Published in : Oncogene (2004), vol. 23, pp. 179-191.
Status : Postprint (Author's version)

Microarray screening for target genes of the proto-oncogene PLAG1

Marianne L Voz[1,2], Janick Mathys[3], Karen Hensen[1], Hélène Pendeville[2], Isabelle Van Valckenborgh[1], Christophe Van Huffel[4], Marcela Chavez[1], Boudewijn Van Damme[5], Bart De Moor[3], Yves Moreau[2] and Wim JM Van de Ven[1]

[1] Laboratory for Molecular Oncology, Center for Human Genetics, KU Leuven & Flanders Interuniversity Institute for Biotechnology, Herestraat 49, Leuven B-3000, Belgium; [2]Laboratoire de Biologie Moléculaire et de Génie Génétique, Université de Liège, Allée du XVI-août, Institut de Chimie, B6, Sart-Tilman B-4000, Belgium; [3]ESAT-SCD, KU Leuven, Kasteelpark Arenberg 10, Leuven B-3001, Belgium; [4] Department of Genomics, Starlab NV/SA, Emgeland, 555 B-1180 Busselles, Belgium; [5] Department of Morphology and Molecular Pathology, Katholieke Universiteit Leuven, Minderbroederstraat 33, Leuven B-3000, Belgium

Abstract

PLAG1 is a proto-oncogene whose ectopic expression can trigger the development of pleomorphic adenomas of the salivary glands and of lipoblastomas. As PLAG1 is a transcription factor, able to activate transcription through the binding to the consensus sequence GRGGC$(N)_{6-8}$GGG, its ectopic expression presumably results in the deregulation of target genes, leading to uncontrolled cell proliferation. The identification of PLAG1 target genes is therefore a crucial step in understanding the molecular mechanisms involved in PLAG1-induced tumorigenesis. To this end, we analysed the changes in gene expression caused by the conditional induction of PLAG1 expression in fetal kidney 293 cell lines. Using oligonucleotide microarray analyses of about 12000 genes, we consistently identified 47 genes induced and 12 genes repressed by PLAG1. One of the largest classes identified as upregulated PLAG1 targets consists of growth factors such as the insulin-like growth factor II and the cytokine-like factor 1. The *in silico* search for PLAG1 consensus sequences in the promoter of the upregulated genes reveals that a large proportion of them harbor several copies of the PLAG1-binding motif, suggesting that they represent direct PLAG1 targets. Our approach was complemented by the comparison of the expression profiles of pleomorphic adenomas induced by PLAG1 versus normal salivary glands. Concordance between these two sets of experiments pinpointed 12 genes that were significantly and consistently upregulated in pleomorphic adenomas and in PLAG1-expressing cells, identifying them as putative PLAG1 targets in these tumors.

Keywords : *PLAG1* gene ; oligonucleotide microarray ; target genes ; pleomorphic adenoma ; salivary gland

Introduction

Oncogenic activation of the *PLAG1* gene on 8q12 is a crucial event in the formation of pleomorphic adenomas of the salivary glands. This activation mainly results from recurrent chromosomal translocations that lead to promoter substitution between *PLAG1*, a gene primarily expressed in fetal tissues, and more broadly expressed genes (Kas *et al*., 1997; Voz *et al*., 1998; Astrom *et al*., 1999). The three translocation partners characterized so far are the *β-catenin* gene (Kas *et al*., 1997), the *leukemia inhibitory factor receptor* gene (Voz *et al*., 1998) and the *elongation factor SII* gene (Astrom *et al*., 1999). Breakpoints invariably occur in the 5' noncoding region of the *PLAG1* gene, leading to an exchange of the regulatory control elements while preserving the *PLAG1* coding sequence. The replacement of the *PLAG1* promoter, inactive in adult salivary glands, by a strong promoter derived from the translocation partner, leads to ectopic expression of PLAG1 in the tumor cells. This abnormal PLAG1 expression presumably results in a deregulation of PLAG1 target genes, causing salivary gland tumorigenesis.

PLAG1 promoter swapping is also a central oncogenic event in lipoblastomas (Hibbard *et al*., 2000). Similarly, the entire PLAG1 coding sequence is placed under the control of an active and ectopic promoter region. The fusion partners discovered to date are the *hyaluronic acid synthase 2* and the *collagen 1α2* genes. These genes were never reported as translocation partners in pleomorphic adenomas, suggesting specific fusion partners for lipoblastomas.

Ectopic PLAG1 expression is not only the result of PLAG1 promoter substitution. PLAG1 overexpression has also been found in tumors without 8q12 translocations, such as pleomorphic adenomas of the salivary glands with 12q15 translocations or normal karyotype, in uterine leiomyomas, leiomyosarcomas and in smooth muscle

tumors (Astrom *et al*., 1999). This emphasizes the importance of *PLAG1* overexpression in tumorigenesis.

Recently, we proved that PLAG1 acts as a proto-oncogene with mitogenic and transforming potential (Hensen *et al*., 2002). Indeed, NIH3T3 cell lines engineered to overexpress PLAG1 were able to proliferate in medium containing 1% serum, suggesting that PLAG1 at least partially abrogates the serum requirement for the growth of NIH3T3 cells. Moreover, these cells displayed the typical features of neoplastic transformation: the cells lost cell-cell contact inhibition, showed anchorage-independent growth and were able to form tumors in nude mice.

PLAG1 is a transcription factor that contains seven canonical C_2H_2 zinc-fingers and a serine-rich C-terminus that exhibits transactivation capacities when fused to the Gal4 DNA-binding domain (Kas *et al*., 1998). PLAG1 binds a bipartite DNA element containing a core sequence, GRGGC, and a G-cluster, GGG, separated by six to eight random nucleotides (Voz *et al*., 2000). Potential PLAG1-binding sites were found in several genes and notably in promoter 3 of the human *insulinlike growth factor II* (*IGF-II*) gene. We have shown that PLAG1 is actually able to bind to the *IGF-II* promoter 3 and can stimulate its activity (Voz *et al*., 2000). Moreover, induction of PLAG1 expression in the fetal kidney cell line 293 leads to a drastic stimulation of the *IGF-II* transcript deriving from the P3 promoter. Finally, this *IGF-II* transcript is highly expressed in salivary gland adenomas overexpressing *PLAG1*, while it is not detectable in adenomas without upregulated *PLAG1* expression or in normal salivary gland (nsg) tissues. All these results indicate that *IGF-II* is a *bona fide* PLAG1 target gene, providing us with the first clue for understanding the role of PLAG1 in salivary gland tumor development.

To identify other target genes of PLAG1, we screened high-density oligonucleotide microarrays representing about 12000 human genes for transcripts whose levels were modified in pleomorphic adenomas with PLAG1 ectopic expression compared to normal salivary glands. We also monitored gene expression profiles shortly after inducing PLAG1 expression in genetically engineered human epithelial kidney 293 cell lines. Such cell lines were generated by isolating independent clones that had stably integrated a DNA fragment enabling zinc-inducible expression of PLAG1 or β-galactosidase (β-gal) used as control. Comparative evaluation of all these results allowed us to identify potential PLAG1 targets in pleomorphic adenomas. Finally, we looked for PLAG1-binding motifs in the promoter region of the identified genes to get an additional indication of whether they are direct targets of PLAG1.

Material and methods

Tumor samples and cell lines

Specimens of primary pleomorphic adenomas of the salivary glands were obtained from patients at the time of surgery. Chromosome metaphases of tumor cells were prepared from short-term primary cultures according to a routine method, and the karyotype of the tumors was determined. Tumor G27 carries a t(3;8)(p21;q12) as sole abnormality. Tumor G18 carries a t(6;8)(p21;q12), and tumor G19 carries a t(4;8)(p21;q12) among other abnormalities. Tumors F32, F36 and G30 present a normal diploid karyotype, while tumor F35 carries multiple chromosomal abnormalities. The karyotype of tumors K5773, K3149 and K3259 was not available. Normal salivary gland specimens H1, H2, H3 and G12 are biopsies corresponding to the normal salivary gland counterpart from patients presenting a pleomorphic adenoma.

The inducible PLAG1 (P1-8 and P1-32 clones) and β-gal (B-1 and B-57 clones)-expressing cell lines were obtained by stable integration of the expression vector pSAR-MT-FLAG-PLAG1 or pSAR-MT-β-gal (Morin *et al*., 1996) in the human fetal kidney epithelial cell line 293 (ATCC, CRL 1573) (Hensen *et al*., 2002). Individual colonies were selected on the basis of high zinc-inducible expression of PLAG1 or β-gal, as estimated by Western blot analysis. For expression profile analysis, cells were grown at midconfluency and either treated with 100 μM $ZnCl_2$ for 16 h or left untreated.

Preparation of RNA and Northern blot analysis

Total RNA was extracted from primary tumors and normal salivary gland biopsies using the guanidine thiocyanate method, and further purified using Rneasy spin columns (Qiagen, Valencia, CA, USA), according to the manufacturer's protocol. The total RNA from the cell lines was directly isolated using Rneasy columns. Northern blot analysis was performed according to standard procedures. For filter hybridizations, probes were radiolabeled with α-^{32}P-dCTP using the megaprime DNA-labeling system (Amersham). A 1.5 kb cDNA probe containing the complete *PLAG1* ORF was used for the detection of the PLAG1 transcripts. The human *IGF-II* exon 9 probe, common to the four different transcripts P1, P2, P3 and P4, was generated by PCR and contained nucleotides 7970-8774 of the published gene sequence (Dull *et al*., 1984) (GenBank/EMBL, accession number

X03562). As for the cellular retinoic acid-binding protein II (*CRABP-II*), the complete cDNA (M68867) obtained from RZPD (Germany) was used as probe. Finally, PCR amplicons corresponding to cytokine-like factor 1 (*CLF-1*) (1247-1614 bp of the sequence AF059293 (GenBank/EMBL)), *CRP2* (305-981 bp of D42123), *PIGF* (398-1396 bp of X54936) and $p57^{kip2}$ (1048-1316 bp of U22398) were used as probes.

Microarray analysis

Preparation of cRNA and subsequent steps leading to hybridization, scanning and data analysis were performed according to Affymetrix guidelines (Affymetrix, Santa Clara, CA, USA) (Lipshutz *et al*., 1999). Briefly, 20 μg of total RNA were converted into double-stranded cDNA using a oligo(dT) primer containing the T7 promoter. This double-stranded cDNA was then used in an *in vitro* cRNA synthesis reaction using T7 RNA polymerase and biotinylated cRNA ribonucleotides from a bioarray high-yield RNA transcript labeling kit (Enzo Diagnostics, Farmingdale, NY, USA). Biotin-labeled cRNA was purified with a RNeasy column (Qiagen) and quantified by spectrophotometry. Fragmentation of cRNA was performed at 95°C for 35 min. The quality of the procedure was evaluated by hybridizing 5 μg of cRNA to a Test3 microarray (Affymetrix). A quantity of 15 μg of the remaining fragmented cRNA was subsequently hybridized to HuGeneFL Array (Affymetrix) for 16 h at 45°C. After washing and staining, the arrays were scanned in an HP/Affymetrix scanner at 570 nm. The scanned images were analysed using Affymetrix Microarray Suite 4.2 software. The image for each GeneChip was scaled such that the average intensity value for all arrays was adjusted to a target intensity of 700; this value corresponds to the average intensity obtained for all the arrays analysed in this study. Scaled average differences values (SADVs) lower than 50 were arbitrarily set to a baseline value of 50 to avoid unrealistic level of stimulation. Therefore, for the probe sets where the values in the reference samples are at the background level, the rates of stimulation or repression could not be accurately calculated and are probably underestimated. In this situation, the fold stimulation obtained is preceded by the sign ' $\geq$ ' and the fold repression by the sign ' $\leq$ '.

Computational search for PLAG1 motifs

Search for the PLAG1-binding consensus sequence $GRGGC(N)_{6-8}GGG$ was performed using regular expressions. The search was done on both strands, without allowing any overlap of the motifs. For each gene whose promoter is not yet defined, we attempted to obtain it by retrieving from the databanks all the cDNAs (including ESTs) available for this gene. After alignment, we selected the one that starts at the most upstream part of the gene. Finally, we retrieved the sequence located directly upstream of this selected cDNA from the human genome database and considered this sequence as the putative promoter region. The random set of 50 known promoter sequences was obtained by retrieving randomly promoter sequences from the EPD database (Praz *et al*., 2002). The χ^2 test, performed to compare the distribution of the PLAG1 motifs in the promoter of the target genes versus the random genes population (see Figure 3), has been performed as described on http://www.georgetown.edu/faculty/ballc/ webtools/web_chi.html, with four degrees of freedom (genes with four or more motifs are grouped together because the numbers become too small and adjustments for continuity should be made).

Results

Expression profiles of pleomorphic adenomas compared to normal salivary glands

To decipher the molecular processes involved in PLAG1-induced oncogenesis, we compared the gene expression profiles obtained from pleomorphic adenomas displaying PLAG1 ectopic expression to those from normal salivary gland specimens. We hybridized oligonucleotide microarrays representing about 12000 genes (HuGeneFL Array, Affymetrix) with biotinylated cRNA obtained from four different RNA samples, two from normal salivary glands (H1 and G12) and two from pleomorphic adenomas (G27 and G19). These tumors carry either a t(3;8)(p21;q12) or a t(4;8)(q35;q12) translocation that leads to ectopic activation of PLAG1 expression, as judged by Northern blot analysis (data not shown). This activation results from promoter swapping between PLAG1 and β-catenin in the t(3;8) and a not yet identified translocation partner in the t(4;8). Comparison of the four samples uncovered 627 genes significantly differentially expressed. In all, 373 genes are upregulated at least three times in both tumors, while 254 genes are downregulated. Based on SWISSPROT keywords, the 627 differentially expressed genes have been classified in 14 functional categories. Nine categories are shown in Table 1, while the other five categories (signaling proteins, metabolism-related proteins, proteins of the extracellular matrix, miscellaneous proteins and proteins of unknown function) are provided online as Supplementary Tables 1a and b. The complete data set, including the values of the microarray experiments for all the four samples, is also provided online as Supplementary Tables 1c and d. The identity of these genes reveals that many aspects of cell physiology are altered in the tumors, as suggested by the change in expression

of growth factors, transcription factors, genes regulating apoptosis, cell cycle genes and signaling proteins. It is interesting to note that growth factors, growth factors receptors and growth factor-binding proteins are mainly found upregulated. Given the cellular diversity found in pleomorphic adenomas, where many different cell types including mesenchymal cells can be present, a battery of genes coding for proteins of the extracellular matrix and of the cytoskeleton were found to be differentially expressed (e.g. keratin, collagen, fibronectin, tenascin and elastin). As expected, PLAG1 transcript is only detected in pleomorphic adenomas while, in normal salivary glands, PLAG1 expression levels do not exceed the background, arbitrarily set to a value of 50 (see Material and methods for the analysis of the microarray data). Therefore, the level of PLAG1 induction in pleomorphic adenoma could not be accurately calculated and was thus estimated to be at least 14-fold (represented in the table by the sign ' ≥ '). A drastic upregulation of *IGF-II* is observed for all the three independent probe sets (the stimulations were 76-, ≥ 15-, ≥ 5-fold). Moreover, genes already identified as upregulated in pleomorphic adenomas such as the apoptosis regulator Bcl2 (Sunardhi-Widyaputra and Van Damme, 1995; Debiec-Rychter *et al*., 2001), the bone morphogenetic protein 2 (BPM2) (Kusafuka *et al*., 1998), tenascin (Sunardhi-Widyaputra and Van Damme, 1993; Shrestha *et al*., 1994) and elastin (Grosso, 1996) were stimulated 8.0-, ≥ 8.5- and ≥ 9.3-, ≥ 9.8-, ≥ 14.7-fold, respectively, which is indicative of the validity of the microarray approach.

Table 1 *Expression profiles of pleomorphic adenomas compared to normal salivary glands*

Upregulated					*Downregulated*
Growth factors					
76, ≥ 15, ≥ 5	J03242	Insulin-like growth factor IP	0.21	U50330	Bone morphogenetic protein 1
≥ 10	AF059293	Cytokine-like factor 1	≤ 0.13	L09753	Tumor necrosis factor ligand superfamily member 8
7.1	M37435	Macrophage colony-stimulating factor 1			
10.9	M60314	Bone morphogenetic protein 5			
≥ 8.5, ≥ 9.3	M22489	Bone morphogenetic protein 2[2]			
7.4	L42379	Bone-derived growth factor			
≥ 5.8	Hs.112432	Muellerian-inhibiting factor			
4.5	U78110	Neurturin			
3.6	M31682	Inhibin beta B chain			
4.2	U43030	Cardiotrophin-1			
Growth factor receptors					
20.8	M64347	Fibroblast growth factor receptor 3	≤ 0.16	M59941	Cytokine receptor common beta chain
4.6, 4.7, 4.9	M34641	Basic fibroblast growth factor receptor 1[3]			
7.9	X75958	BDNF/NT-3 growth factor receptor			
		Growth factor-binding proteins			
13.2	M34057	Latent transforming growth factor beta-binding protein 1			
6.4	Z37976	Latent transforming growth factor beta-binding protein 2			
6.4, 5.0	M62403	Insulin-like growth factor-binding protein 4[2]			
Growth regulation proteins					
28.8	0.22	Neuromodulin	0.22	AF078077	Growth arrest and DNA-damage-inducible protein GADD45 beta
5.1	U35139	Necdin			
13.6, 6.4	L13698	Growth-arrest-specific protein 1[2]			
4.0	L13720	Growth-arrest-specific protein			
Cell division and cell-cycle-related proteins					
5.5	S78187	M-phase inducer phosphatase 2	≤ 0.08	Z36714	G2/mitotic-specific cyclin F
7.1	M92287	G1/S-specific cyclin D3	≤ 0.2	M81933	M-phase inducer phosphatase 1
6.7	U73379	Ubiquitin-conjugating enzyme E2 C	0.17	AF004709	Mitogen-activated protein kinase 13
≥ 5.4	AF035811	Septin 4			
≥ 5.2	U29725	Mitogen-activated protein kinase 7			

4.5, 3.8	X57348	14-3-3 protein sigma			
Apoptosis-related proteins					
≥ 7.2	X98172	Caspase-8	0.2	AB007619	Receptor-binding cancer antigen (RCAS1)
≥ 6.2	Y09392	Tumor necrosis factor receptor superfamily member 12	0.15	U19599	BAX protein, cytoplasmic isoform delta
			≤ 0.10	U64863	Programmed cell death protein 1
Proto-oncogenes					
≥ 13.9	U65002	Zinc-finger protein PLAG1	0.25	X57110	Signal transduction protein CBL
≥ 21.6	X12949	Proto-oncogene tyrosine-protein kinase receptor ret	0.23	D78579	Nuclear hormone receptor NOR-1
≥ 21.3	U16954	Proto-oncogene AF1Q	0.10	M69199	Putative lymphocyte G0/G1 switch protein 2
≥ 11.8, 5.9	M14333	Proto-oncogene tyrosine-protein kinase FYN[2]			
4.7	D43969	Runt-related transcription factor 1			
4.3	M73554	G1/S-specific cyclin D1			
8.0	M13995	Apoptosis regulator Bcl-2			
4.6	X61118	Rhombotin-2			
Tumor suppressor					
5.6	U81992	Zinc-finger protein PLAGL1			
≥ 7.2	M22898	Cellular tumor antigen p53			
7.1, 3.5, 3.2	D64137	Cyclin-dependent kinase inhibitor 1C ($p57^{KIP2}$)[3]			
Transcriptional regulators			*Transcription factors*		
≥ 23.7, 3.7	M97915	Retinoic acid-binding protein II (CRABP-II)[2]	0.24, 0.17	Hs. 149923	X box-binding protein-1[2]
			0.04	X83877	DNA-binding protein ABP/ZF
≥ 32.9	Hs.99348	Distal-less homeo box 5 (DLX5)	0.07	M97287	DNA-binding protein SATB1
			0.18	AJ012611	Homeobox protein SIX3
≥ 11.2	L07919	Distal-less homeo box 2 (DLX2)	0.04	X76732	Nucleobindin 2 [Precursor]
≥ 6.2	X61755	Homeobox protein Hox-C5	≤ 0.11	U80987	T-box transcription factor TBX5
6.2	U66619	SWI/SNF complex 60 kDa subunit	0.14	L19871	Cyclic-AMP-dependent transcription factor ATF-3
7.5, 6.8	X77956	DNA-binding protein inhibitor ID-1[2]	0.22	AB020639	Estrogen-related receptor gamma
6.2	Hs.34853	DNA-binding protein inhibitor ID-4	0.17	M83667	CCAAT/enhancer binding protein delta
7.1	Hs.76884	DNA-binding protein inhibitor ID-3			
≥ 18.8	AB002305	Aryl hydrocarbon receptor nuclear translocator 2			
≥ 13.6	AB018303	Smad-and Olf-interacting zinc-finger protein			
≥ 7.8, 4.7	D13969	DNA-binding protein Mel-18[2]			
3.8	J03258	Vitamin D3 receptor			
5.3	AF041210	Midline 1 protein			
14.5	X96381	Ets-related protein ERM			
4.2	U15655	ETS-domain transcription factor ERF			
7.6, ≥ 9.6	AF055376	Short-form transcription factor C-MAF[2]			
11.4	X70683	Transcription factor SOX-4			
≥ 23.5	L31881	Nuclear factor 1 X-type			
13.8	X53390	Nucleolar transcription factor 1			
6.4	AB006909	Microphthalmia-associated transcription factor			
8.8	AF035528	Mothers against decapentaplegic homolog 6			
5.3	AF010193	Mothers against decapentaplegic homolog 7			

8.7	X70991	NGFI-A-binding protein 2
14.7	U49857	Transcriptional activator
7.7	L07592	Peroxisome proliferator-activated receptor delta

Based on SWISSPROT keywords, the 627 genes differentially expressed have been classified into 14 functional categories. Nine categories are shown in Table 1, while the other five categories (signaling proteins (e.g. kinases, phosphatases, receptors), metabolism-related proteins (e.g. enzymes), proteins of the extracellular matrix (e.g. collagen, fibronectin, tenascin), miscellaneous proteins and proteins of unknown function) are provided online as Supplementary Tables 1a and b. The complete data set, including the values of the microarray experiments for all the four samples, is also provided online as Supplementary Tables 1c and d. Those genes that were upregulated or downregulated at least threefold in both pleomorphic adenoma compared to both controls are considered to be differentially expressed, with the restriction that the scaled average difference value (SADV) has to be at least 200 for both tumors in the case of the upregulated genes and for both normal salivary specimens in the case of downregulated genes. The values on the left are the ratios (SADV PA1 + SADV PA2/SADV normal salivary gland 1 + SADV normal salivary gland 2) obtained for each gene. For genes represented by several probe sets, the value in the superscript above its name indicates the number of independent probe sets corresponding to this gene, and the stimulation or repression has been calculated independently for each probe set. The ≥ symbol indicates that the SADV values for normal salivary glands are at the background level (arbitrarily fixed to 50, see Material and methods), which will lead to an underestimation of the calculated stimulation, and the ≤ symbol indicates that the SADV values for the tumor are at the background level, which will lead to an underestimation of the repression

Identification of PLAG1 target genes

In order to identify in this pool of genes differentially expressed in pleomorphic adenomas, those directly under the control of PLAG1, expression profiles were monitored shortly after induction of PLAG1 expression in genetically engineered human epithelial kidney 293 cell lines. Such cell lines were generated by isolating independent clones that had stably integrated a DNA fragment enabling zinc-inducible expression of PLAG1 (clones P1-8 and P1-32) or of β-gal (clones B-1 and B-57) (Hensen *et al*., 2002). When these clones are grown in the absence of zinc, low exogenous *PLAG1* or β-gal expression is detected, while upon induction with 100 μM of zinc, the *PLAG1* and β-gal transcripts are efficiently synthesized. Induction of PLAG1 expression results in a drastic upregulation of *IGF-II* transcripts that can be already visualized after 5 h of zinc induction, but which reaches a maximal level of stimulation at 16 h (data not shown). This delay in obtaining maximal IGF-II induction could be due to the fact that the steady state of PLAG1 protein reaches a plateau at 4h and remains equal for at least 12 h (data not shown). Therefore, to get maximal induction for the potential PLAG1 targets, expression profiles were performed after 16 h of zinc induction. RNA samples were isolated from PLAG1-expressing clones P1-8 and P1-32 treated with zinc or left untreated, and from the zinc-treated β-gal clones B-1 and B-57. The comparison between treated versus untreated PLAG1-expressing clones highlighted 85 genes induced at least 2.5-fold (Figure 1, red circle). As this pool of genes could comprise genes induced by the zinc treatment, we also compared zinc-treated clones expressing PLAG1 versus β-gal clones and identified 111 genes stimulated at least threefold (blue circle). This stricter criterion compared to the one applied to the first comparison (2.5-fold) was chosen due to the fact that the untreated P1-8 and P1-32 clones expressed already low levels of PLAG1, which are likely to reduce the level of stimulation. In total, 47 genes were significantly induced in all the four comparative evaluations (Figure 1). As shown in Table 2, a drastic stimulation of the *bona fide* PLAG1 target IGF-II was observed upon PLAG1 induction; the stimulation obtained was of 39- and 32-fold when comparing PLAG1 versus β-gal clones, both zinc treated (columns G and L), and of four- and 20-fold for the comparison of treated versus untreated PLAG1-expressing clones (columns F and K). The induced genes were distributed among nine functional categories including growth factors, transcription factors, proto-oncogenes, tumor suppressors and signal transduction proteins. Note that *PLAG1* transcripts were not found upregulated because the probe set used is located in the 3' UTR of the gene, a region that is not present in the PLAG1 expression vector. This analysis also reveals a small group of downregulated genes (Table 3). Strikingly, all the 12 genes identified were repressed after 16 h of zinc induction (Table 3, columns F, G, K and L) but not at 9h (column E), suggesting that they are not direct downstream targets of PLAG1. In contrast, most of the upregulated genes were already highly induced after 9 h of zinc treatment (31 out of the 47 genes, column E in blue in Table 2).

Published in : Oncogene (2004), vol. 23, pp. 179-191.
Status : Postprint (Author's version)

Figure 1 Venn diagram of the number of genes altered in the three sets of expression profiles. For the upregulated genes, the red circle denotes the genes that are induced at least 2.5-fold when comparing treated versus untreated PLAG1-expressing clones (i.e. P1-8 + /P1-8- ≥2.5 and P1-32 + /P1-32-≥2.5). The blue circle represents the genes that are induced at least threefold when comparing treated PLAG1-expressing clones versus treated *β*-gal-expressing clones (i.e. P1-8 + /B-1 + ≥3.0 and P1-32 + /B-57 + ≥3.0). The green circle denotes the genes which are induced at least threefold when comparing pleomorphic adenoma of the salivary glands versus normal salivary glands (i.e. PA G19/nsg H1 ≥3.0 and PA G19/nsg G12≥3.0 and PA G27/nsg H1≥3.0 and PA G27/nsg G12≥3.0) and a SADV for PA G19 and PA G27≥200. For the down-regulated genes, the red circle denotes the genes that are repressed at least 2.5-fold when comparing treated versus untreated PLAG1-expressing clones (i.e. P1-8 ± ≤0.4 and P1-32 ± ≤0.4). The blue circle represents the genes that are repressed at least threefold when comparing treated PLAG1-expressing clones versus treated *β*-gal-expressing clones (i.e. P1-8 + /B-1 + ≤1/3 and P1-32 + /B-57 + ≤1/ 3). The green circle denotes the genes which are repressed at least threefold when comparing pleomorphic adenoma of the salivary glands versus nsgs (i.e. PA G19/nsg H1≤1/3 and PA G19/nsg G12≤1/3, and PA G27/nsg H1≤1/3 and PA G27/nsg G12≤1/3) and a SADV for nsg H1 and nsg G12≥200

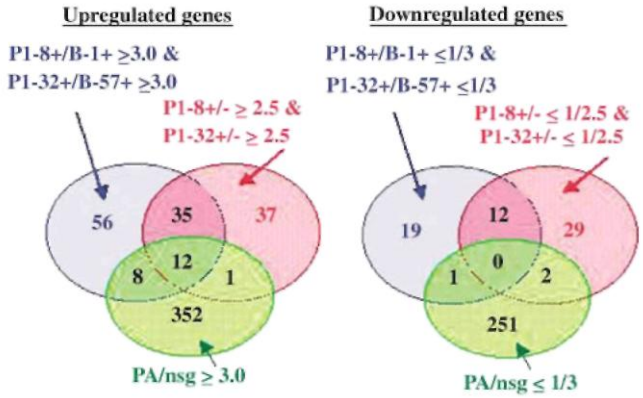


Table 2 *Identification of 47 genes consistently and significantly induced by PLAG1*

AB		*C*	*D*	*E*	*F*	G	*H*	*IJ*		*K*	*L*	*M*	*N*	*O*
B1+	*P8-*	*P8+*	P8+9h	P8+9h/P8-	*P8+/-*	*P8+/B1+*	*B57+*	*P32-*	*P32+*	*P32+/-*	*P32+/B57+*	*Accession*	*Genes*	*BS*
Growth factors														
144	1462	5604	*3505*	***2.4***	3.8	38.9	132	208	4164	20.0	31.6	J03242	Insulin-like growth factor II (IGF-II)	8
211	169	2319	*889*	***5.3***	13.8	11.0	72	107	1249	11.6	17.3	AF059293	Cytokine-like factor-1(CLF-1)	8
130	108	462	*173*	***1.6***	4.3	3.5	234	314	836	2.7	3.6	L42379	Bone-derived growth factor (BPGF-1)	3
50	50	864	*50*	***1.0***	17.3	17.3	50	52	150	2.9	3.0	J00117	Choriogonadotropin beta chain (CGB)	NA
376	296	1193	*976*	***3.3***	4.0	3.2	145	50	461	9.2	3.2	AF024710	Vascular endothelial growth factor (VEGF)	1
96	120	301	*522*	***4.4***	2.5	3.1	50	74	328	4.4	6.6	X54936	Placental growth factor (PIGF)	1
Transcriptional regulators														
153	1363	9166	*2463*	***1.8***	6.7	60.0	53	1240	7512	6.1	141.2	M97815	Retinoic acid-binding protein II (CRABP-II)	1
82	842	4589	*1771*	***2.1***	5.5	55.9	50	517	2485	4.8	49.7	M97815		
50	50	256	*50*	***1.0***	5.1	5.1	50	50	150	3.0	3.0	D12765	Adenovirus E1 A enhancer binding protein (E1 A-1)	*1*
67	225	834	*1195*	***5.3***	3.7	12.4	53	59	314	5.3	5.9	U08015	Nuclear factor of activated T-cells (NF-ATc)	*2*
164	191	1127	*210*	***1.1***	5.9	6.9	196	135	628	4.7	3.2	AB021663	Cyclic-AMP-dependent transcription factor ATF-5	3
258	390	994	*717*	***1.8***	2.5	3.9	170	417	1048	2.5	6.2	U66619	SWI/SNF complex 60kDa submit	2
50	50	591	*58*	***1.2***	11.8	11.8	50	50	167	3.3	3.3	AF055009	OASIS protein	2
Oncogene														
52	57	373	*103*	***1.8***	6.5	7.2	50	52	195	3.7	3.9	M13995	Apoptosis regulator Bcl-2	5
Tumor suppressor														
470	138	1303	*6281*	***4.5***	9.4	27.7	279	694	2997	4.3	10.8	U22398	Cyclin-dependent kinase	6

Published in : Oncogene (2004), vol. 23, pp. 179-191.
Status : Postprint (Author's version)

	6	4											inhibitor 1C ($p57^{KIP2}$)
95	565	2848	*2146*	***3.8***	5.0	30.1	104	113	328	2.9	3.1	U22398	
139	68	837	*899*	***13.2***	12.3	6.0	72	121	372	3.1	5.1	D64137	
Proteins involved in signaling													
124	55	1086	*294*	***5.4***	19.8	8.8	187	179	648	3.6	3.5	BC027933	Mitogen-activated protein kinase11 (MAPK11) 0
50	222	910	*433*	***2.0***	4.1	18.2	91	50	311	6.2	3.4	U09303	Ephrin-B1 (EFNB1/EPLG2) 3
301	664	2532	*1174*	***1.8***	3.8	8.4	240	494	1365	2.8	5.7	AF100153	Connector enhancer of KSR-like (CNK1) 2
342	630	3041	*2235*	***3.6***	4.8	8.9	50	471	1280	2.7	25.6	W27541	Cdc42 effector protein4 (CEP4) 3
119	286	1492	*944*	***3.3***	5.2	12.5	64	176	850	4.8	13.3	AF027406	Serine/threonine protein kinase 23 (STK23/MSSK1) 3
112	50	378	*193*	***3.9***	7.6	3.4	50	58	164	2.8	3.3	X77533	Activin receptor type IIB 4
50	64	486	*193*	***3.0***	7.6	9.7	50	55	160	2.9	3.3	AF025533	Leucocyte immunoglobulin-like receptor-3 0
50	50	234	*85*	***1.7***	4.7	4.7	50	73	256	3.5	5.1	X66171	Leukocyte immunoglobulin-like receptor CMRF35 1
50	169	475	*419*	***2.5***	2.8	9.5	50	50	253	5.1	5.1	NM_058178	Neuronal Pentrax in Receptor (NPTXR) 6
50	50	176	*171*	***3.4***	3.5	3.5	50	50	205	4.1	4.1	U19261	TNF receptor associated factorl 1
70	58	222	*122*	***2.1***	3.8	3.2	50	58	181	3.1	3.6	U19261	
50	50	342	*61*	***1.2***	6.8	6.8	50	56	352	6.2	7.0	M33552	Lymphocyte-specific protein 1 (LSP1) 2
209	428	1201	*1040*	***2.4***	2.8	5.7	196	229	730	3.2	3.7	AF041434	Protein tyrosine phosphatasetype4A,3 (PTP4A3) NA
Proteins involved in metabolism													
237	300	1953	*1765*	***5.9***	6.5	8.2	55	102	382	3.8	6.9	AB003791	Keratan sulfate Gal-6-sulfotransferase 1
50	57	601	*528*	***9.2***	10.5	12.0	50	165	611	3.7	12.2	M55153	Transglutaminase 2 (TGM2) NA
Extra cellular matrix/cytoskeleton proteins													
50	50	343	*234*	***4.7***	6.9	6.9	50	55	280	5.1	5.6	L41162	Collagen alpha 3 (IX) chain 3
50	50	2628	*1016*	***20.3***	52.6	52.6	50	50	1089	21.8	21.8	U53204	Plectin 1 3
50	228	583	*612*	***2.7***	2.6	11.7	50	50	410	8.2	8.2	Y00503	Keratin, type 1 cytoskeletal 19 3
77	98	740	*55*	***0.6***	7.5	9.6	50	50	157	3.1	3.1	S76756	Microtubule associated protein 0
50	187	1325	*415*	***2.2***	7.1	26.5	57	102	1092	10.7	19.0	M21984	Tropon in T3, fast skeletal muscle isoform β 1
645	899	4144	*3304*	***3.7***	4.6	6.4	594	642	4099	6.4	6.9	AJ012737	Filamin, muscleiso form Miscellaneous 2
80	194	1360	*1247*	***6.4***	7.0	16.9	50	50	1386	27.7	27.7	D42123	Cysteine-rich protein 2 (CRP2/ESP1) 5
50	50	705	*207*	***4.1***	14.1	14.1	50	50	307	6.1	6.1	U45448	Purino receptor P2X1 2
50	50	235	*240*	***4.8***	4.7	4.7	50	50	191	3.8	3.8	U41518	Aquapo rin-CH IP 0
357	264	1241	*1415*	***5.4***	4.7	3.5	74	50	369	7.4	4.9	U07364	Inward rectifier potassium channel 4 1
50	50	613	*313*	***6.3***	12.3	12.3	68	51	539	10.6	7.9	L37792	Syntaxin 1A 2
841	604	3743	*3717*	***6.2***	6.2	4.4	500	1196	4352	3.6	8.7	AF022813	Tetraspan (NAG-2) NA
318	413	1183	*540*	***1.3***	2.9	3.7	381	441	1331	3.0	3.5	M20469	Brain-type clathrin light-chain b 2
50	79	931	*484*	***6.1***	11.7	18.6	70	98	519	5.3	7.4	M96759	ROD outer segment membrane protein 1 6
243	288	1168	*586*	***2.0***	4 .1	4.8	185	208	594	2.9	3.2	M87068	Calcium-binding protein S100A2 1
Unknown													
50	50	363	*298*	***6.0***	7.3	7.3	50	56	191	3.4	3.8	AB023171	KIAA09 54 0
50	50	257	*89*	***1.8***	5.1	5.1	50	50	181	3.6	3.6	M60614	Wilms tumor-associated 0

Published in : Oncogene (2004), vol. 23, pp. 179-191.
Status : Postprint (Author's version)

91	179	1115	*923*	***5.2***	6.2	12.2	94	106	338	3.2	3.6	AI391564	protein WIT-1 tg16b02.x1 Homo sapiens cDNA	NA
50	119	543	*106*	***0.9***	4.6	10.9	166	58	502	8.7	3.0	AA165701	zo75g08.s1 Homo sapiens cDNA	1

The 47 genes induced at least threefold when comparing PLAG1 versus *β*-galactosid ase-expressing clones, both zinc-treated (columns G and L) and at least 2.5-fold when comparing treated versus untreated PLAG1-expressing clones (columns F and K) are listed. The columns include the following information : the scaled average differences values (SADVs) for the probe sets in the different conditions (columns A-D, H-J); the accession number of the gene corresponding to the probe set (column M); the number of PLAG1-binding motifs found in the first 1000bp of their putative promoter region (column O), numbers in italics meaning that the first 1000bp of the promoter were not available in its entirety (only 280 and 158 bp were available for NF-ATc and E1 A-1. respectively), and NA meaning that the promoter was not available at all. SADVs of less than 50 were arbitrarily set to a baseline value of 50 to avoid unrealistic levels of stimulation (see Material and methods). The data of this table came from two independent microarray experiments, the first one performed on RNAs from samples A to D treated at the same moment, and the second from samples H to J. For this reason, we compared P1-8 with B-1 and P1-32 with B-57

Table 3 *Identification of 12 genes consistently and significantly repressed by PLAG1*

AB		*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*N*	*O*
B1+	*P8-*	*P8+*	P8+9h	P8-/P8+9h	*P8-/+*	*B1+/P8+*	*B57+*	*P32-*	*P32+*	*P32-/+*	*B57+/P32+*	*Accession*	*Genes*	*BS*
Growth factors														
287	279	50	*279*	***1.0***	5.6	5.7	155	157	50	3.1	3.1	Y07867	Pirin	0
Proteins involved in metabolism														
174	170	50	*289*	***0.6***	3.4	3.5	400	163	50	3.3	8.0	AF039652	Ribonuclease H type II	1
752	685	248	*740*	***0.9***	2.8	3.0	606	499	130	3.8	4.7	S71018	Peptidyl-prolyl-isomerase C	0
459	286	50	*347*	***0.8***	5.7	9.2	215	152	50	3.0	4.3	M33494	Triptase -beta 1	4
Miscellaneous														
358	257	50	*458*	***0.6***	5.1	7.2	483	762	50	15.2	9.7	AJ133769	Nuclear transport receptor	1
709	429	165	*518*	***0.8***	2.6	4.3	347	344	96	3.6	3.6	X04325	Gap junction beta-1 protein	3
237	327	50	*310*	***1.1***	6.5	4.7	247	406	68	6.0	3.6	U90546	Butyrophilin	0
172	254	50	*188*	***1.3***	5.1	3.4	179	169	50	3.4	3.6	AB015332	Neighbor of A-kinase anchoring protein 95	1
Unknown														
317	377	56	*488*	***0.8***	6.8	5.7	485	365	50	7.3	9.7	AI133727	Habcs0217 homo sapiens cDNA	0
297	329	75	*476*	***0.7***	4.4	3.9	423	355	50	7.1	8.5	AB020674	mRNA for KIAA0867 protein	NA
161	197	50	*480*	***0.4***	3.9	3.2	326	320	50	6.4	6.5	AI344681	Qp09h03.x1 homo sapiens cDNA	1
192	252	50	*286*	***0.9***	5.0	3.8	228	307	50	6.1	4.6	AL031432	Hypothetical protein DJ465N24.1	NA

The 12 genes repressed at least threefold when comparing PLAG1 versus *β*-galactosidase -expressing clones, both zinc-treated (columns G and L), and at least 2.5-fold when comparing treated versus untreated PLAG1-expressing clones (columns F and K) are listed. Columns include the following information : the scaled average difference values (SADVs) for the probe sets in the different conditions (columns A-D, H-J); the accession number of the gene corresponding to the probe set (column M); the number of PLAG1-binding motifs found in the first 1000 bp of their putative prom oter region (column O), NA meaning that the promoter sequences were not available. SADVs of less than 50 were arbitrarily set to a baseline value of 50 to avoid unrealistic levels of repression (see Material and methods). The data of this table came from two independent microarray experiments, the first one performed on RNAs from samples A to D treated at the same moment, and the second from samples H to J. For this reason, we compared P1-8 with B-1 and P1-32 with B-57

Identification of candidate PLAG1 target genes in pleomorphic adenomas

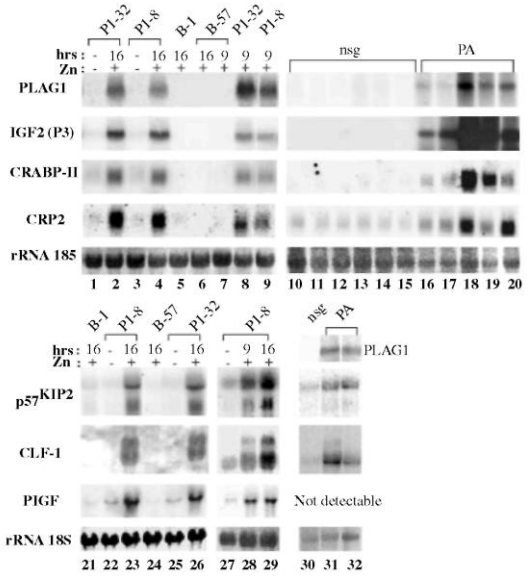
The comparison of the two sets of expression profiles (tumors *versus* normal salivary glands and PLAG1-expressing cells *versus* control cells) allowed us to identify candidate PLAG1 targets in pleomorphic adenomas (Figure 1). A total of 12 genes were found to be stimulated in PLAG1-expressing cells as well as in the tumors (genes in green in Table 2). These genes are coding for the IGF-II, the CLF-1, the bone-derived growth factor 1 (BPGF-1), the CRABP-II, the SWI/ SNF complex 60 kDa subunit, the apoptosis regulator Bcl-2, the cyclin-dependent kinase inhibitor 1C ($p57^{KIP2}$), ephrin B-1, the neuronal pentraxin receptor, the collagen alpha 3 (IX) chain, the muscle isoform of filamin and the tetraspan (NAG-2). In contrast, none of the 12 downregulated genes have their expression significantly altered in the pleomorphic adenoma of the salivary glands.

Published in : Oncogene (2004), vol. 23, pp. 179-191.
Status : Postprint (Author's version)

Validation of the oligonucleotide microarray results

In order to confirm the microarray analyses, we performed Northern blot analyses using six putative target genes as probes, that is, *IGF-II, CRABP-II, CRP2, p57^{KIP2}, CLF-1* and *PIGF*. As shown in Figure 2, expression levels of the six transcripts are extremely low or undetectable in zinc-treated *β*-gal-expressing clones (B-1+ and B-57+ ; lanes 5, 6, 7, 21 and 24). The transcript levels increase slightly in untreated PLAG1-expressing clones (P1-8— and P1-32—; lanes 1, 3, 22, 25 and 27). This increase might be due to a leaky expression of PLAG1 in this system, which can be visualized after a longer exposition of the blots (data not shown). A strong induction in the expression of all the six transcripts is solely seen when PLAG1 expression is induced (lanes 2, 4, 23, 26 and 29), as predicted by the microarray results. High induction of the six targets is also obtained after 9 h of zinc treatment, which reaches approximately half of the stimulation observed at 16 h (lanes 8, 9 and 28). Assessment of the transcript levels in the tumors also correlated with the microarray data as the transcripts for *PLAG1*, *CRABP-II*, *IGF-II* and *CLF-1* are solely detected in pleomorphic adenoma, while no expression is observed in normal salivary glands. The p57^{kip2} transcript is also overexpressed in the tumors compared to normal salivary glands, but to a lesser extent (about 2.1-fold stimulation). As for *CRP2*, its transcript is on average overexpressed 3.2-fold in tumors compared to normal salivary glands, which is roughly similar to the 2.2-fold stimulation obtained by the microarray analysis. Finally, the *PIGF* transcript was not detectable, either in tumors or in normal salivary glands, which is in agreement with the microarray values close to the background.

Figure 2 Validation of the oligonucleotides microarray results by Northern blot analyses. Northern blot analysis of total RNA isolated from genetically engineered human epithelial kidney 293 cell lines that have stably integrated a DNA fragment enabling zinc-inducible expression of PLAG1 (clones P1-8 and P1-32) or of *β*-gal (clones B-1 and B-57) grown without (-) or with (+) 100 μM $ZnCl_2$ for the indicated times (lanes 1-9, 21-29). Also depicted are Northern blot analyses of total RNAs isolated from pleomorphic adenomas of the salivary glands (PA) corresponding to biopsies F32 (lane 16), F36 (lane 17), F35 (lane 18), G30 (lane 19), G18 (lanes 20 and 32) and from biopsies K5773, K3149 and K3259 pooled together (lane 31). The total RNA from normal salivary gland (nsg) H3 (lane 10), H1 (lane 30) and H2 (lanes 11-15 corresponding to different RNA preparations obtained from different part of the biopsy H2) was used for comparison. Blots were hybridized sequentially with different ^{32}P-labeled human probes, as indicated on the left side of the figure. The sizes of the transcripts detected are the following: 7.5 kb for *PLAG1*; 6.0 kb for *IGF-II* corresponding to the P3 transcript; 1.2 kb for *CRABP-II*; 1.4kb for CRP2; 1.7 and 1.4kb for *p57^{kip2}* in the PLAG1-expressing cell lines, while only the 1.7 transcript is detected in the tumors; 1.7 and 1.4 kb for *CLF-1* in the 293 cell lines, while only the 1.4 transcript is detected in the tumors and, finally, 1.7 kb for *PIGF*. Methylene blue-stained rRNA levels demonstrate similar loading and RNA quality in each lane



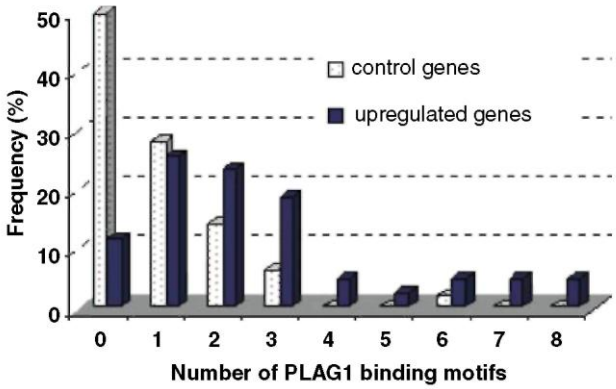
Search for PLAG1 motifs in the PLAG1 target genes

To get a clue whether the 47 upregulated genes found in this study could be direct targets of PLAG1, we exploited the fact that PLAG1 is a sequence-specific DNA-binding protein that interacts with the motif GRGGC$(N)_{6-8}$GGG. Therefore, we scanned the promoter regions of these genes for the presence of this binding motif, banning any mismatches. As the promoter region of most of these genes was not yet defined, we attempted to obtain them by retrieving from the human genome database the sequences located directly upstream

of the cDNAs with the 5' end starting at the most upstream part of the genes (see Material and methods). However, it remains possible that some sequences do not correspond to the real promoter, which could be located upstream of an unidentified exon. As shown in Table 2 (column O), the 47 upregulated genes often harbor several PLAG1-binding sites in the first 1000 bp of their putative promoter regions, with an average of 2.4 binding sites per upregulated gene. As demonstrated by the χ^2 test, this value is significantly higher than the mean obtained for a random population of 50 promoters (1.1 binding site per gene). The distribution of the motifs is also clearly different in the upregulated genes compared to the random population, as illustrated in Figure 3. For example, 50% of the control promoters do not display any PLAG1-binding sites, while this percentage drops to 10% for the upregulated genes. On the contrary, 40% of the upregulated genes contain at least three PLAG1 motifs in comparison to 8% in the random population of promoters.

The downregulated genes were also scanned for the presence of the *PLAG1* motifs. In all, 11 consensus sequences were found in the 10 promoter sequences available, defining a frequency of 1.1 consensus per gene, which is identical to the one obtained for the random population of 50 promoters.

Figure 3 Histograms illustrating the abundance of PLAG1 motifs in the promoter regions of the upregulated genes. Histogram depicting the percentage of genes containing zero to eight PLAG1 consensus sequences ($GRGGC(N)_{6-8}GGG$) in their promoter for each population (the 47 upregulated PLAG1 targets and the random pool of 50 control promoters) . The χ^2 test demonstrates a > 99.75% chance that there is a significant difference in the number of *PLAG1* motifs between upregulated and control genes



Discussion

PLAG1 is a proto-oncogene recently discovered, whose ectopic expression can trigger the development of pleomorphic adenomas of the salivary glands (Kas *et al*., 1997) and of lipoblastomas (Hibbard *et al*., 2000).

As PLAG1 is a transcription factor, its ectopic expression presumably leads to the deregulation of target genes. Their identification is therefore essential for the understanding of PLAG1-induced tumorigenesis. To date, only one PLAG1 target has been identified, the IGF-II gene (Voz *et al*., 2000). Several lines of evidence indicate that IGF-II is a direct target of PLAG1. First, PLAG1 is able to bind to the promoter 3 of *IGF-II*, which contains eight consensus-binding sites, and activates its promoter activity efficiently. Secondly, *IGF-II* transcripts deriving from the P3 promoter are highly expressed in salivary gland adenomas over-expressing *PLAG1*. In contrast, they are undetectable in adenomas without abnormal *PLAG1* expression and in the normal salivary gland tissue. Finally, a drastic upregulation of the *IGF-II* transcript originating from the P3 promoter is observed shortly after inducing PLAG1 expression in the fetal kidney cell line 293 (Hensen *et al*., 2002).

In order to identify other PLAG1 targets, we analysed the expression profiles of about 12000 genes after conditional induction of PLAG1 expression in 293 cell lines. Such cell lines have stably integrated a DNA fragment enabling zinc-inducible expression of PLAG1 or of *β*-gal (Hensen *et al*., 2002). In the absence of zinc, they expressed low levels of PLAG1 and *β*-gal proteins, while a strong overexpression is observed upon treatment with zinc chloride. PLAG1 induction leads to a strong induction of the *IGF-II* transcript which can already be visualized after about 5 h, but which peaks at 16 h. This quite tardy peak is presumably due to the fact that in this system, the level of PLAG1 protein reaches a plateau at 4h and remains equal for at least 12 h (data not shown). For most of the microarray analyses, we therefore extracted RNA after 16 h of induction, reasoning

that, at that time, we should have maximal induction for PLAG1 targets, with a still large proportion of direct targets. This choice was reinforced by the microarray analyses performed on fibroblasts constitutively expressing a MYC-ER protein (Coller *et al*., 2000). In the presence of tamoxifen, the MYC-ER protein, sequestered in the cytoplasm, is transported into the nucleus, where it can activate the transcription of Myc target genes. In this system, where the MYC-ER protein does not have to be produced, most of the MYC targets detected after 9 h of tamoxifen treatment were still direct targets.

Expression profile analyses allowed us to identify 47 genes that are consistently induced upon zinc treatment of PLAG1-expressing clones, and also overexpressed in zinc-treated PLAG1 clones compared to the treated *β*-gal-expressing clones. As illustrated in Figure 1, both comparisons were essential to identify only genuine PLAG1 targets. Indeed, the comparison between treated *versus* untreated PLAG1-expressing clones highlighted, on top of the 47 PLAG1 target genes, 38 genes, some of which are presumably induced by the zinc treatment alone. Surprisingly, comparisons between PLAG1 *versus β*-gal-expressing clones, both zinc treated, reveal 64 genes not obtained by the comparisons of treated *versus* untreated PLAG1 clones. These genes could be differentially expressed due to clone disparities and/or because, even in the absence of zinc, PLAG1-expressing clones proliferate more rapidly than the *β*-gal-expressing clones. It is possible that the faint leaky PLAG1 expression obtained in the absence of zinc is sufficient to slightly stimulate cell proliferation, resulting in changes of expression of some growth-related genes.

The search for PLAG1-binding sites in these 47 upregulated genes indicated that they contain significantly more binding sites than a random population of promoters (Figure 3). This is a strong indication that at least some of them are direct PLAG1 targets. This hypothesis is reinforced by the fact that expression profiles and Northern blot analysis indicate that a majority of them (30 out of 47) are already highly stimulated after 9 h of zinc induction. Contrary to the upregulated genes, the 12 downregulated genes do not appear to be direct PLAG1 targets. Indeed, the promoter of these genes does not contain significantly more PLAG1 motifs compared to the random promoter population; secondly, none of them are already repressed after 9h of zinc induction. Based on this study, it thus appears that PLAG1 acts essentially as an activator of transcription. This is in accordance with the data showing that the C-terminal part of PLAG1 acts as a transactivation domain when fused to the GAL4 DNA-binding domain (Kas *et al*., 1998).

Interestingly, one of the largest classes identified as upregulated PLAG1 targets consists of growth factors such as IGF-II, the CLF-1 and the BPGF-1. These three cytokines are also upregulated in pleomorphic adenomas of the salivary glands displaying PLAG1 over-expression, suggesting that they are PLAG1 targets in the tumors. We can thus postulate that one of the roles of PLAG1 in tumor formation is to influence cell proliferation via induction of growth factors. The role of IGF-II as a potent stimulator of cell proliferation during embryonic development and tumorigenesis is well established (DeChiara *et al*., 1990; Baker *et al*., 1993; Toretsky and Helman, 1996; Burns and Hassan, 2001), and it is likely that at least part of the PLAG1 oncogenic potential passes through the induction of IGF-II expression. This is supported by our transformation studies, where we showed that PLAG1 is able to induce the neoplastic transformation of NIH3T3 but not of R^-cells, which are devoid of the IGFI-R receptor, the main mediator of IGF-II effect (Sell *et al*., 1994; Hensen *et al*., 2002). The role of the CLF-1 on cell proliferation is not yet documented. This recently discovered growth factor associates with the cardiotrophin-like cytokine (CLC) to form a heterodimeric cytokine that binds and activates the trimeric complex constituted of the CNTFR receptor, the leukemia-inhibitory factor receptor and gpl30 (Elson *et al*., 1998, 2000). Although the signaling cascade resulting from the binding to the receptor is already well defmed (Lelievre *et al*., 2001), the role of the CLF-1 factor in terms of cellular responses has not yet been investigated. Until now, the only information available comes from the $CLF\text{-}1^{-/-}$ mice that have reduced numbers of hemopoietic progenitor cells, fail to suckle and die within 24 h after birth (Alexander *et al*., 1999). Whole mount *in situ* hybridization reveals CLF-1 expression at several sites in the developing embryo and notably in the secretory buds and ducts of the submandibular salivary glands at 14.5 and 18.5 dpc (Alexander *et al*., 1999). This brings the hypothesis that PLAG1 could be one of the regulators of CLF-1 expression in the developing salivary gland, controlling the salivary gland organogenesis via this cytokine. As for the BPGF-1, also called quiescin Q6, this gene is specifically expressed when cultured human lung fibroblasts begin to leave the proliferative cycle and enter quiescence (Coppock *et al*., 1993). BPGF-1 thus seems to be associated with growth inhibition, which presumably interferes with tumor formation. The significance of BPGF-1 overexpression in pleomorphic adenomas and also in some malignant breast cell lines (Coppock *et al*., 2000) is therefore not yet understood. However, Coppock *et al*. propose that one potential function of quiescins might be to ensure viability under stressful conditions that might otherwise lead to apoptosis. On the other hand, it appears from our microarray data that PLAG1 induces not only target genes able to promote tumor formation, but also genes that in contrary seem to interfere with it. The most striking example is the induction by PLAG1 of the putative tumor suppressor $p57^{kip2}$. This factor is indeed a tight-binding inhibitor of several G1 cyclin/Cdk complexes, resulting

in a negative regulation of cell proliferation (Lee *et al*., 1995; Matsuoka *et al*., 1995). Diminished expression of $p57^{kip2}$ has been found in a variety of tumors (Oya and Schulz, 2000; Schwienbacher *et al*., 2000; Ito *et al*., 2002a) and, to our knowledge, pleomorphic adenoma is one of the few cases with hepatoblastomas (Hartmann *et al*., 2000) and thyroid neoplasms (Ito *et al*., 2002b) of a $p57^{kip2}$ overexpression. Another example of PLAG1-induced genes with growth-inhibitory potential is the gene coding for the CRABP-II. CRABP-II potentiates the transcriptional activity of the retinoic acid receptor via its ability to channel retinoic acid to the receptor (Budhu and Noy, 2002). Overexpression of CRABP-II in MCF-7 mammary carcinoma cells dramatically enhances their sensitivity to retinoic acid-induced growth inhibition. Conversely, diminished expression of CRABP-II renders MCF-7 and squamous cell carcinoma SCC25 much less sensitive to retinoic acid-mediated inhibition of proliferation (Vo and Crowe, 1998; Budhu and Noy, 2002). These seemingly contradictory actions of PLAG1 targets on tumor formation could be one of the reasons as to why PLAG1 triggers only the formation of benign tumors such as pleomorphic adenomas of the salivary glands and lipoblastomas. We could hypothesize that the balance between PLAG1 targets promoting and inhibiting tumor formation is in favor of a weak induction of cell proliferation as observed in pleomorphic adenoma (Zhu *et al*., 1997; Frade Gonzalez *et al*., 2001). The low rate of proliferation, scarcely propitious to accumulate additional mutations necessary for the tumor to become malignant, could explain the fact that pleomorphic adenomas of the salivary glands rarely give rise to a malignant tumor.

Our expression profile analyses also reveal that PLAG1 could be involved in a variety of cellular processes, and notably in vasculogenesis and/or angio-genesis. Indeed, the identification as upregulated PLAG1 targets of the vascular endothelial growth factor (VEGF) and the placental growth factor (PIGF), both known as potent mitogen for endothelial cells (Conway *et al*., 2001), is highly suggestive of a role of PLAG1 in controlling blood vessel formation. This hypothesis is reinforced by the upregulation by PLAG1 of Ephrin B1, which is involved in the sprouting of new vessels (Conway *et al*., 2001). Such hypothesis will be verified in mice with a targeted disruption of the PLAG1 gene, as well as in transgenic mice with conditional induction of PLAG1 expression. In the same way, the biological significance of the other PLAG1 targets will be evaluated in these mouse models. The identification of a number of PLAG1 targets in this study, besides giving us insights into understanding how PLAG1 activation contributes to salivary gland tumorigenesis, will also be a valuable and indispensable tool for the molecular comprehension of PLAG1 function during embryogenesis.

Abbreviations

PLAG1, pleomorphic adenoma gene 1; IGF-II, insulin-like growth factor II; PA, pleomorphic adenoma of the salivary glands; nsg, normal salivary gland; $p57^{kip2}$, cyclin-dependent kinase inhibitor 1C; CRABP-II, cellular retinoic acid-binding protein II; CLF-1, cytokine-like factor 1; CRP2, cysteine-rich protein 2; PIGF, placental growth factor; SADV, scaled average difference value.

Acknowledgements

We gratefully acknowledge N Agten, Grigory Dubois and E Meyen for technical assistance, and Bernard Peers and Sabine Tejpar for critical review of the manuscript. We also like to thank the head and neck surgeons Professor P Delaere, Professor V Vander Poorten, Dr Chr Vanclooster, Dr E Annys, Department of Otorhinolaryngology, Head and Neck Surgery and Professor J Schoenaers, Department of Oral and Maxillofacial Surgery, University Hospitals, Faculty of Medicine, Catholic University of Leuven, Leuven, Belgium for cooperation in providing the surgical specimens. We are also grateful to Professor R Sciot, Pathology Department University Hospitals, Faculty of Medicine, Catholic University of Leuven, Leuven, Belgium for histological diagnosis and Prof A Hagemmeijer and Dr M Debiec-Rychter, Laboratory for Cytogenetics and Molecular Genetics of Human Malignancies, Catholic University of Leuven, Leuven, Belgium for the cytogenetic analyses. This work was supported by the 'Geconcerteerde Onderzoekacties 2002-2006', the 'Fonds voor Wetenschappelijk Onderzoek Vlaanderen (FWO)' and the 'Fonds National de la Recherche Scientifique' (Télévie 7.4533.01 and 7.4552.02; Crédit Chercheur 1.5.108.02). MLV is a Chercheur Qualifié and HP is a Collaborateur Scientifique from the Fonds National de la Recherche Scientifique. IVV is an aspirant fellow and YM a postdoctoral fellow of the FWO. JM is a postdoctoral researcher and BDM is a full professor at the KUL, Belgium.

References

Alexander WS, Rakar S, Robb L, Farley A, Willson TA, Zhang JG, Hartley L, Kikuchi Y, Kojima T, Nomura H, Hasegawa M, Maeda M, Fabri L, Jachno K, Nash A, Metcalf D, Nicola NA and Hilton DJ. (1999). *Curr. Biol.*, 9, 605-608.

Astrom AK, Voz ML, Kas K, Roijer E, Wedell B, Mandahl N, Van de Ven W, Mark J and Stenman G. (1999). *Cancer Res.*, 59, 918-923.

Baker J, Liu JP, Robertson EJ and Efstratiadis A. (1993). *Cell,* 75, 73-82.

Budhu AS and Noy N. (2002). *Mol. Cell. Biol,* 22, 2632-2641.

Burns JL and Hassan AB. (2001). *Development,* 128, 3819-3830.

Coller HA, Grandori C, Tamayo P, Colbert T, Lander ES, Eisenman RN and Golub TR. (2000). *Proc. Natl. Acad. Sci. USA,* 97, 3260-3265.

Conway EM, Collen D and Carmeliet P. (2001). *Cardiovasc. Res.,* 49, 507-521.

Coppock D, Kopman C, Gudas J and Cina-Poppe DA. (2000). *Biochem. Biophys. Res. Commun.,* 269, 604-610.

Coppock DL, Kopman C, Scandalis S and Gilleran S. (1993). *Cell Growth Differ.,* 4, 483-493.

Debiec-Rychter M, Van Valckenborgh I, Van den Broeck C, Hagemeijer A, Van de Ven WJ, Kas K, Van Damme B and Voz ML. (2001). *Lab. Invest,* 81, 1289-1297.

DeChiara TM, Efstratiadis A and Robertson EJ. (1990). *Nature,* 345, 78-80.

Dull TJ, Gray A, Hayflick JS and Ullrich A. (1984). *Nature,* 310, 777-781.

Elson GC, Graber P, Losberger C, Herren S, Gretener D, Menoud LN, Wells TN, Kosco-Vilbois MH and Gauchat JF. (1998). *J. Immunol,* 161, 1371-1379.

Elson GC, Lelievre E, Guillet C, Chevalier S, Plun-Favreau H, Froger J, Suard I, de Coignac AB, Delneste Y, Bonnefoy JY, Gauchat JF and Gascan H. (2000). *Nat. Neurosci.,* 3, 867-872.

Frade Gonzalez C, Garcia-Caballero T, Lozano Ramirez A and Labella Caballero T. (2001). *Acta Otorrinolaringol. Esp.,* 52, 456-460.

Grosso LE. (1996). *J. Oral Pathol. Med.,* 25, 5-9.

Hartmann W, Waha A, Koch A, Goodyer CG, Albrecht S, von Schweinitz D and Pietsch T. (2000). *Am. J. Pathol.,* 157, 1393-1403.

Hensen K, Van Valckenborgh IC, Kas K, Van de Ven WJ and Voz ML. (2002). *Cancer Res.,* 62, 1510-1517.

Hibbard MK, Kozakewich HP, Dal Cin P, Sciot R, Tan X, Xiao S and Fletcher JA. (2000). *Cancer Res.,* 60, 4869- 4872.

Ito Y, Takeda T, Sasaki Y, Sakon M, Yamada T, Ishiguro S, Imaoka S, Tsujimoto M, Monden M and Matsuura N. (2002a). *Liver, 11,* 145-149.

Ito Y, Yoshida H, Nakano K, Kobayashi K, Yokozawa T, Hirai K, Matsuzuka F, Matsuura N, Kuma K and Miyauchi A. (2002b). *Int. J. Mol. Med.,* 9, 373-376.

Kas K, Voz ML, Hensen K, Meyen E and Van de Ven WJ. (1998). *J. Biol. Chem.,* 273, 23026-23032.

Kas K, Voz ML, Roijer E, Astrom AK, Meyen E, Stenman G and Van de Ven WJ. (1997). *Nat. Genet.,* 15, 170-174.

Kusafuka K, Yamaguchi A, Kayano T, Fujiwara M and Takemura T. (1998). *Virchows Arch.,* 432, 247-253.

Lee MH, Reynisdottir I and Massague J. (1995). *Genes Dev.,* 9, 639-649.

Lelievre E, Plun-Favreau H, Chevalier S, Froger J, Guillet C, Elson GC, Gauchat JF and Gascan H. (2001). *J. Biol. Chem., 216,* 22476-22484.

Lipshutz RJ, Fodor SP, Gingeras TR and Lockhart DJ. (1999). *Nat. Genet.,* 21, 20-24.

Matsuoka S, Edwards MC, Bai C, Parker S, Zhang P, Baldini A, Harper JW and Elledge SJ. (1995). *Genes Dev.,* 9, 650-662.

Morin PJ, Vogelstein B and Kinzler KW. (1996). *Proc. Natl. Acad. Sci. USA,* 93, 7950-7954.

Oya M and Schulz WA. (2000). *Br. J. Cancer,* 83, 626-631.

Praz V, Perier R, Bonnard C and Bucher P. (2002). *Nucleic Acids Res.,* 30, 322-324.

Schwienbacher C, Gramantieri L, Scelfo R, Veronese A, Calin GA, Bolondi L, Croce CM, Barbanti-Brodano G and Negrini M. (2000). *Proc. Natl. Acad. Sci. USA,* 97, 5445-5449.

Sell C, Dumenil G, Deveaud C, Miura M, Coppola D, DeAngelis T, Rubin R, Efstratiadis A and Baserga R. (1994). *Mol. Cell. Biol.*, 14, 3604-3612.

Shrestha P, Sumitomo S, Ogata K, Yamada K, Takai Y, Yang L and Mori M. (1994). *Eur. J. Cancer B Oral Oncol.*, 30B, 393-399.

Sunardhi-Widyaputra S and Van Damme B. (1993). *Pathol. Res. Pract.*, 189, 138-143.

Sunardhi-Widyaputra S and Van Damme B. (1995). *Pathol. Res. Pract.*, 191, 1186-1191.

Toretsky JA and Helman LJ. (1996). *J. Endocrinol.*, 149, 367-372.

Vo HP and Crowe DL. (1998). *Anticancer Res.* ,18, 217-224.

Voz ML, Agten NS, Van de Ven WJ and Kas K. (2000). *Cancer Res.*, 60, 106-113.

Voz ML, Astrom AK, Kas K, Mark J, Stenman G and Van de Ven WJ. (1998). *Oncogene*, 16, 1409-1416.

Zhu Q, White FH and Tipoe GL. (1997). *Oral Oncol.*, 33, 29-35.