al.*, 2006a; Shen *et al.*, 2007).

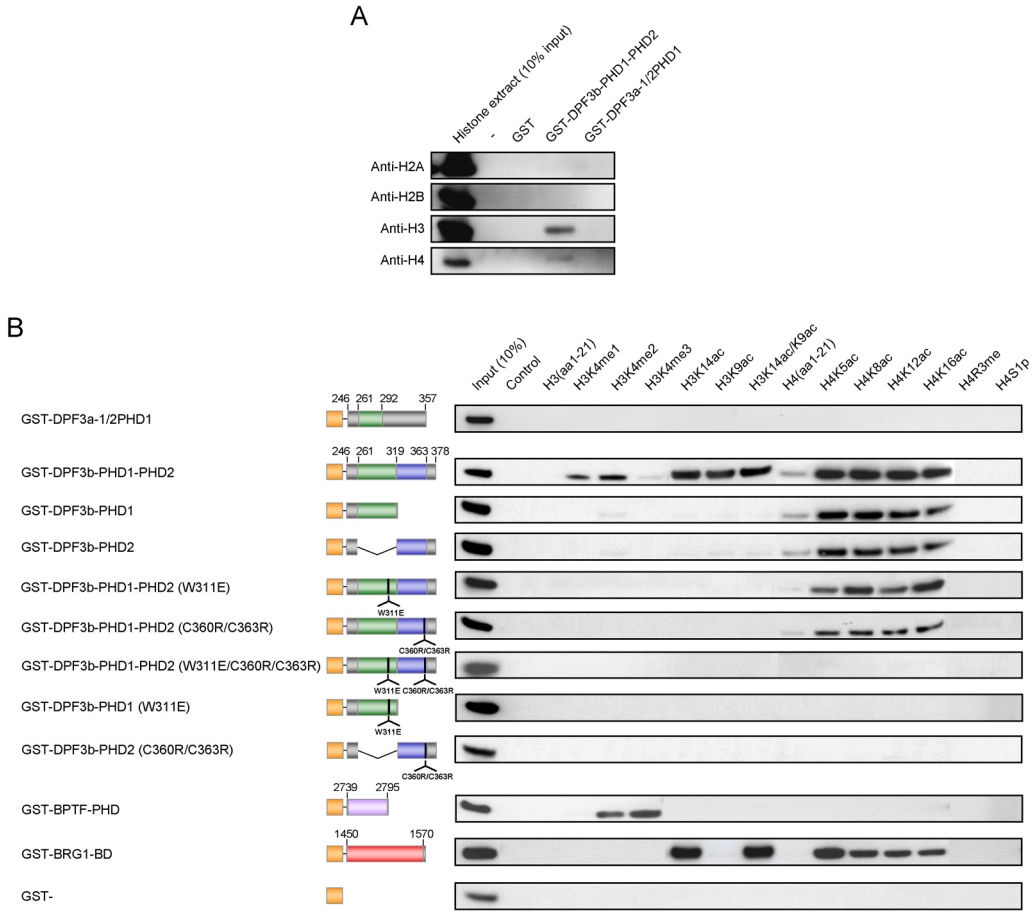

Figure 3.7: The plant-homeodomains (PHD) of DPF3 interact with modified histone tail residues. A) GST-pulldown experiments using recombinant GST-DPF3a-1/2PHD1 and GST-DPF3b-PHD1-PHD2 in combination with calf thymus histone extract. B,C) GST-pulldown experiments using recombinant GST-DPF3b-PHD1-PHD2 and truncated versions, GST- or GST-DPF3a-1/2PHD1 and mutated constructs in combination with modified or unmodified histone peptides followed by immunodetection of GST. Binding properties were compared to known histone methyllysine- and acetyllysine-binders BPTF-PHD and BRG1-bromodomain (BD).

In summary, an intact double PHD finger was sufficient and necessary for histone 3 methyl- and acetyl-interactions. Moreover, histone 4 recognition was still achieved by single PHD

fingers, suggesting a different mode of recognition for H3 and H4, respectively. Further support for distinct binding pockets comes from the mutation analysis of tryptophan 358, which is essential for H3 interactions in the GST-DPF3b-PHD1-PHD2 (W358E) construct, but not for H4 interactions of the DPF3b-PHD2 (W358E) construct.

3.5 The role of DPF3 in developmental signaling pathways

3.5.1 *DPF3* promoter analysis

In order to characterize the mechanisms and pathways controlling DPF3 gene expression, the human *DPF3* promoter region was analyzed for the presence of evolutionary conserved elements and potential transcription factor binding sites. Given the evolutionary conserved gene expression pattern of *Dpf3*, it was assumed that regulatory elements are well conserved among species. As a first step, the evolutionary conserved promoter regions among mammals were identified by preparing VISTA plot alignments of human, mouse and opossum *DPF3* promoter sequences. VISTA is a set of programs used to compare DNA sequences from two or more species and visualize the alignments with annotation information (http://genome.lbl.gov/vista/servers.shtml) (Mayor *et al.*, 2000). Two promoter regions of high evolutionary conservation were identified comparing human and mouse sequences, which became even more obvious when comparing the more distinct species human and opossum. The region directly adjacent to the transcription start site was called "core promoter", while a region further upstream of the transcription start site was called "enhancer" (Figure 3.8).

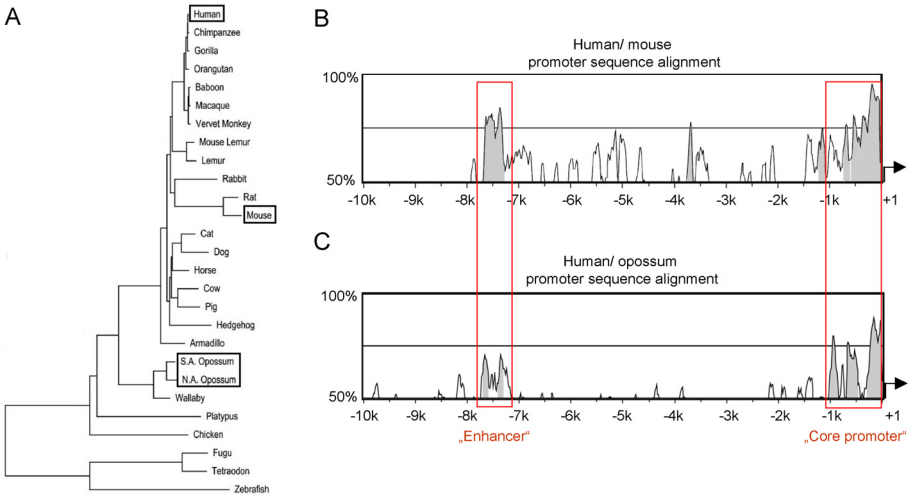

Figure 3.8: A) Phylogenetic tree of vertebrate species. The last common ancestor between primates (human) and rodents (mouse) is estimated at 75 mya (million years ago) and between eutherians and metatherians (opossum) at 185 mya. B, C) Comparative sequence alignments between human/mouse

and human/opossum *DPF3* promoter region, respectively, depicting evolutionary conserved regions (labeled in red). The y-axis depicts sequence identity in %; the x-axis shows the distance from the transcription start site. Figure A modified from (Margulies *et al.*, 2005).

The *DPF3* core promoter was analyzed regarding its potential to active transcription in a reporter gene assay.

In a reportergene assay, a potential regulatory element is tested for its ability to active transcription by fusing it to a reporter gene such as luciferase, which can be measured easily by transfection of mammalian cells. The complete region and 5´ truncations of the core promoter were tested to identify the minimal region sufficient for transcriptional activation (Figure 3.9).

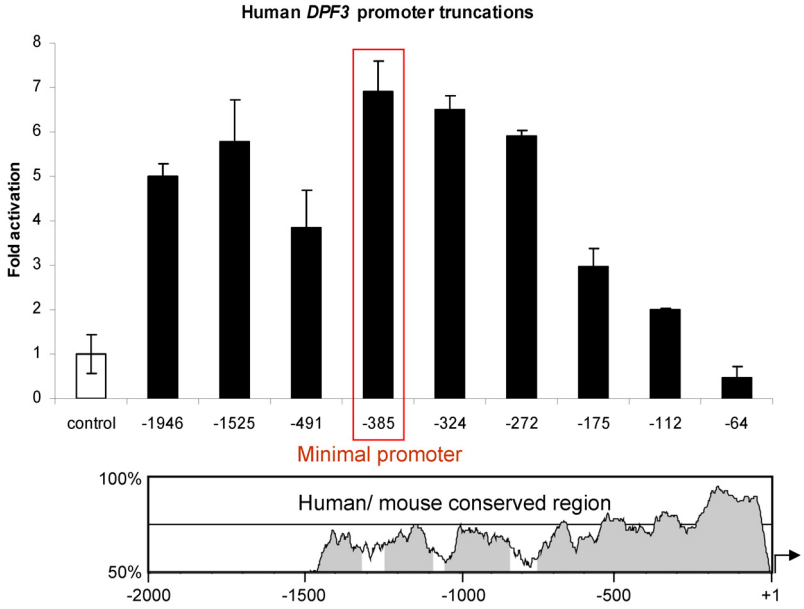

Figure 3.9: Transcriptional activity of the human *DPF3* core promoter. Graph depicts results from reportergene assays of various fragments of the *DPF3* core promoter fused to luciferase and tested for transcriptional activity in the human cell line HEK293T. Fold activation is calculated by the ratio of reporter luciferase and internal *Renilla* control luciferase relative to empty reporter vector. VISTA plot below shows the degree of sequence conservation between human and mouse (non-scaled).

The results established that a minimal region of ~175 bp is sufficient for transcriptional activation of the reporter gene, as luciferase activity drops to background levels after further truncation of the -175bp construct and is completely lost in the -64 bp construct. The highest transcriptional activity was obtained with the -385 bp construct (the sequence is located at chr14:72,430,563-72,430,943 of the Human Mar. 2006 Assembly). Consequently, this promoter region was called *DPF3* "minimal promoter" and analyzed in more detail regarding potential transcription factor binding sites (TFBS) using the prediction tool TRANSFAC Professional 11.4 (Kel *et al.*, 2003).

In line with *Dpf3* mRNA expression data in mouse and zebrafish embryos, putative binding sites for transcription factors implicated in regulating heart and skeletal muscle development were identified (Figure 3.10). Among these were evolutionary conserved binding sites for Mef2, Nkx2.5, Serum Response Factor (SRF), COUP-TF/RXR-α, T-Box factors and transcription factors of the bHLH group, which recognize E-box elements.

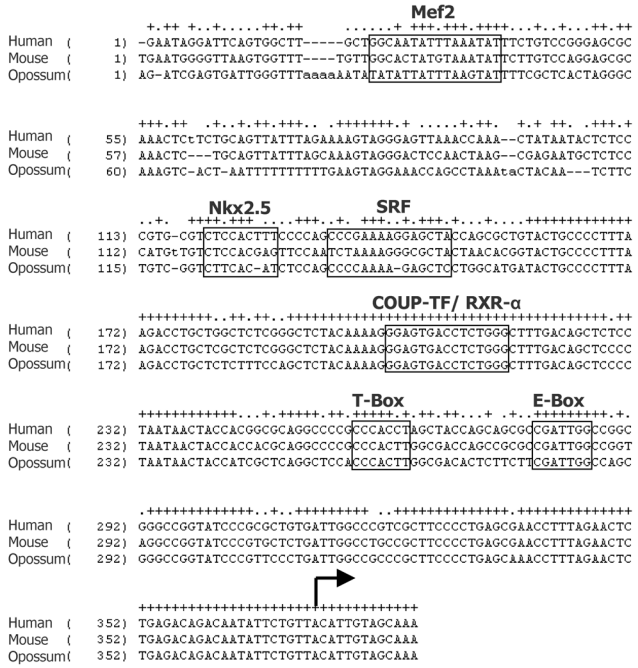

Figure 3.10: Transcription factor binding site (TFBS) prediction of the *DPF3* minimal promoter region using TRANSFAC Professional 11.4. Potential TFBSs are indicated relative to the position of the transcription start site (arrow).

The functional relevance of the predicted TFBSs were tested *in vitro* and *in vivo*. First, reportergene assays were carried out using the *DPF3* minimal promoter (-385 bp). The wildtype construct was tested in cell culture alone or in co-transfection experiments with a broad panel of potentially relevant transcription factors (TFs). Apart from TFs with a clearly identified TFBS in the *DPF3* minimal promoter, additional TFs known to play a role in heart and skeletal muscle development were chosen, as TFBS are sometimes poorly defined or even unknown.

Furthermore, signaling molecules of the BMP and Wnt pathways were tested for their ability to regulate *DPF3* expression, as they have also been shown to be important for heart and skeletal muscle development (Figure 3.11).

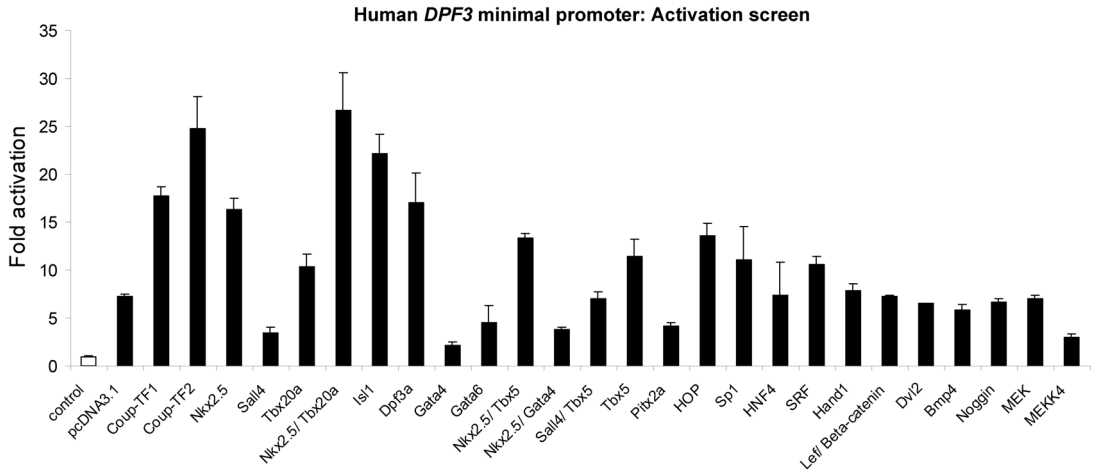

Figure 3.11: Luciferase reportergene assay experiments using the *DPF3* minimal promoter in HEK293T cells. The reporter construct was co-transfected with expression vectors for the transcription factors indicated on the x-axis.

If a transcription factor showed potential to regulate the reporter, site-directed mutagenesis was performed to delete the predicted TFBS. Furthermore, electromobility shift assays were performed to test for the ability of the transcription factor to bind to the according DNA sequence. Summarized below are the transcription factors which could be confirmed to activate and bind to the *DPF3* promoter by luciferase assays combined with site-directed mutagenesis of binding sites, electromobilityshift assays or chromatin immunoprecipitation.

COUP-TFs

Among the factors tested were COUP-TF1 and COUP-TF2. The COUP-TFs belong to the family of nuclear receptors and bind to the consensus DNA sequence TGACCYNTGACCY (Calonge *et al.*, 2004). The role of COUP-TFs in heart development has been demonstrated by genetic inactivation in mouse embryos (Guo *et al.*, 2001; Pereira *et al.*, 1999). Co-transfection of the reporter construct and increasing amounts of COUP-TF2 expression vector in HEK293T cells revealed a dose-dependent transcriptional activation by COUP-TF2 (Figure 3.12A). In line with this, deletion of the potential COUP-TF binding site (GTGACCTCTGGG) showed that transcriptional activity was indeed mediated through the analyzed TFBS, as activation of the construct was lost in the mutated construct. Both COUP-TF1 and COUP-TF2 were able to activate the reporter to a comparable degree (Figure 3.12B). Furthermore, an electromobility shift assay was performed using a labeled oligonucleotide containing the COUP-TF binding site and HEK293T nuclear cell extract previously transfected with COUP-TF1 or COUP-TF2 expression vectors. The specificity of the binding reaction is shown by addition of 100 fold excess unlabeled oligonucleotides, which abolishes

binding of labeled oligonucleotides (Figure 3.12C). These experiments confirmed that COUP-TFs were capable of binding to their target sequence within the *DPF3* promoter *in vitro*.

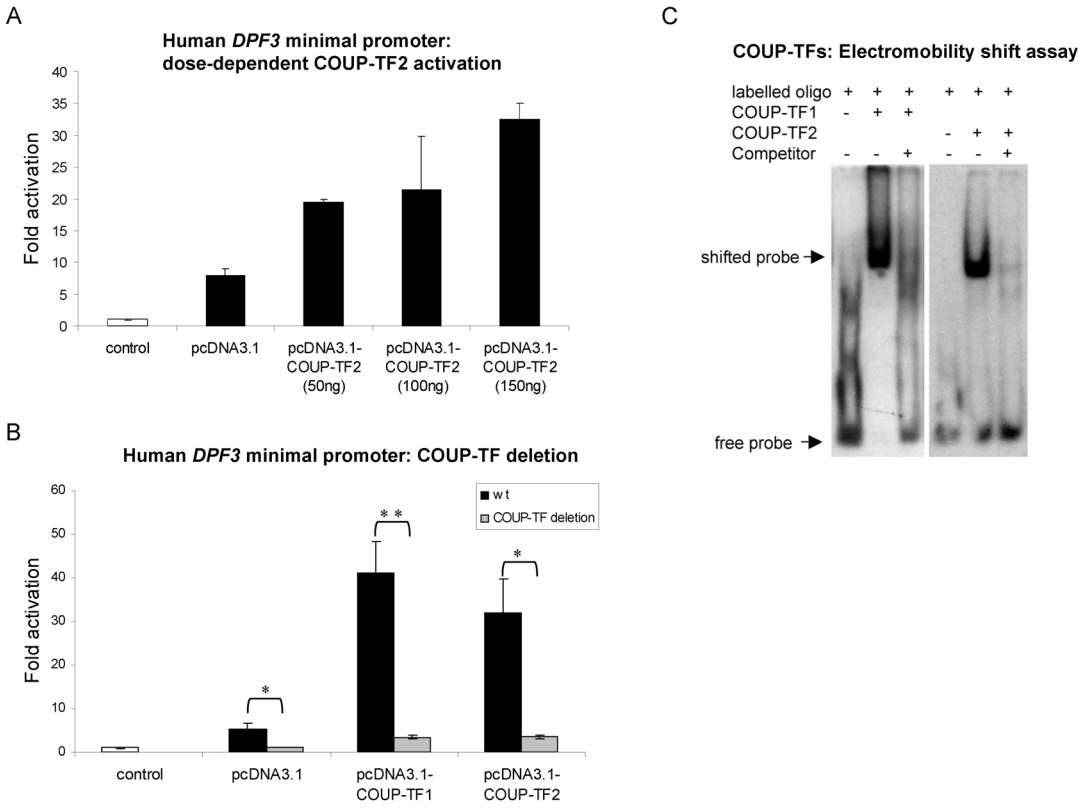

Figure 3.12: COUP-TFs activate the human *DPF3* minimal promoter. A) Luciferase reportergene assay experiments in HEK293T cells. Co-transfection of increasing amounts of COUP-TF2 expression vector and a fixed concentration of the *DPF3* minimal promoter showed a dose-dependent activation by COUP-TF2. B) Deletion of the COUP-TF binding site abolished activation by COUP-TF1 and COUP-TF2 (COUP-TF pcDNA wt vs. mut: p-val 0.0232; COUP-TF1 wt vs. mut: p-val 0.009; COUP-TF2 wt vs. mut: p-val 0.018). C) Electromobility shift assay using a digoxygenin-labeled oligonucleotide containing the COUP-TF binding site and HEK293T nuclear cell extract previously transfected with COUP-TF1 or COUP-TF2 expression vectors.

Nkx2.5

Another factor tested was Nkx2.5. This transcription factor sits up high in the genetic program controlling heart development. Gene inactivation studies in mouse have shown that Nkx2.5 is necessary for chamber myocardium formation and maintenance of conduction system function (Lyons *et al.*, 1995; Prall *et al.*, 2007).

Co-transfection of the reporter construct and increasing amounts of Nkx2.5 expression vector in HEK293T cells revealed a dose-dependent transcriptional activation by Nkx2.5 (Figure 3.13A). In line with this, deletion of the potential *Nkx2.5* binding element (NKE), (TCCACTTTCC) showed that transcriptional activity was indeed mediated through the analyzed TFBS, as activation of the construct was lost in the mutated construct (Figure 3.13B).

Furthermore, an electromobility shift assay was performed using a labeled oligonucleotide containing the NKE and HEK293T nuclear cell extract previously transfected with Nkx2.5 expression vectors (Figure 3.13C). These experiments confirmed that Nkx2.5 is able to bind to its target sequence within the *DPF3* promoter *in vitro*.

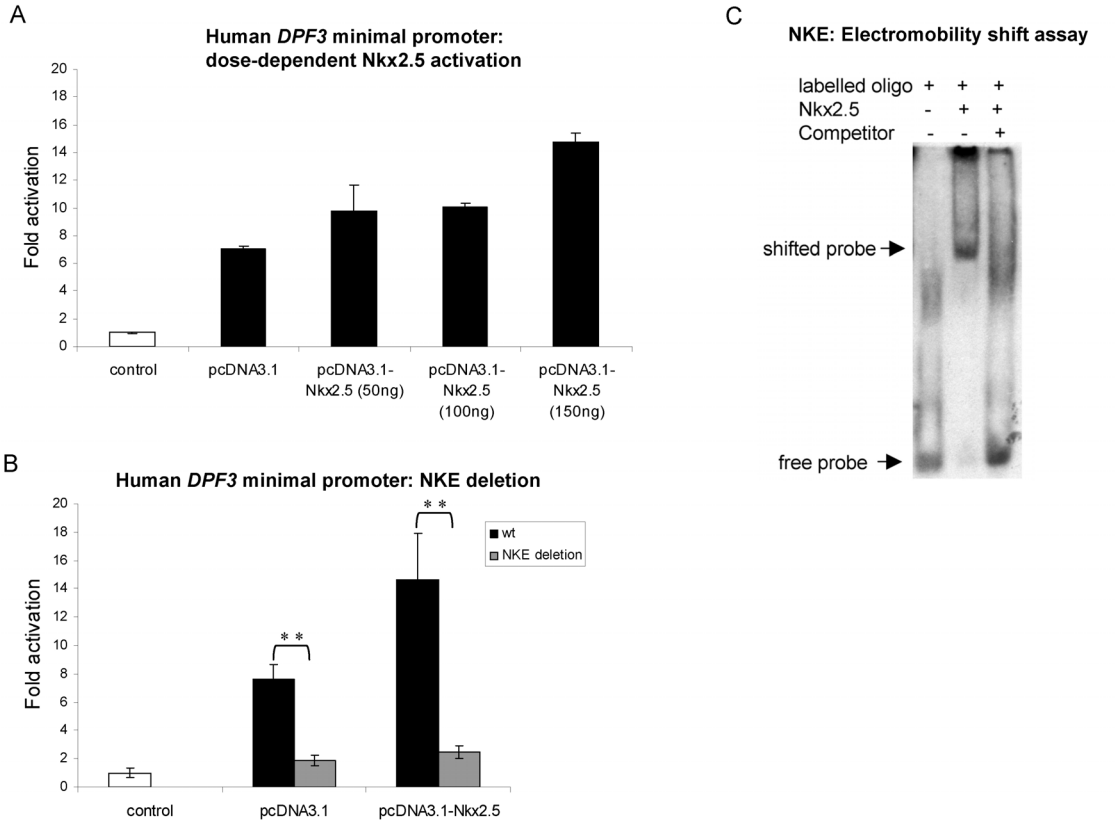

Figure 3.13: Nkx2.5 activates the human *DPF3* minimal promoter. A) Luciferase reportergene assay experiments in HEK293T cells. Co-transfection of increasing amounts of Nkx2.5 expression vector and a fixed concentration of the *DPF3* minimal promoter showed a dose-dependent activation by Nkx2.5. B) Deletion of the NKE abolished activation by Nkx2.5 (pcDNA3.1 wt vs. mut: p-val 0.0381; pcDNA3.1-Nkx2.5 wt vs. mut: p-val 0.001). C) Electromobility shift assay using a digoxygenin-labeled oligonucleotide containing the NKE and HEK293T nuclear cell extract previously transfected with Nkx2.5 expression vector.

Sall4

In situ hybridization experiments in mouse embryos had previously shown that *Dpf3* mRNA expression is reduced in the developing interventricular septum of the mouse heart (see Chapter 1.1). Interestingly, the Spalt-family zinc-finger transcription factor Sall4 has been shown to be expressed in this region of the heart and to function as a transcriptional repressor (Koshiba-Takeuchi *et al.*, 2006). Furthermore, *SALL4* mutations cause Okihiro syndrome, which is characterized by absent abducens neurons, congenital deafness, renal abnormalities, anal stenosis and CHDs, especially ventricular septal defects (VSDs) (Kohlhase *et al.*, 2002).

Hence, the effect of Sall4 overexpression in HEK293T cells on the *DPF3* minimal promoter was tested in a reportergene assay (Figure 3.14). The results showed that Sall4 overexpression reduces the *DPF3* minimal promoter activity by 56 %, indicating that Sall4 might by involved in regulating the regional specificity of *Dpf3* mRNA expression in the developing heart. The Sall4 DNA binding motif is currently uncharacterized, hampering further experiments to analyze the specific binding properties of Sall4.

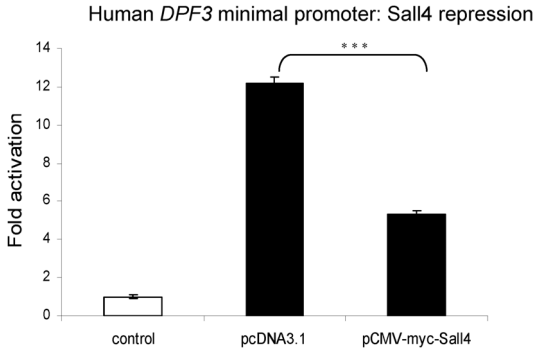

Figure 3.14: Sall4 acts as a repressor of the human *DPF3* minimal promoter. Luciferase reportergene assay experiments in HEK293T cells. Co-transfection with Sall4 expression vector and the *DPF3* minimal promoter showed a repression by Sall4 (pcDNA3.1 vs. pCMV-Sall4: p-val 2.144E-06).

Mef2a

Myocyte enhancer factor 2a (Mef2a) belongs to the MADS-family of transcription factors and plays multiple roles in development and disease (McKinsey *et al.*, 2002). Mef2a deficient mice and zebrafish embryos are phenotypically similar to the observed myofibrillar disarray in *dpf3* knockdown embryos ((Naya *et al.*, 2002; Potthoff *et al.*, 2007; Wang *et al.*, 2005) (accompanied manuscript)). Consequently, the *DPF3* core promoter was screened for potential Mef2 binding sites. Within a conserved 1.2 kbp promoter region three *Mef2* matrices were found using TRANSFAC MATCH with stringent settings (Kel *et al.*, 2003) (Figure 3.15A). Transcriptional regulation of the *Dpf3* gene by Mef2a was also tested in luciferase reportergene assays: The previously characterized human *DPF3* minimal promoter (Figure 3.9), which contains a putative Mef2 binding site (Mef2.3) was tested either alone or in combination with 4 consecutive repeats of additional putative Mef2 binding sites (Mef2.1 and Mef2.2) located upstream of the core promoter and in the ChIP-enriched region (Figure 3.15A). Co-transfections in HEK293T cells revealed an endogenous activation of the core promoter alone, which was additionally enhanced by Mef2a overexpression. Fusion constructs of the core promoter and the Mef2.1 and Mef2.2 sites showed that transcriptional activity was additionally enhanced only by the Mef2.1 site, supporting a role for Mef2a as a regulator of

Dpf3 through combinatorial effects on the Mef2.1 and Mef2.3 sites (Figure 3.15B). Chromatin immunoprecipitation against Mef2a combined with chip detection (ChIP-chip) in mouse cardiomyocytes (HL-1 cells) showed a significant peak of Mef2a binding in the *Dpf3* promoter region that could also be confirmed by real-time PCR (1.8 fold change) (Figure 3.15A). Knockdown of Mef2a in HL-1 cells using two different siRNAs led to a reduction of *Dpf3* expression of up to 40%, demonstrating that Mef2a functionally binds the *Dpf3* promoter and activates its expression (Figure 3.15C, Figure 3.15D).

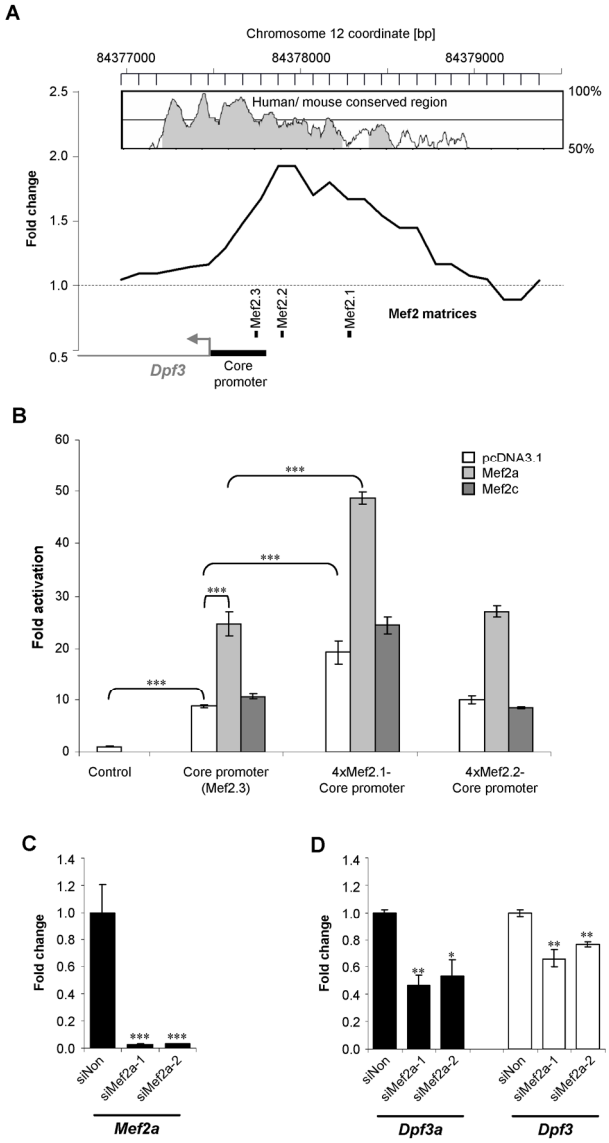

Figure 3.15: Mef2a regulates *Dpf3* expression. A) VISTA plot showing the human/mouse conserved regions of the *Dpf3* promoter. Chromatin immunoprecipitation followed by chip analysis shows binding of Mef2a to an evolutionary conserved region of the *Dpf3* promoter *in vivo*. Mef2a matrices obtained by TRANSFAC MATCH are indicated. Conservation of promoter sequence is shown. B) Luciferase Reportergene assay using the *DPF3* core promoter containing the Mef2.3 site alone or in combination with 4 repeats of the conserved, putative Mef2 binding sites (Mef2.1, Mef2.2) fused to luciferase. Activity of the reporter was measured alone or in co-transfections with Mef2a/Mef2c expression vectors in HEK293T cells. p-value<0.05 (*), p-value<0.01 (**), p-value<0.001 (***). C) Knockdown of Mef2a in HL-1 cells using two different siRNAs. Knockdown efficiency was analyzed by real-time PCR. D) Knockdown of Mef2a in HL-1 cells led to reduced expression of *Dpf3* and *Dpf3a*.

3.5.2 *DPF3* minimal promoter analysis in transgenic mice

In order to analyze the spatio-temporal expression pattern regulated by the *DPF3* minimal promoter *in vivo*, transgenic mice containing the human *DPF3* minimal promoter fused to beta-galactosidase were generated by injection of transgene DNA into the pronucleus and subsequent retransfer in pseudopregnant foster mothers. Offspring of these mice was genotyped and used as founders of transgenic mouse lines. In case a male founder was obtained, it was mated with female Bl6 mice and mating was controlled by plug-checks in the morning. Pregnant females were sacrificed during gestation, embryos extracted and analyzed for the presence of transgene expression by X-Gal staining. The transgenic constructs were either wildtype or contained the Nkx2.5 or COUP-TF binding site deletion.
For the wildtype *DPF3* minimal promoter, two stable transgenic lines (Line1001 and 1004) and one litter of transient transgenic mouse embryos were obtained. Lines 1001 and 1004 have been analyzed thoroughly at all stages of embryonic development.

DPF3 minimal promoter: Line 1001

Embryos derived from line 1001 expressed beta-galactosidase under the control of the *DPF3* minimal promoter at early stages of heart development. Expression was first observed in the anterior portion of the heart tube at E8.5 (Figure 3.16A). At E9.0, expression was observed in the future right ventricle of the looping heart (Figure 3.16B-C´). From stage E10.0 onwards, cells expressing the transgene became gradually restricted to the developing interventricular septum of the heart. In addition, cells of the right and left atrium also expressed the transgene (Figure 3.16D-O). Expression was restricted to the myocardium of the heart.

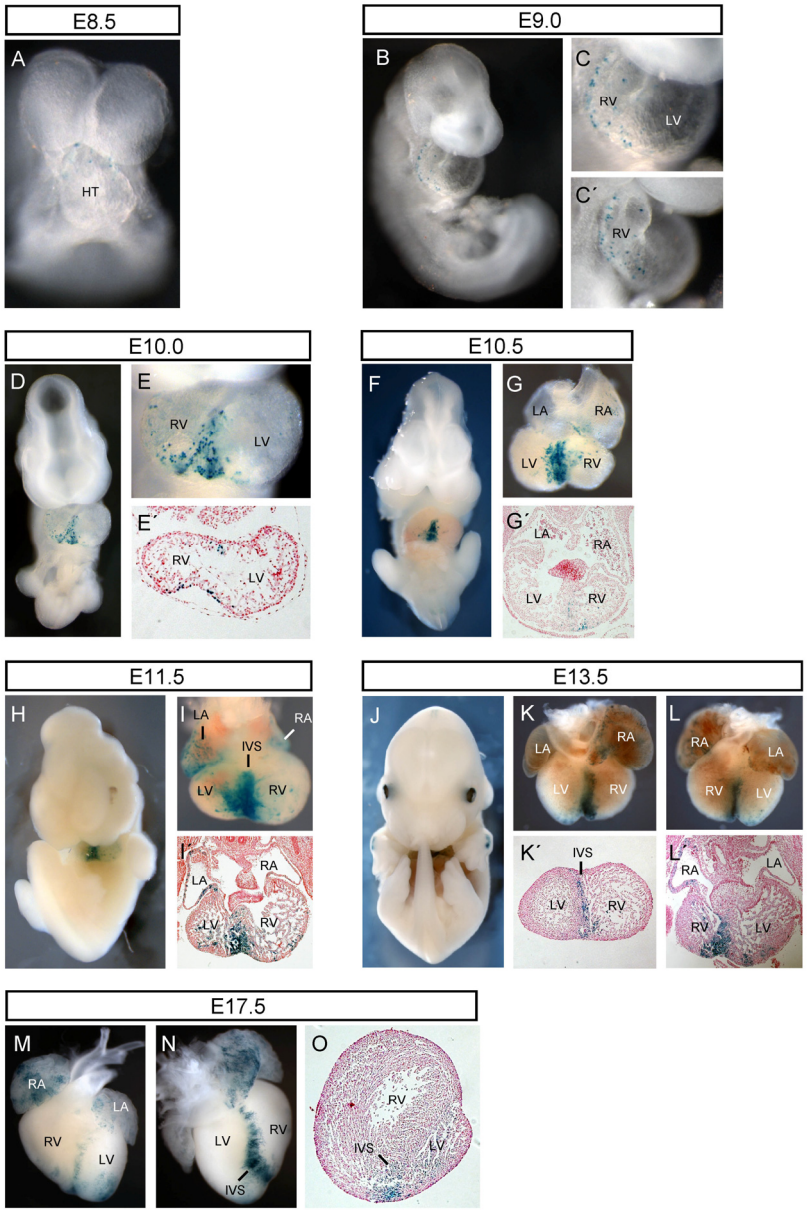

Figure 3.16: Expression patterns of the human wildtype *DPF3* minimal promoter in transgenic mouse embryos of line 1001. X-Gal staining of whole mount embryos and sections at indicated stages. Dorsal views of whole mount embryos and close-ups. HT, heart tube; RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; IVS, interventricular septum.

Determination of the transgene integration sites within the host genome

The transgenic mouse lines were generated by injection of transgene DNA into the pronucleus, which leads to random integration of the DNA into the host genome. The site of integration can have substantial influence of the expression pattern of the transgene with no relationship to transgene copy number. This phenomenon is referred to as position effect (Palmiter and Brinster, 1986). Therefore, the integration sites of the transgenes were analyzed by inverse PCR. This variation of the PCR can be used to determine unknown DNA sequence adjacent to a stretch of known sequence (Ochman *et al.*, 1988). It is based on the formation of

small circular DNA fragments by endonuclease digest followed by dilution of the DNA and subsequent self-ligation to form circles. These fragments can then be used in a PCR with two primers located in the known region, which point outwards, thus amplifying the region adjacent to the known sequence. PCR products can subsequently be sequenced to determine the integration site.

Transgene integration site of line 1001

The transgene integration site of line 1001 was identified in intron1 of the *Clock* gene. A PCR product that maps to mouse chr5:76,695,646-76,696,584 has been amplified, which is flanked on both ends by the appropriate HinfI restriction sites used to produce circular DNA fragments followed by transgene DNA sequence (Figure 3.17).

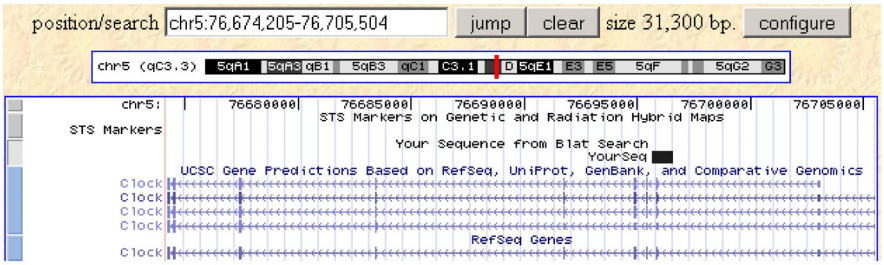

Figure 3.17: Transgene integration site of line 1001 visualized in the UCSC Genome Browser.

The *Clock* gene is an essential regulator of circadian rhythms. It encodes a member of the basic helix-loop-helix/PER-ARNT-SIM family of transcription factors known to play a central role in the control of diverse cellular events (Steeves *et al.*, 1999). It has been shown to be ubiquitously expressed (King *et al.*, 1997).

***DPF3* minimal promoter: Line 1004**

The expression pattern observed in wildtype DPF3 minimal promoter line 1004 was characterized by strong transgene activity in the developing somites of the embryo. Expression was first observed at E7.5 (Figure 3.18A). Somitic expression persisted throughout development (Figure 3.18B-R). Transgene expression was also observed in the developing heart. In contrast to line 1001, activity was seen throughout the heart (Figure 3.18G), although restricted to the endocardial layer (Figure 3.18F,I,J,N,O). Additionally, the *DPF3* minimal promoter was active in the developing eye at E12.5 (Figure 3.18L) as well as forelimbs and hindlimbs at E15.5 (Figure 3.18Q-R´).

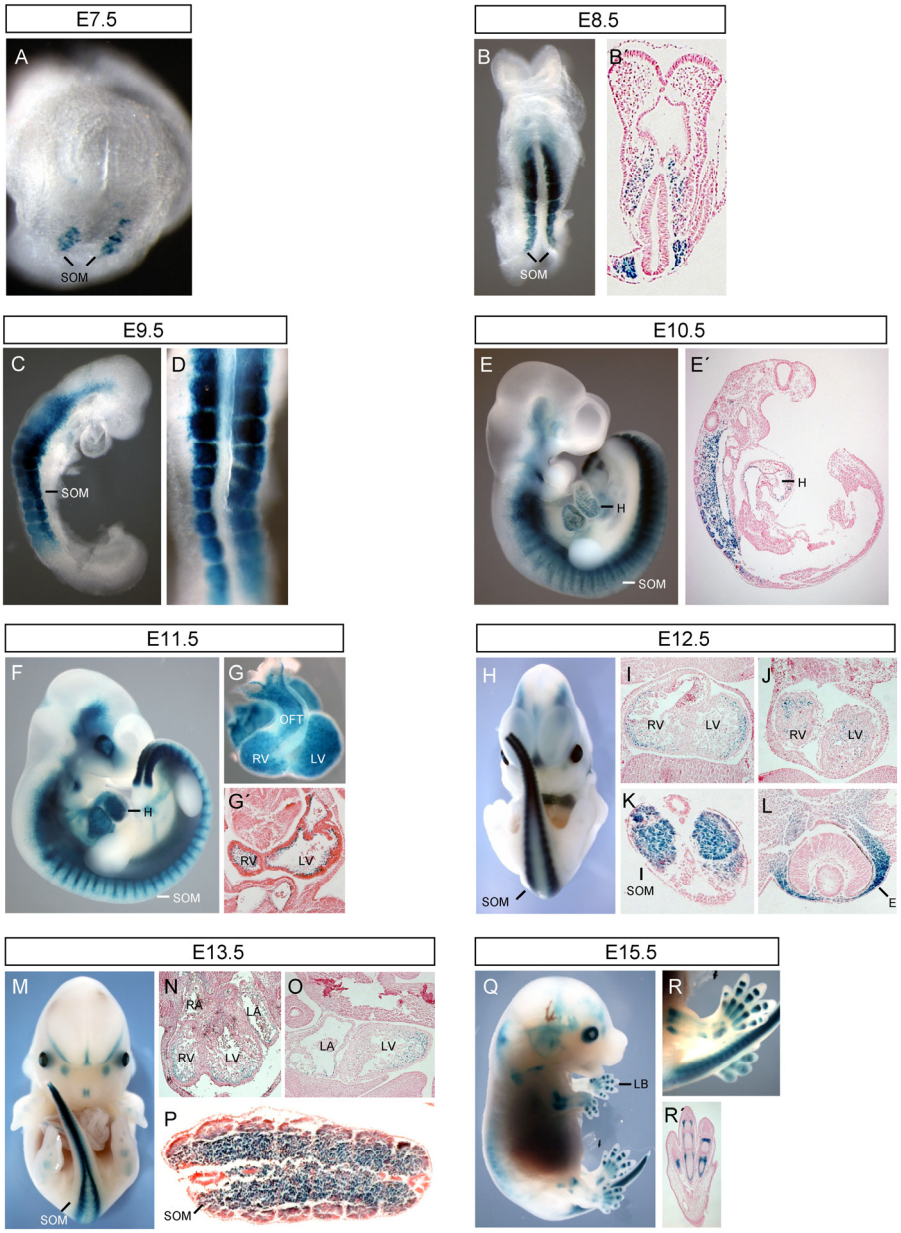

Figure 3.18: Expression patterns of the human wildtype *DPF3* minimal promoter in transgenic mouse embryos of line 1004. X-Gal staining of whole mount embryos and sections at indicated stages. Dorsal/ lateral views of whole mount embryos and close-ups. SOM, somites; H, heart; RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; OFT, outflow tract; LB limb bud; E, eye.

Transgene integration site of line 1004

The transgene integration site of line 1004 was identified to be ~36kb upstream of a locus encoding the Forkhead transcription factor genes *Foxc2* and *Foxl1*. A PCR product that maps to mouse chr8:124052092-124052420 has been amplified, which is flanked on both ends by the appropriate HinfI restriction sites used to produce circular DNA fragments followed by transgene DNA sequence (Figure 3.19).

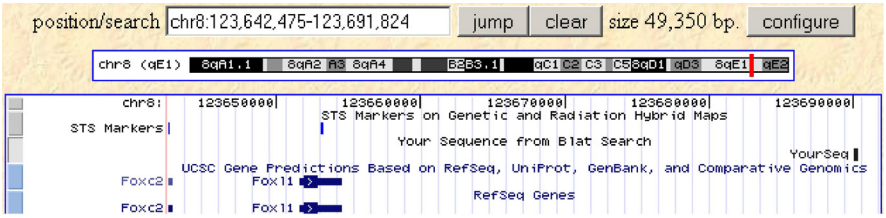

Figure 3.19: Transgene integration site of line 1004 visualized in the UCSC Genome Browser.

Foxc2 is expressed in endothelial cells of the heart, somites and vasculature, while *Foxl1* is a marker for the mesodermal compartments of the developing gut (Kaestner *et al.*, 1996). Comparison of the expression patterns of line 1004 and *Foxc2* mRNA during embryogenesis revealed a striking overlap, indicating that the *lacZ* gene was under the control of *Foxc2* regulatory elements (see also Figure 4.4).

DPF3 minimal promoter: transient transgenics

Due to the discrepancies observed in the expression patterns of line 1001 and 1004, further transgenesis was attempted. In contrast to generating stable transgenic lines and analyzing their offspring, embryos derived from pronuclear injections and retransfer to pseudopregnant foster mothers were directly analyzed by X-Gal staining. Out of 23 litters of injected embryos, only one transgenic litter was obtained. A partially overlapping pattern of expression as in the stable transgenic lines was observed with expression in the developing right ventricle of the heart, somites as well as the apical ectodermal ridge of the limb buds (Figure 3.20).

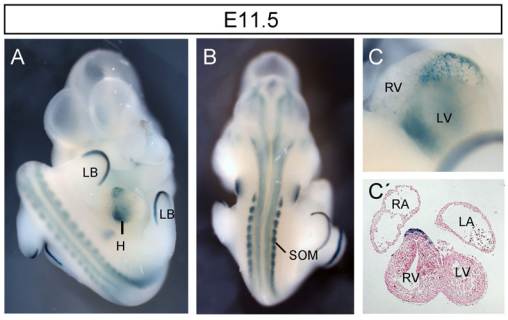

Figure 3.20: Expression patterns of the human wildtype *DPF3* minimal promoter in transiently transgenic mouse embryos. X-Gal staining of whole mount embryos and sections at E11.5. Dorsal/ ventral views of whole mount embryos and close-ups. SOM, somites; H, heart; RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; LB limb bud.

3.5.3 *DPF3* minimal promoter: COUP-TF binding site deletion

Data obtained from reportergene assays in cell culture as well as *in vitro* binding experiments indicated that COUP-TFs are regulators of *DPF3*. Moreover, *Coup-TF2* is expressed in the posterior region of the heart, namely the sinus venosus and atria, during murine development,

overlapping partially with *Dpf3* expression. To test the importance of the identified COUP-TF binding site *in vivo*, transgenic embryos were generated using the COUP-TF binding site deletion construct fused to beta-galactosidase. Two stable mouse lines (Lines 2002, 2027) were obtained and analyzed.

DPF3 minimal promoter: COUP-TF binding site deletion line 2002

Transgene expression of line 2002 was still observed within the heart, although restricted to the anterior portion of the looping heart tube at E9.5 (Figure 3.21A,B). Transgene activity remained specific to the OFT region of the heart throughout development, manifested in strong expression in the aorta and pulmonary artery at E15.5 (Figure 3.21J,K). In addition to the heart, transgene expression was observed in the developing limbs, specifically in the zone of polarizing activity (ZPA) at E11.5 and E12.5 (Figure 3.21E-I). At E15.5, transgene expression in the limbs was still detectable in the former ZPA, but also observed in the progression zone of the outgrowing fingertips (Figure 3.21L, L´).

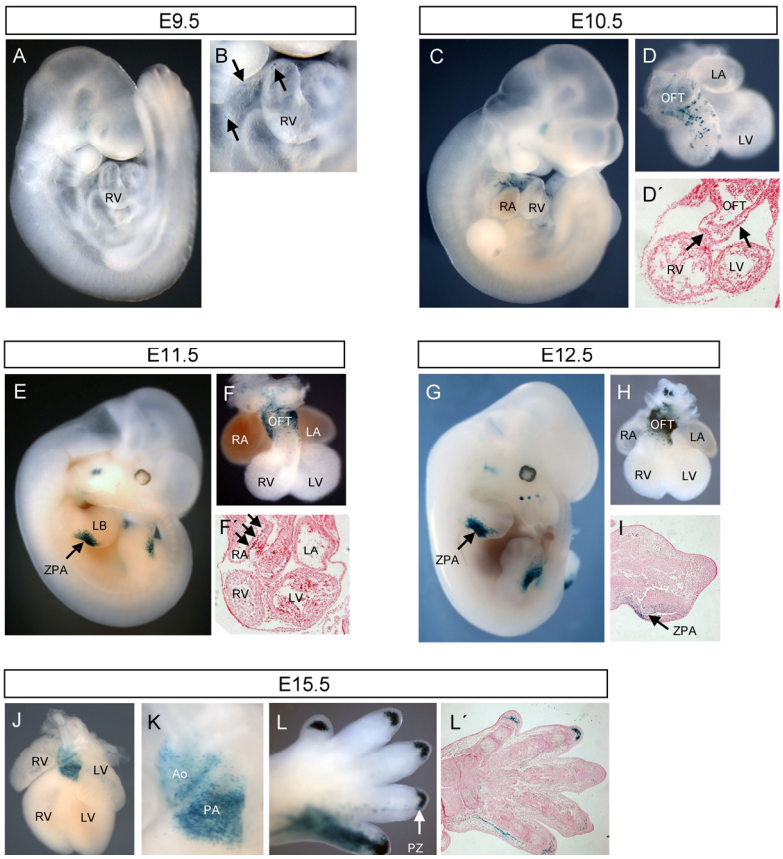

Figure 3.21: Expression patterns of the human *DPF3* minimal promoter COUP-TF deletion in transgenic mouse embryos of line 2002. X-Gal staining of whole mount embryos and sections at indicated stages. Dorsal/ lateral views of whole mount embryos and close-ups. OFT, outflow tract; RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; Ao, aorta; PA, pulmonary artery; ZPA, zone of polarizing activity; LB, limb bud; PZ, progression zone.

Transgene integration site of line 2002

To determine possible position effects on transgene activity, the integration site of the transgene within the host genome was determined by inverse PCR. The transgene of line 2002 had integrated into intron 5 of the unannotated gene 6820431F20Rik on chr8:19,991,120-19,991,510 (Figure 3.22).

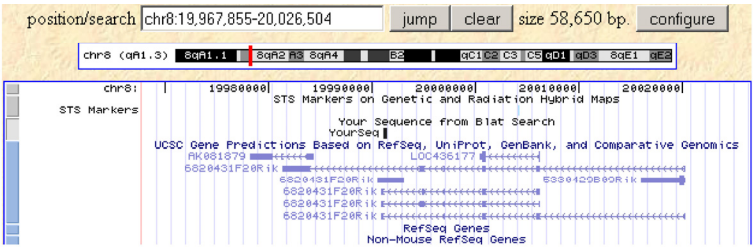

Figure 3.22: Transgene integration site of line 2002 visualized in the UCSC Genome Browser.

This gene is similar to Cadherin-11 (Osteoblast-cadherin, OSF-4). Cdh-11 is one of the factors involved in the process of calcification of growth plate chondrocytes. Interestingly, it is expressed in a proximo-distal gradient within the limb buds as well as the trabeculae of the outflow tract of the heart (Simonneau *et al.*, 1995). The overlapping expression patterns of Cadherin-11 and line 2002 in the outflow tract of the heart indicated that the reporter gene may have been under the control of regulatory elements of the integration site.

DPF3 minimal promoter: COUP-TF binding site deletion line 2027

Transgene expression in line 2027 was almost undetectable, with a faint expression domain observed in the developing eyes. Notably, no expression in the heart was observed (Figure 3.23).

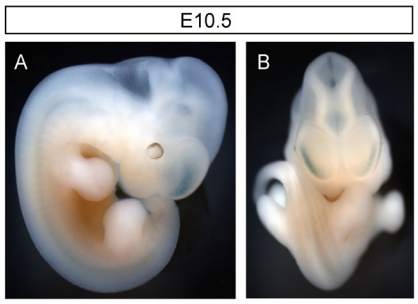

Figure 3.23: Expression patterns of the human *DPF3* minimal promoter COUP-TF deletion in transgenic mouse embryos of line 2027. X-Gal staining of whole mount embryo at E10.5.

Transgene integration site of line2027

The transgene of line 2027 had integrated in a gene-depleted region of chromosome 2 (chr2:102,777,798-102,777,933) with no gene found in a region of 300kb (Figure 3.24). The lack of expression

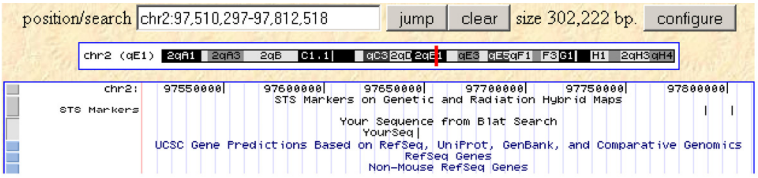

Figure 3.24: Transgene integration site of line 2027 visualized in the UCSC Genome Browser.

3.5.4 *DPF3* minimal promoter: Nkx2.5 binding element (NKE) deletion

Reportergene assays and *in vitro* binding experiments showed that Nkx2.5 is a regulator of *DPF3*. Moreover, *Nkx2.5* transcripts can be observed, among other regions, in the cardiac crescent of the developing heart from mouse E7.5 onwards, overlapping with expression of *Dpf3*. To test the function of the identified NKE *in vivo*, transgenic embryos were generated using the NKE deletion construct fused to beta-galactosidase. Two stable cell lines (Line 2511, 2521) were generated and embryos derived from these lines were analyzed for transgene activity by X-Gal staining.

***DPF3* minimal promoter: NKE deletion line 2511**

No expression in the heart was observed in line 2511 (Figure 3.25C). Instead, a clear expression domain was identified in the developing liver of embryos (Figure 3.25A,A´,B,B´).

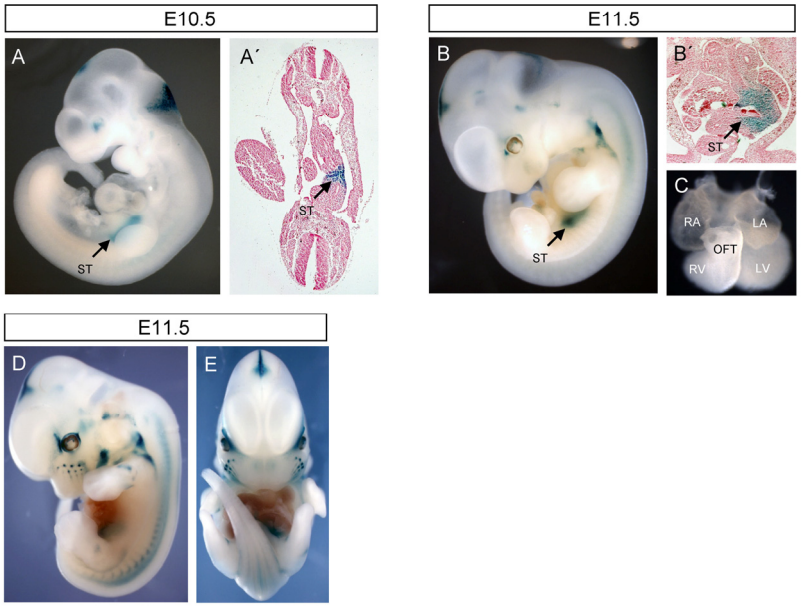

Figure 3.25: Expression pattern of the human *DPF3* minimal promoter NKE deletion in transgenic mouse embryos of line 2511. X-Gal staining of whole mount embryos and sections at indicated stages.

Dorsal/ lateral views of whole mount embryos and close-ups. ST, septum transversum; RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; OFT, outflow tract.

DPF3 minimal promoter: NKE deletion line 2521

As in line 2511, no expression of the transgene was observed in the heart. Expression was faintly observed in the midbrain region of the embryo at E11.5 (Figure 3.26).

Figure 3.26 Expression patterns of the human *DPF3* minimal promoter NKE deletion in transgenic mouse embryos of line 2521. X-Gal staining of whole mount embryo at E11.5.

Transgene integration sites of lines 2511 and 2521

The transgene integration sites of these two lines could not be determined as it was not possible to generate PCR products from digested and recircularized genomic DNA.

3.5.5 A fate-map of *DPF3* minimal promoter activity: Cre transgenics

An interventricular septum (IVS)-specific expression pattern such as the one observed in *DPF3* minimal promoter transgenic Line 1001 (Figure 3.16) had not been described previously, as no molecule is known to be specifically expressed only in this region of the developing heart. The interventricular septum is of great interest for heart development and congenital heart disease, as malformations of the IVS are commonly observed in humans with congenital heart disease.

Therefore, an attempt was made to generate a transgenic mouse line that expresses Cre-Recombinase under control of the *DPF3* minimal promoter (*DPF3*-Cre). Using the Cre/LoxP-system, this would allow for the IVS-specific inactivation of genes during mouse development. Mice were generated using a strategy similar to the one used for lacZ-transgenic mice, with the exception that *lacZ* was replaced by *Cre*-Recombinase. One transgenic line was obtained from pronuclear injections followed by retransfer into pseudopregnant foster mothers. This male founder was mated with female mice of the strain Rosa26R, which contains a version of lacZ in the Rosa-locus, which is activated upon exposure to Cre-Recombinase. Offspring of these matings was used to analyze the activity of Cre-Recombinase

depicted by beta-galactosidase-induced X-Gal-staining. In contrast to direct analysis of beta-galactosidase expression under control of a certain promoter, the expression pattern obtained from matings of Rosa26R mice with Cre-transgenic mice represents all cells in which the promoter was active during life and is thus a fate-map by nature. The transgene was active in the IVS of the developing heart of mouse embryos at E10.5 (Figure 3.27A-C´), although expression was not entirely restricted to the IVS. Expression was expanded into the right and left ventricle as well as the outflow tract region including parts of the atria. Furthermore, additional expression was observed in the sclerotome (Figure 3.27A). In the neonatal heart, Cre-Recombinase expression was most prominent in the atria and the aorta. Due to the expression domains outside the interventricular septum, this Cre-Recombinase transgenic line seemed unsuitable for IVS-specific gene ablations using the Cre/LoxP-System.

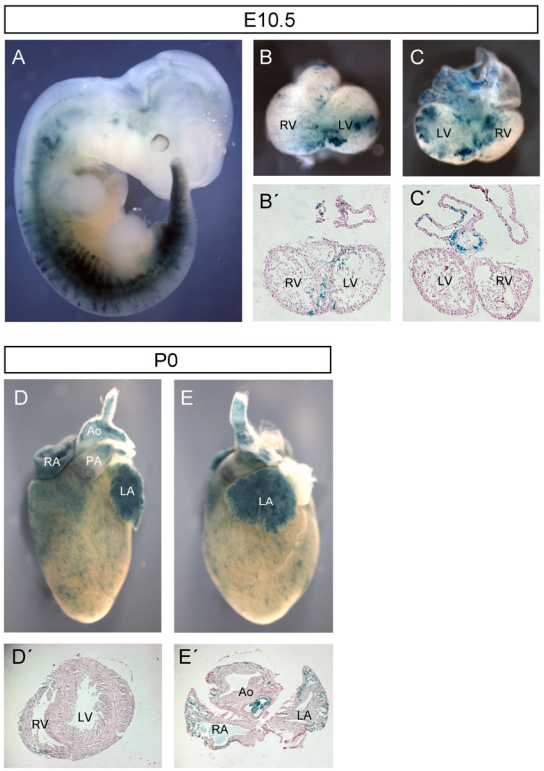

Figure 3.27: Fate map of *DPF3* minimal promoter activity in transgenic mouse embryos of line *DPF3*-Cre. X-Gal staining of whole mount embryos and sections at indicated stages. Lateral views of whole mount embryos and close-ups. ST, septum transversum; RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; OFT, outflow tract.

3.5.6 *DPF3* enhancer

Apart from the *DPF3* minimal promoter present in the region directly adjacent to the transcription start site, a second region of evolutionary sequence conservation was identified between -7 to -8 kb upstream of the TSS of *DPF3* (sequence is located on chr14:72,437,834-72,438,342 of Human Mar.2006 Assembley) (Figure 3.8B). This region, called *DPF3* enhancer, was tested for its ability to influence gene transcription using reportergene assays in cell culture and in transgenic mice.

DPF3 enhancer activity in cell lines

In order to compare the activity of the *DPF3* enhancer in distinct cell lines, luciferase reportergene assay experiments were carried out in HEK293T, COS7, C2C12 and H9c2 cells. Transfection of a construct containing the *DPF3* enhancer fused to an SV40 minimal promoter followed by luciferase revealed that the enhancer is most active in the muscle-specific cell lines C2C12 and H9c2 (Figure 3.28). Little to no activity was observed in HEK293T or COS7 cells, which are derived from kidney of human and chimp, respectively.

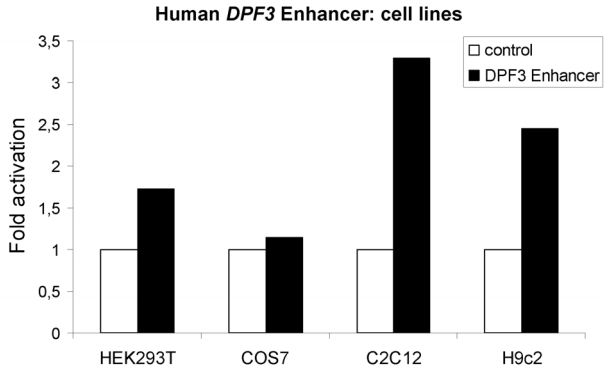

Figure 3.28: Luciferase reportergene assay experiments in cell lines HEK293T, COS7, C2C12 and H9c2. Transfection of a construct containing the *DPF3* enhancer fused to an SV40 minimal promoter followed by the luciferase gene.

DPF3 enhancer in transgenic mice

To test whether a specificity of the *DPF3* enhancer for heart or skeletal muscle could be observed *in vivo*, transgenic mice were generated which expressed beta-galactosidase under the control of the *DPF3* enhancer fused to the *hsp68*-minimal promoter. Only one transgenic founder (Line 17) was obtained from pronuclear injections followed by retransfer into pseudopregnant foster mothers. This male founder was mated with female mice and its offspring analyzed by X-Gal staining (Figure 3.29). Analysis of a midgestation embryo showed activity of the *DPF3* enhancer in the left ventricle of the developing heart as well as faint

expression in the neural tube. Expression in the left ventricle was observed within the myocardium (Figure 3.29E). These results suggested that the *DPF3* enhancer is involved in regulating the cardiac-specific expression pattern of *DPF3*.

The integration site of this transgene could not be determined by means of inverse PCR.

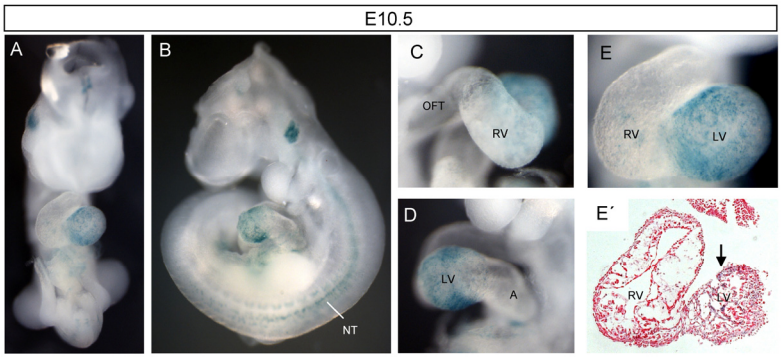

Figure 3.29: Expression pattern of the human *DPF3* enhancer in transgenic mouse embryos of line 17. X-Gal staining of whole mount embryos and sections at E10.5. Dorsal and lateral views of whole mount embryos and close-ups. RV, right ventricle; LV, left ventricle; A, atrium; OFT, outflow tract; NT, neural tube.

3.5.7 DPF3 downstream targets: Analysis of signaling pathways

At the time of the study, DPF3 was annotated in the literature as a transcription factor, due to the presence of a C2H2-type zinc finger. Moreover, only the short isoform DPF3a was known in human. C2H2-type zinc fingers are frequently found in transcription factors and serve as DNA-binding as well as protein-protein interaction domains. Consequently, the effect of DPF3 overexpression on several luciferase reporters for distinct signaling pathways was tested in cell culture. In addition to HEK293T cells, assays were also carried out in the mouse skeletal muscle cell line C2C12. Pathways analyzed included the BMP-, Wnt-, and Mitogen-activated-protein-kinase (MAPK) -pathway. A positive control known to activate the different reporter constructs was included in the experimental setup.

The BMP-pathway was tested using the BRE-luc reporter construct, which consists of repetitive BMP-responsive elements (BREs), namely Smad-binding sites. This reporter is activated by BMP signaling, induced by co-transfection of a constitutively active form of the BMP-receptor 1b (caBMPR1b). CaBMPR1b activated BRE-luc in both HEK293T and C2C12 cells. Co-transfection of BRE-luc with an expression vector for DPF3a led to a reduction of reporter activity, indicating that DPF3a overexpression inhibits BMP signaling (Figure 3.30A).

The Wnt-pathway was analyzed using the TOPFLASH reporter system. It consists of repetitive Lef1-binding sites fused to luciferase, which responds to active Wnt-signaling. The positive control used in this assay was a constitutively active form of Lef1/beta-catenin, which

activated the reporter in both HEK293T and C2C12 cells. Overexpression of DPF3a had no significant effect on the reporter, indicating that DPF3a does not affect Wnt-signaling (Figure 3.30B).

Finally, MAPK signaling was tested using the pFR-Luc / cJun-Gal system. The pFR-Luc vector contains a synthetic promoter with five tandem repeats of yeast GAL4 binding sites and the luciferase gene. The cJun-Gal vector expressing the fusion *trans*-activator protein consists of the activation domain of the c-*Jun* fused with the DNA binding domain of the yeast GAL4. Overexpression of MEKK4, a kinase of the MAPK-pathway led to activation of the reporter system only in C2C12 cells, suggesting that HEK293T cells lack essential additional molecules necessary for proper activation. Overexpression of DPF3a had no effect on MAPK signaling in the analyzed system (Figure 3.30C). Taken together, the analysis of the three pathways showed that the BMP-pathway was repressed by DPF3a overexpression, while Wnt signaling and the MAPK pathway were unaffected.

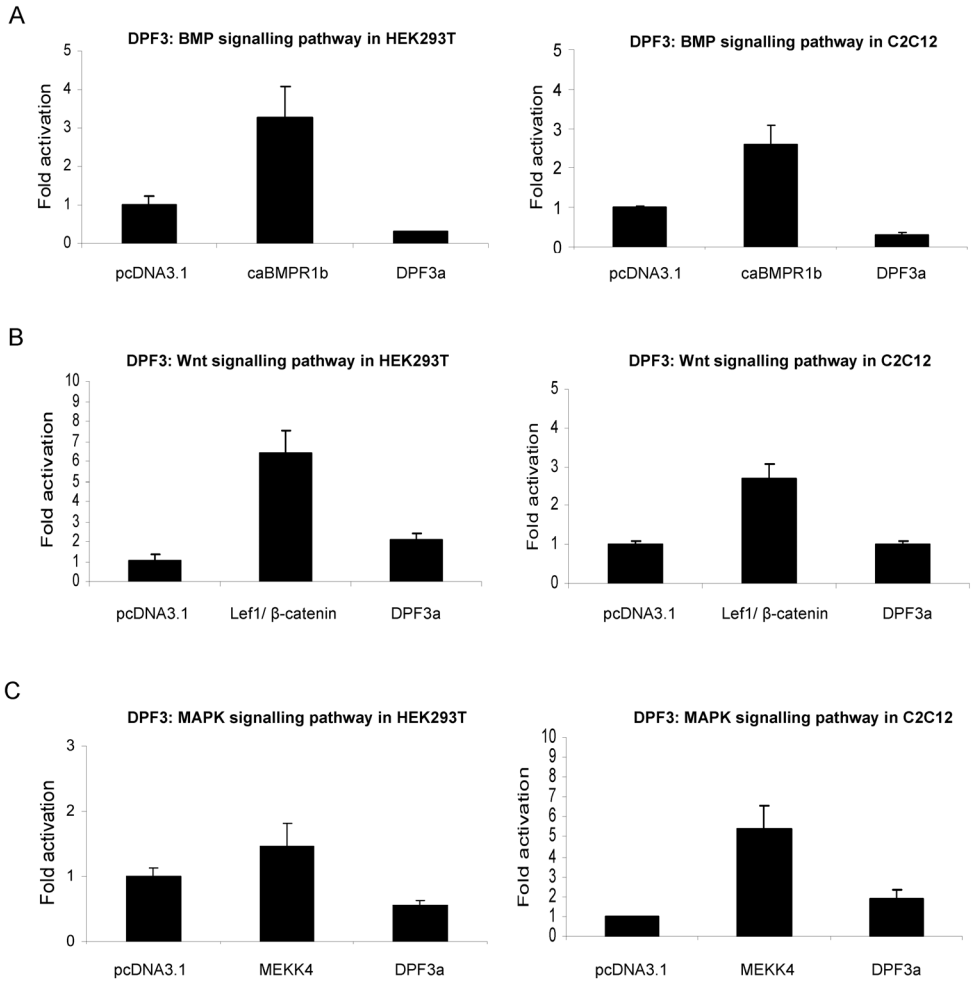

Figure 3.30: Reportergene assays for different signalling pathways and the effect of DPF3a overexpression. Assays have been carried out in HEK293T and C2C12 cells. A) BMP-pathway analysis using BRE-luc. B) Wnt-pathway analysis using TOPFLASH. C) MAPK-pathway analysis using pFR-Luc/cJun-Gal.

3.6 The evolution of the d4 protein family

DPF3 is a member of the d4 protein family. In order to analyze the evolutionary relationship between DPF3, the other d4 family members and their predecessors, sequence database searches were performed. d4 family protein sequences of species representing key steps of evolution were extracted and compared with regard to sequence similarities (Table 3.4).

Table 3.4: Protein evolution of the d4 family.

	Species	Description	Accession number	Conservation to human in %					Human ortholog		
				2/3domain	C2H2	PHD1	PHD2	total	DPF1	DPF2	DPF3
DPF3a	H. sapiens (human)	DPF3a	NP_036206								
	M. musculus (mouse)	Dpf3a	NP_478119.1	100	100	100*	-	**99**	+	+	+
	G. gallus (chicken)	Dpf3a1	AAK51969.1	97	100	96*	-	**87**	+	+	+
	G. gallus (chicken)	Dpf3a2	AAK51970	97	100	96*	-	**84**	+	+	+
DPF3b	H. sapiens (human)	DPF3b	AAX20019.1								
	M. musculus (mouse)	Dpf3b	BAC30204.1	100	100	100	100	**99**	+	+	+
	M. domestica (opposum)	Dpf3b	XP_001375927	100	100	98	100	**96**	+	+	+
	G. gallus (chicken)	Dpf3b1	AAK51967	97	100	94	95	**91**	+	+	+
	G. gallus (chicken)	Dpf3b2	NP_989970	97	100	94	95	**80**	+	+	+
	X. tropicalis (frog)	Dpf3	ENSXETP00000047250	93	95	89	91	**74**	+	+	+
	D. rerio (zebrafish)	dpf3	EU245032	86	95	75	88	**70**	+	+	+
	G. aculeatus (stickleback)	dpf3	ENSGACP00000017212	85	85	75	87	**67**	+	+	+
	T. rubripes (pufferfish)	dpf3	SINFRUP00000135931	82	95	75	86	**60**	+	+	+
									Similarity to human in %		
d4 family	S. purpuratus (sea urchin)	hypothetical protein	XP_788653	65	72	70	84	**49**	46	48	49
	A. mellifera (honey bee)	similar to d4 CG2682-PA	XP_395098	57	68	73	84	**46**	46	47	46
	C. intestinalis (sea squirt)	zinc finger protein	NP_001071860	61	63	75	75	**43**	41	43	43
	D. melanogaster (fruitfly)	d4	CG2682-PA	55	-	73	82	**39**	38	41	39
	C. elegans (worm)	C28H8.9a	NP_498281	26	-	61	66	**34**	33	34	34
C2H2 zinc finger	S. cerevisiae (yeast)	zinc finger protein Sfp1	NP_013507	-	61	-	-	**-**	61	66	61
plant homeodomain 1	S. cerevisiae (yeast)	Snt2_PHD1	NP_011384	-	-	31	-	**-**	27	28	31
plant homeodomain 2	S. cerevisiae (yeast)	Jhd2p_PHD	NP_012653	-	-	-	39	**-**	39	34	39

The shorter isoform DPF3a is found only in human, mouse and chicken. DPF3b, the second isoform, is present in a wider range of species and can be traced back to the origin of vertebrates with orthologs found in zebrafish, stickleback and pufferfish. Moreover, orthologs of the other members of the d4 family, namely DPF1 and DPF2, are also present in all vertebrates analyzed to date. In invertebrate species, only a single d4 protein is found, defined by the presence of the unique N-terminal 2/3-domain. The most derived sequence can be found in the roundworm *C. elegans,* which represents the founder of the d4 family. Interestingly, d4 proteins of protostome species *C. elegans* and *D. melanogaster* lack the C2H2 type zinc finger, whereas it is present in the deuterostomes sea squirt (*C. intestinalis*) and sea urchin (*S. purpuratus*), suggesting that this domain has been added later in evolution. The C2H2-type zinc finger and the PHD finger are domains found in all eukaryotes, including unicellular organisms such as yeast (*S. cerevisiae).* Unlike the 2/3 domain though, they are present in a broad range of protein families. The yeast C2H2-type zinc finger protein with the highest similarity to the C2H2-type zinc finger of human DPF3 is Sfp1. This protein has been

shown to play a role in transcriptional regulation as well as cell size control (Cipollina *et al.*, 2008). The yeast PHD finger proteins with the highest similarity to the human DPF3 PHD fingers are Sant-domain-containing 2 (Snt2) for human PHD1 and jumanji homeodomain 2 protein (Jhd2p) for human PHD2. Snt2 has been implicated in stress response signaling in yeast (Rokhlenko *et al.*, 2007). Aside from the PHD finger, it contains a Bromo adjacent homology domain (BAH), which has been shown to have DNA and nucleosome binding properties (Connelly *et al.*, 2006).Yeast Jhd2p is a histone 3 lysine 4 trimethyldemethylase (Liang *et al.*, 2007). The jumonji family regulates chromatin, gene expression, and controls development through various signaling pathways (Takeuchi *et al.*, 2006). Taken together, the d4 protein family, which can first be detected in *C. elegans* appears to be a compound protein of several domains found in species as low as yeast. Sequence comparisons revealed that sea urchin d4 (*Strongylocentrotus purpuratus*) is the protein most closely related to human DPF3 before the occurrence of a gene duplication event (Table 3.4). As the d4 family members in higher vertebrates show distinct expression patterns (see Chapter 1.1), the expression pattern of sea urchin *d4* was analyzed by *in situ* hybridization. This allowed for a comparison of the initial domains of expression to the maintained and possibly newly acquired expression domains of d4 family members. *In situ* hybridization was performed on early blastula to late gastrula stages (Figure 3.31). Expression of *d4* mRNA was first observed at the early blastula stage in the vegetal plate of the embryo populated by cells of the skeletogenic lineage and the small micromere precursors of adult mesoderm. During late blastula stages, *d4* expression was maintained at the vegetal plate in the region of the developing primitive gut, which is of endodermal origin as well as in cells that contribute to the mesodermal lineage, which later also give rise to muscle cells. At late gastrula stage, strong *d4* expression was observed in the fore- and hindgut, in which cells of mesodermal as well as endodermal origin are located.

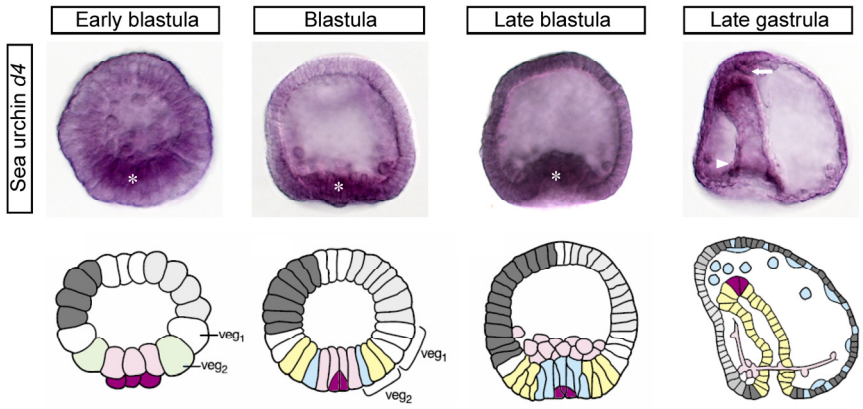

Figure 3.31: *In situ* hybridization of *d4* mRNA in sea urchin embryos at indicated stages. Asterisk, vegetal plate; arrow, foregut, arrowhead, hindgut. Shown below are schematic diagrams of *S. purpuratus* color-coding specified regions of endomesoderm. lavender, skeletogenic lineage; darker

purple, the small micromere precursors of adult mesoderm; green, endomesodermal veg2 lineage that later gives rise to endoderm, yellow, and to mesoderm, blue. Diagrams modified from (Davidson *et al.*, 2002).

Taken together, *d4* expression in sea urchin embryos was observed in tissues of mesodermal and endodermal origin, whereas expression was not observed in cells of the ectoderm.