

Additional Data File 4

Biological Background

The data set consisted of 350 genes that were upregulated in pleiomorphic adenoma of the salivary glands of mice compared to normal mice. These adenoma are benign **epithelial tumours** of the salivary glands. Microscopically, they show a marked histological diversity with **epithelial** and **mesenchymal** components in a variety of patterns. They form in the cartilage and consist mainly of **fibroblasts**. They are the result of recurrent chromosome rearrangements, particularly reciprocal translocations, with breakpoints at 8q12, 3p21 and 12q13-15. It was demonstrated that the t(3;8)(p21;q12) translocation results in promoter swapping between PLAG1, a **developmentally** regulated **zinc finger** gene and the constitutively expressed gene for beta-catenin (CTNNB1), a protein functioning in the **WG/WNT** signalling pathway and specification of cell fate during **embryogenesis**. The promoter swapping leads to activation of PLAG1 expression and reduced expression of CTNNB1 [1]. In normal conditions, PLAG1 is expressed in **placenta** and **fetal** tissues, with no detectable expression in adult salivary glands [2]. The abnormal PLAG1 expression in the adenoma results in deregulation of PLAG1 target genes causing salivary gland tumorigenesis [4]. The main target of PLAG1 was identified as the insulin-like growth factor II (**IGF2**) gene [3]. IGF2 perfectly fits in the picture of a restarted **developmental** program with concomitant loss of **differentiation**, the typical hallmark for any tumour [4]. Furthermore, glial **intermediate filament** proteins [10] coexpressed with various **keratins** such as cytokeratin 14 [5] and cytokeratin 18, a known suppressor of **apoptosis** [6], were found to be effective markers of this type of **tumours**. Additionally, **proteoglycans** [9], **collagen** and glycosaminoglycans such as **chondroitin-sulfate**, were found characteristic for pleiomorphic adenoma and are thought to influence the proliferation, **differentiation**, secretory activity and shape of these **tumour** cells [7]. It was also indicated that oncogene **p53** is mutated in these cells and may serve as a **tumour** marker [8].

TXTGate profiles

GENERIF-GO				Medline-eVOC			
	mean		variance		mean		variance
express	0.08172	express	0.01339	organ	0.28121	organ	0.01383
cell	0.05138	keratin	0.01087	intern	0.15775	intern	0.00672
develop	0.04191	develop	0.00899	normal	0.15710	growth	0.00599
function	0.04129	frameshift	0.00866	red	0.13482	development	0.00455
activ	0.03759	pattern	0.00752	male	0.12830	fibroblast	0.00449
regul	0.03753	embryogenesi	0.00719	femal	0.12795	tumour	0.00391
induc	0.03283	phosphoryl	0.00660	visual	0.11712	red	0.00386
pathwai	0.02807	elem	0.00648	capillari	0.11152	nucleu	0.00378
interact	0.02678	induc	0.00589	system	0.11027	normal	0.00373
protein	0.02563	embryon	0.00579	optic	0.10717	embryo	0.00358
gene	0.02375	function	0.00577	retina	0.10631	teratoma	0.00347
respons	0.02365	apoptosi	0.00551	viral	0.09760	depend	0.00339
region	0.02347	cell	0.00541	bacteri	0.09706	stem_cell	0.00337
modul	0.02270	direct	0.00536	adult	0.07234	kidnei	0.00321
signal	0.02263	enhanc	0.00534	chain	0.06789	epithelium	0.00303
apoptosi	0.02195	interact	0.00518	cell	0.04885	visual	0.00300
site	0.02114	regul	0.00502	growth	0.04799	multipl	0.00298
pattern	0.02071	region	0.00492	tissu	0.04328	skin	0.00294
increas	0.02001	activ	0.00487	development	0.03822	muscl_cell	0.00286
direct	0.01981	helicas	0.00481	metabol	0.03736	system	0.00271
inhibit	0.01918	heart	0.00475	embryo	0.03675	capillari	0.00269
control	0.01909	mechan	0.00472	fibroblast	0.03272	mammari	0.00269
embryon	0.01908	excretion	0.00455	tumour	0.03180	type_ii	0.00269
mechan	0.01818	modul	0.00454	depend	0.03009	bacteri	0.00266
effect	0.01787	effect	0.00449	genet	0.02810	male	0.00260
level	0.01781	increas	0.00446	alpha	0.02809	small	0.00260
result	0.01769	antibodi	0.00437	brain	0.02804	viru	0.00259
brain	0.01768	tissu	0.00431	kidnei	0.02621	femal	0.00258
enhanc	0.01761	copper	0.00429	multipl	0.02396	trophoblast	0.00256
tissu	0.01739	target	0.00426	nucleu	0.02074	testi	0.00253
tumor	0.01660	transform	0.00426	lung	0.02061	optic	0.00238
promot	0.01653	absorpt	0.00414	liver	0.02052	pool	0.00237
earli	0.01627	gene	0.00413	primari	0.01981	retina	0.00234
target	0.01627	pathwai	0.00403	multipl_scleros	0.01943	chain	0.00232
import	0.01598	chondroitin_su	0.00403	muscl_cell	0.01905	brain	0.00231
lung	0.01582	tip	0.00396	skin	0.01890	skelet	0.00226
involv	0.01541	respons	0.00385	epithelium	0.01835	adipocyt	0.00225
bind	0.01526	site	0.00376	heart	0.01826	teratocarcino	0.00223
domain	0.01496	dna	0.00374	spleen	0.01697	central_nervc	0.00222
mediat	0.01474	relat	0.00369	affect	0.01688	multipl_sclerc	0.00217
embryogenes	0.01465	hnrnp	0.00367	small	0.01687	intermedi	0.00217
induct	0.01439	iron	0.00364	neuron	0.01610	neuron	0.00216
elem	0.01438	mrna_transcrip	0.00362	stem_cell	0.01589	lung	0.00214
growth	0.01407	hair	0.00355	diseas	0.01553	heart	0.00211
differenti	0.01402	util	0.00354	carcinoma	0.01506	mesoderm	0.00206
keratin	0.01381	control	0.00354	skelet	0.01504	macrophag	0.00203
depend	0.01319	p53	0.00352	mesoderm	0.01433	placenta	0.00203

Medline-MESH				Medline-HUGO			
	mean		variance		mean		variance
encyclopedia	0.18233	encyclopedia	0.00652	cast	0.37695	cast	0.03576
reproduc_of_resul	0.12165	reproduc_of_result	0.00290	red	0.27270	rod	0.01882
gene_librari	0.11409	project	0.00262	rod	0.26967	red	0.01825
project	0.10267	comput_biologi	0.00243	poli	0.02181	poli	0.00935
standard	0.09417	gene_librari	0.00237	span	0.01945	marck	0.00846
comput_biologi	0.09308	proteoglycan	0.00211	sp1	0.01573	p53	0.00480
organ	0.09291	cyclin	0.00185	e12	0.01215	cd44	0.00474
extend	0.08998	extend	0.00175	p53	0.01190	race	0.00456
mous	0.08609	standard	0.00172	marck	0.01114	sp1	0.00453
mice	0.08307	organ	0.00152	abl	0.00950	p19	0.00435
rna_cap	0.07551	collagen	0.00125	cat	0.00928	ap_2	0.00420
method	0.07529	rna_cap	0.00110	p19	0.00818	glut	0.00399
base	0.07374	monosaccharid_transport	0.00105	cap	0.00805	span	0.00398
singl	0.06920	keratin	0.00102	cord	0.00791	zap	0.00389
prepar	0.06711	membran_glycoprotein	0.00101	ap_2	0.00775	hnrnp	0.00382
imag_cytometri	0.06650	subtract	0.00099	race	0.00758	ptn	0.00377
domest_anim	0.06393	immedi_earli_gene	0.00097	trait	0.00755	stat5	0.00376
subtract	0.06245	base	0.00097	ptn	0.00742	ltbp	0.00369
frozen_section	0.06195	basem_membran	0.00096	wnt	0.00692	e12	0.00351
collect	0.06132	helix_loop_helix_motif	0.00096	ap2	0.00662	id4	0.00340
sensit	0.06080	phosphoglycer_kinas	0.00093	ebp	0.00657	cadherin	0.00333
opsin	0.05666	collect	0.00092	hnrnp	0.00654	trait	0.00330
blue	0.05635	teratoma	0.00090	fat	0.00642	ebp	0.00322
sequenc_analysi	0.05533	fetal_develop	0.00089	cadherin	0.00640	adm	0.00311
intern	0.05509	imag_cytometri	0.00088	bhlh	0.00622	pit	0.00309
specif	0.05493	prepar	0.00083	stat5	0.00612	ap2	0.00305
pick	0.05437	tran_activ	0.00082	erk	0.00591	sre	0.00298
genom	0.05400	domest_anim	0.00081	adm	0.00586	gas1	0.00295
physic	0.05342	heterogen_nuclear_ribon	0.00080	p21	0.00576	wnt	0.00285
proport	0.05329	physic	0.00079	call	0.00551	igfbp	0.00285
econom	0.05295	ceruloplasmin	0.00078	glut	0.00547	apc	0.00280
dna	0.05206	intern	0.00077	zap	0.00510	bpag2	0.00272
ultraviolet	0.05094	3t3_cell	0.00077	e10	0.00505	fat	0.00264
cone	0.05041	zinc_finger	0.00077	fgf	0.00497	notch1	0.00260
pregnanc	0.05030	cytokin	0.00076	cd44	0.00496	ubc	0.00259
rhodopsin	0.05018	frozen_section	0.00075	face	0.00479	cyr61	0.00257
develop	0.04956	singl	0.00075	ndr1	0.00473	rac1	0.00257
comparison	0.04835	squalen	0.00075	il_4	0.00470	bhlh	0.00250
care	0.04821	thrombospondin	0.00074	sac	0.00464	arc	0.00245
rna	0.04784	dna_bind_protein	0.00073	pdz	0.00456	mpf	0.00245
bacteriophag_lam1	0.04757	lactat	0.00072	il_6	0.00454	tdd5	0.00245
tertiari	0.04715	sensit	0.00071	dmt1	0.00446	net	0.00243
sequenc	0.04711	stem_cell_factor	0.00070	112	0.00446	abl	0.00242
cast	0.04708	alpha	0.00070	sre	0.00446	cap	0.00238
genet_model	0.04693	microfila_protein	0.00070	igfbp	0.00444	cord	0.00236
reagent	0.04680	mammari	0.00068	id3	0.00440	ahr	0.00230
system	0.04672	immedi_earli_protein	0.00068	met	0.00440	met	0.00224

Medline-GO				Medline-OMIM			
	mean		variance		mean		variance
project	0.15852	project	0.00634	sensit	0.19887	sensit	0.00839
extend	0.14440	extend	0.00463	famili	0.16792	famili	0.00607
rna_cap	0.13473	rna_cap	0.00383	chromosom	0.15650	chromosom	0.00600
organ	0.12202	organ	0.00265	dna	0.15448	pick	0.00496
sensit	0.10797	sensit	0.00242	specif	0.13607	dna	0.00472
singl	0.10722	singl	0.00212	blue	0.12842	glycoprotein	0.00385
accuraci	0.08623	collagen	0.00197	select	0.12305	phosphatas	0.00382
rna	0.08425	cyclin	0.00182	enhanc	0.12141	blue	0.00359
pregnanc	0.08393	notch	0.00176	isol	0.11045	kinas	0.00351
intern	0.07990	syndecan	0.00174	cone	0.10423	growth	0.00335
develop	0.07857	intern	0.00173	rod	0.10302	lactat	0.00314
biosynthesi	0.07838	accuraci	0.00162	sever	0.10183	mammari	0.00292
specif	0.07605	myristoyl	0.00154	capillari	0.10107	adhesion	0.00290
opsin	0.07540	basem_membran	0.00149	optic	0.10037	skin	0.00288
genom	0.07538	keratin	0.00146	imag	0.09782	induc	0.00287
start	0.07437	proteoglycan	0.00145	system	0.09037	fetal	0.00285
blue	0.07241	pregnanc	0.00142	doubl	0.08944	alpha	0.00282
cone	0.06768	monosaccharid_transp	0.00132	male	0.08635	earli	0.00281
rhodopsin	0.06734	protein_tyrosin_phosph	0.00130	femal	0.08573	select	0.00281
dna	0.06650	opsin	0.00124	short	0.07868	membran	0.00279
translat	0.06589	cytoskelet_protein	0.00124	long	0.07065	factor	0.00272
tertiari	0.06589	p53	0.00123	cluster	0.06595	pattern	0.00241
transcript	0.06553	growth_factor	0.00120	combin	0.06409	rod	0.00238
optic	0.06498	lactat	0.00114	type	0.06394	capillari	0.00237
coloni	0.06368	platelet_deriv_growth_f	0.00113	adult	0.06315	specif	0.00236
select	0.06067	blue	0.00112	process	0.06009	junction	0.00235
enhanc	0.06051	tertiari	0.00112	chain	0.05818	imag	0.00231
templat	0.05982	laminin	0.00112	factor	0.04105	cone	0.00230
red	0.05901	casein	0.00111	tissu	0.03925	kidnei	0.00226
rod	0.05885	gestat	0.00111	growth	0.03819	optic	0.00221
branch	0.05852	embryon	0.00111	acid	0.03770	enhanc	0.00219
plant	0.05844	transcript_factor	0.00110	cell	0.03761	respons	0.00212
contain	0.05786	dna_bind	0.00109	earli	0.03600	transport	0.00209
potenti	0.05653	phosphoglycer	0.00104	bind	0.03589	complex	0.00207
system	0.05507	heterogen_nuclear_ribc	0.00104	pattern	0.03319	mesenchym	0.00204
visual	0.05483	translat	0.00101	induc	0.03268	brain	0.00204
green	0.05455	embryo	0.00099	membran	0.03045	doubl	0.00198
high	0.05443	rna	0.00099	activ	0.03013	total	0.00197
3_end	0.05335	genom	0.00099	mai	0.02963	receptor	0.00195
integr	0.05303	dystroglycan	0.00098	metabol	0.02928	depend	0.00193
cycl	0.05271	rhodopsin	0.00097	respons	0.02836	tumor	0.00191
mrna	0.05241	cone	0.00097	brain	0.02650	multipl	0.00188
doubl	0.05137	start	0.00096	kinas	0.02598	link	0.00188
carboxyl	0.05134	zinc	0.00094	receptor	0.02583	phase	0.00187
sever	0.05111	embryon_develop	0.00094	fetal	0.02433	skelet	0.00186
5_end	0.05106	attach	0.00091	tumor	0.02404	transient	0.00185
increas	0.05098	optic	0.00090	complex	0.02401	inhibitor	0.00184

Func-GO				GO-GO			
	mean		variance		mean		variance
protein_fold	0.00276	protein_fold	0.00109	activ	0.10012	extracellular_space	0.03270
saccharomyc	0.00255	saccharomyc	0.00092	extracellular_space	0.09947	membran	0.02141
er	0.00250	er	0.00089	membran	0.06788	integr	0.01717
oxid	0.00231	oxid	0.00076	integr	0.05652	nucleu	0.01502
endoplasm_reticu	0.00230	endoplasm_retic	0.00075	nucleu	0.05203	extracellular_matrix	0.01148
integr_membran_	0.00218	integr_membran	0.00068	regul	0.03824	atp_bind	0.01077
gene	0.00192	transmembran_r	0.00057	atp_bind	0.03129	calcium_ion_bind	0.01060
protein	0.00184	characterist	0.00036	dna_bind	0.03028	intracellular_signal_c	0.00956
involv	0.00181	involv	0.00035	cell_adhesion	0.02855	nucleic_acid_bind	0.00911
encod	0.00151	transform	0.00034	dna_depend	0.02674	calmodulin_bind	0.00903
similar	0.00135	gene	0.00032	transcript	0.02584	cell_cycl	0.00864
transmembran_re	0.00115	protein	0.00029	extracellular_matrix	0.02565	regul	0.00860
characterist	0.00092	similar	0.00026	calcium_ion_bind	0.02330	dna_bind	0.00857
transform	0.00090	cellular	0.00025	cell_cycl	0.02018	intermedi_filam	0.00841
cellular	0.00076	encod	0.00025	protein_bind	0.01889	cell_adhesion	0.00836
homolog	0.00075	homolog	0.00024	cytoskeleton	0.01849	molecular_function_u	0.00772
defens	0.00069	defens	0.00021	nucleic_acid_bind	0.01779	dna_depend	0.00760
kinas	0.00068	kinas	0.00020	molecular_function_u	0.01729	biolog_process_unkr	0.00724
compound	0.00068	compound	0.00020	cell_growth	0.01597	receptor_activ	0.00691
resist	0.00064	resist	0.00018	intermedi_filam	0.01567	actin_bind	0.00691
membran_associ	0.00063	membran_assoc	0.00017	kinas	0.01510	transcript	0.00670
mrp	0.00063	mrp	0.00017	hydrolas	0.01492	cellular_compon_unl	0.00669
atp_bind_cassett	0.00063	atp_bind_casset	0.00017	biolog_process_unkr	0.01478	cell_growth	0.00649
abc	0.00063	abc	0.00017	cellular_compon_unk	0.01459	cytoskelet_protein_b	0.00645
tap	0.00062	tap	0.00017	translat_regul	0.01445	cytoskeleton	0.00630
protein_transport	0.00061	protein_transpor	0.00016	oxidoreductas	0.01423	translat_regul	0.00592
major	0.00051	major	0.00011	intracellular_signal_c	0.01421	protein_bind	0.00586
membran	0.00049	membran	0.00010	peptidolysi	0.01364	copper_ion_bind	0.00579
molecul	0.00049	molecul	0.00010	proteolysi	0.01364	oxidoreductas	0.00577
transport	0.00048	transport	0.00010	microfibril	0.01345	lipid_bind	0.00541
				extracellular	0.01320	microfibril	0.00535
				actin_bind	0.01227	ferrooxidas	0.00507
				calmodulin_bind	0.01207	cytoskeleton_organ	0.00498
				transferas	0.01196	transferas	0.00487
				biogenesi	0.01180	peptidolysi	0.00452
				intracellular	0.01178	proteolysi	0.00452
				signal_transduct	0.01161	activ	0.00452
				extracellular_matrix_:	0.01140	electron_transport	0.00452
				cytoskeleton_organ	0.01136	hydrolas	0.00448
				transcript_factor	0.01101	heparin_bind	0.00448
				structur_molecul	0.01055	growth_factor_bind	0.00423
				structur_constitu	0.01045	cytokinesi	0.00421
				copper_ion_bind	0.01045	biogenesi	0.00420
				receptor_activ	0.01032	cell_cycl_arrest	0.00416
				plasma_membran	0.01021	extracellular	0.00410
				endopeptidas	0.01018	kinas	0.00408
				peptidas	0.01013	gtp_bind	0.00401

References

1. Nat Genet. 1997 Feb;15(2):170-4. Promoter swapping between the genes for a novel zinc finger protein and beta-catenin in pleomorphic adenomas with t(3;8)(p21;q12) translocations. Kas K, Voz ML, Roijer E, Astrom AK, Meyen E, Stenman G, Van de Ven WJ
2. Lab Invest. 1999 May;79(5):583-9. Pleomorphic adenoma gene 1 is expressed in cultured benign and malignant salivary gland tumor cells. Queimado L, Lopes C, Du F, Martins C, Bowcock AM, Soares J, Lovett M.
3. Cancer Res. 2000 Jan 1;60(1):106-13. PLAG1, the main translocation target in pleomorphic adenoma of the salivary glands, is a positive regulator of IGF-II. Voz ML, Agten NS, Van de Ven WJ, Kas K.
4. Verh K Acad Geneesk Belg. 2001;63(1):35-40. Molecular techniques lead to the first insights into the pathophysiology of salivary gland adenomas. Kas K.
5. Appl Immunohistochem Mol Morphol. 2000 Sep;8(3):195-202. Application of immunohistochemistry to the diagnosis of salivary gland tumors. de Araujo VC, de Sousa SO, Carvalho YR, de Araujo NS.
6. Anticancer Res. 2000 Jul-Aug;20(4):2485-7. Immunohistochemical detection of cytokeratin 18 and its neo-epitope in human salivary glands and pleomorphic adenomas. Kusama K, Jiang Y, Ohno J, Shikata H, Ishikawa F, Taguchi K, Kikuchi K, Mori K, Sakashita H, Sakagami H, Kaneko T, Yamamoto Y.
7. J Oral Pathol Med. 1998 Jul;27(6):272-7. Localization of glycosaminoglycans (GAGs) in pleomorphic adenoma (PA) of salivary glands: an immunohistochemical and histochemical evaluation. Zhao M, Takata T, Ogawa I, Miyauchi M, Ito H, Nikai H
8. Virchows Arch. 2002 Nov;441(5):428-36. Epub 2002 Sep 11. Expression profiles of p53, p63, and p73 in benign salivary gland tumors. Weber A, Langhanki L, Schutz A, Gerstner A, Bootz F, Wittekind C, Tannapfel A.
9. J Oral Pathol Med. 1999 Jan;28(1):37-42. Immunohistochemical evaluation of the small and large proteoglycans in pleomorphic adenoma of salivary glands. Zhao M, Takata T, Ogawa I, Yada T, Kimata K, Nikai H.
10. Differentiation. 1986;31(3):206-27. Expression of glial filament protein (GFP) in nerve sheaths and non-neural cells re-examined using monoclonal antibodies, with special emphasis on the co-expression of GFP and cytokeratins in epithelial cells of human salivary gland and pleomorphic adenomas. Achstatter T, Moll R, Anderson A, Kuhn C, Pitz S, Schwechheimer K, Franke WW.