

available at [www.sciencedirect.com](http://www.sciencedirect.com)journal homepage: [www.intl.elsevierhealth.com/journals/arob](http://www.intl.elsevierhealth.com/journals/arob)

## Changes in gene-expression during development of the murine molar tooth germ

Harald Osmundsen <sup>\*</sup>, Maria A. Landin, Sigurd H. From, Kristin M. Kolltveit, Steinar Risnes

Department of Oral Biology, University of Oslo, Box 1052 Blindern, 0316 Oslo, Norway

### ARTICLE INFO

#### Article history:

Accepted 9 February 2007

#### Keywords:

Microarray  
Real-time RT-PCR  
Odontogenesis  
Timecourse  
Acox1  
Shh  
Ambn  
Amlx  
Isbp  
Clu  
Ameloblasts  
Odontoblasts  
Mesenchyme  
Dentin  
Enamel  
Oral epithelium  
Mineralisation

### ABSTRACT

In a matter of a few days the murine tooth germ develops into a complex, mineralized, structure. Murine 30 K microarrays were used to examine gene expression in the mandibular first molar tooth germs isolated at 15.5 dpc and at 2 DPN. Microarray results were validated using real-time RT-PCR.

The results suggested that only 25 genes (3 without known functions) exhibited significantly higher expression at 15.5 dpc compared to 2 DPN. In contrast, almost 1400 genes exhibited significantly ( $P < 0.015$ ) higher expression at 2 DPN compared to 15.5 dpc, about half of which were genes with unknown functions. More than 50 of the 783 known genes exhibited higher than 10-fold increase in expression at 2 DPN, amongst these were genes coding for enamel matrix proteins which were expressed several 100-fold higher at 2 DPN. GO and KEGG analysis showed highly significant associations between families of the 783 known genes and cellular functions relating to energy metabolism, protein metabolism, regulation of cell division, cell growth and apoptosis. The use of bioinformatics analysis therefore yielded a functional profile in agreement with known differences in tissue morphology and cellular composition between these two stages. Such data is therefore useful in directing attention towards genes, or cellular activities, which likely are worthy of further studies as regards their involvement in odontogenesis.

© 2007 Elsevier Ltd. All rights reserved.

## 1. Introduction

Tooth development results from interactions between oral epithelium and cranial neural crest derived mesenchyme<sup>1,2</sup> involving a large number of genes, the expression of which are regulated in time and space. So far expression of more than 250 genes and/or their translated proteins have been detected during tooth germ development by in situ hybridization and immunocytochemistry.<sup>3</sup> Main events in odontogenesis are dentition patterning, establishment of tooth morphology and differentiation of ectomesenchymal cells into dentin-produ-

cing odontoblasts and epithelial cells into enamel-producing ameloblasts. Synthesis of dentin and enamel, the two main hard tissues of the tooth, involves secretion of two different types of extra-cellular matrices, both of which are subjected to controlled mineralization.

An overview of total gene expression in the tooth organ at different developmental stages should help our understanding of the cellular phenomena operating during odontogenesis. Expression profiling may be useful to highlight differences in gene expression between various stages of odontogenesis. To highlight developmental differences we

<sup>\*</sup> Corresponding author. Tel.: +47 22 84 03 51; fax: +47 22 84 03 02.

E-mail address: [haraldo@odont.uio.no](mailto:haraldo@odont.uio.no) (H. Osmundsen).

0003-9969/\$ – see front matter © 2007 Elsevier Ltd. All rights reserved.

doi:10.1016/j.archoralbio.2007.02.008

have compared two distinctly different stages in the development of the mouse mandibular first molar; 15.5 dpc (days post coitum) (late cap stage) and 2 DPN (2 days postnatal) (late bell/early secretory stage).<sup>4</sup> To this end 30 k murine microarrays were used.

## 2. Methods

### 2.1. Experimental animals

Tooth germs were isolated from mouse embryos and pups (CD-1 strain) at 15.5 days post coitum (dpc) and at 2 DPN (i.e. 2 days postnatal), the day of vaginal plug being set to 0.5 dpc. Embryos were staged according to the Theiler criteria.<sup>5</sup> First mandibular molar tooth germs were micro-dissected from 10 phenotypical embryos/pups at the various developmental stages. Most of the tooth follicle remained with the tooth germ.

The pregnant mice were sacrificed by cervical dislocation, the pups by decapitation. Embryos and pup heads were immediately immersed in RNAlater (Ambion Inc., TX, USA). The dissection of tooth germs was subsequently carried out in diluted RNAlater (diluted 1:1 with PBS). While still immersed, the first right mandibular molar germs were dissected out and transferred to undiluted RNAlater<sup>TM</sup>.

The animal-house had a 12-h light:12-h dark cycle, and was thermostatted at 21 °C with a relative humidity of 60%. Fodder and water were supplied ad lib. The animals were kept according to the regulations of the Norwegian Gene Technology Act of 1994.

### 2.2. Isolation of RNA from tooth germs

Total RNA was extracted from single tooth germs by using the Qiagen RNA Mini-kit<sup>TM</sup>. This yielded RNA fractions exhibiting a ratio of OD<sub>260</sub>/OD<sub>280</sub> of at least 1.7. Concentrations of RNA (dissolved in RNA storage solution<sup>TM</sup> (Ambion Inc.)) were assayed using ribo-green<sup>TM</sup> (Molecular Probes, OR, USA). All reagents used were of molecular biology grade.

The quality of isolated RNA was assessed by using the RNA for RT-PCR. RNA solutions giving a signal of correct length from various primer-pairs (i.e. ribosomal protein L27,  $\beta$ -actin or glutamate dehydrogenase) were considered suitable for use in microarray analysis.

### 2.3. Microarray analysis of mRNAs from tooth germs

Murine oligo(30 k)-microarrays were purchased from the NTNU Microarray Core Facility, Trondheim, Norway. These slides had been printed using the Operon murine v. 3 oligo set (Qiagen GmbH, Hilden Germany).

cDNA synthesis, labelling and hybridization were carried out as described for the use of the Cy3 and Cy5 labelling using the Genisphere 3DNA Array 900<sup>TM</sup> detection kit (Genisphere, PA, USA). Each slide was hybridized with cDNA obtained starting with 1  $\mu$ g of total RNA. The microarrays were scanned in a Packard Bioscience Scanarray Lite microarray scanner (Perkin-Elmer Life and Analytical Sciences, Inc., MA, USA). The Cy3 and Cy5 fluorescence signals were quantified by using the ScanArray Express v. 2.2<sup>TM</sup> program (Perkin-Elmer Life and

Analytical Sciences, Inc.). The resulting fluorescence data (contained in a .gpr-file) was analysed using the Spotfire<sup>TM</sup> Functional Genomics v. 8 (Spotfire, MA, USA). The .gpr-files have been deposited in the ArrayExpress database with reference no. E-MEXP-947.

In this investigation total RNA was extracted from each of three tooth germs removed at 15.5 dpc and at 2 DPN. RNA from the three separate tooth germs, isolated at the same stage, was mixed in a 1:1:1 ratio (w/w basis). Each of three microarrays was hybridized with cDNA prepared from 1  $\mu$ g of RNA derived from each of the resulting two RNA solutions. Although this approach is statistically less desirable, it consumed fewer slides compared to an approach involving triplicates of three individual RNA-samples. In these experiments cDNA from 15.5 dpc was Cy3 labelled, and cDNA from 2 DPN was Cy5 labelled.

The printed arrays included probes for 10 different mRNAs from *Arabidopsis thaliana*, enabling checks of hybridization to be carried out using a spike mixture of 10 different mRNAs from *A. thaliana* (purchased from Stratagene, CA, USA). Pairs of the *A. thaliana* mRNAs were mixed in 10 different ratios ranging from 0.1 to 5. The ratios of the resulting Cy5/Cy3 signal ratios were always about 15–20% less than the expected values.

### 2.4. Validation of microarray results using real-time RT-PCR

Levels of selected mRNAs (*Acox1*, *Arpc3*, *Actb*, *Ambn*, *Amlx*, *Aplp1*, *Brd7*, *Catnb*, *Clu*, *Fzd6*, *Gna11*, *Ibsp*, *Itrp3*, *Pfn2*, *Ppif*, *Rbp2*, *Rpl27*, *Shh*, *Slc25a5*, *Wif1*) were also assayed by real-time quantitative RT-PCR using primers designed using Primer3.<sup>6</sup> cDNA was synthesized by oligo-dT priming using First Strand Synthesis Kit (Fermentas GmbH, St. Leon-Rot, Germany). Real-time PCR assays were carried out using Stratagene Mx 4000 PCR system (CA, USA) using a qPCR core kit (Eurogentec, Seraing, Belgium). RNA isolated from three to five separate tooth germs from each stage (13.5, 15.5, and 19.5 dpc; 2 and 4 DPN) was used in this analysis, and every analysis was carried out in triplicates. Statistical evaluation of the significance of differences between measured Ct-values was carried out using the REST 2005 program.<sup>7</sup>

Assay of expression of *Ambn*, *Amelx*, *Brd7*, *Clu*, *Ibsp*, *Ppif* and *Rbp2* were also carried out at several stages of foetal development (Fig. 4). The primers used here were purchased from Applied Biosystems Inc. (CA, USA). Here the Ct-values were normalized to that of *Ppif* which was found to exhibit relatively stable expression.

### 2.5. Statistical analysis

Statistical evaluation (two-tailed t-test) of non-microarray data was carried out using GraphPad InStat (Graphpad Software Inc., CA, USA). Statistical analysis of microarray data was carried out on data derived from sets of triplicate slides which were combined into one data-set. Genes exhibiting net fluorescence values of less than 300 in both channels were not included in further analysis. The ANOVA facility of the Spotfire program was used to select genes which exhibited statistically significant differences in levels of expression ( $P < 0.015$ ) with respect to developmental stage. For this purpose the measured fluorescence intensities

(median values, with background subtracted) were converted to  $\log_2$ -scale, and the  $\log_2$ -values were subjected to z-score normalization.<sup>8</sup> The z-score normalized values were used in subsequent ANOVA analysis.

### 2.6. Functional analysis of differentially expressed genes

In an attempt to provide functional interpretations of significant changes in expression profiles we have used the DAVID<sup>9</sup> program for GO analysis (WebGestalt has a limit of 500 genes for GO analysis), and the WebGestalt<sup>10</sup> program for KEGG analysis using the WebGestalt mouse reference data. The data input for these programs was files listing Entrez gene IDs for genes which exhibited significant higher levels of expression at either 15.5 dpc or 2 DPN.

## 3. Results

### 3.1. Microarray results

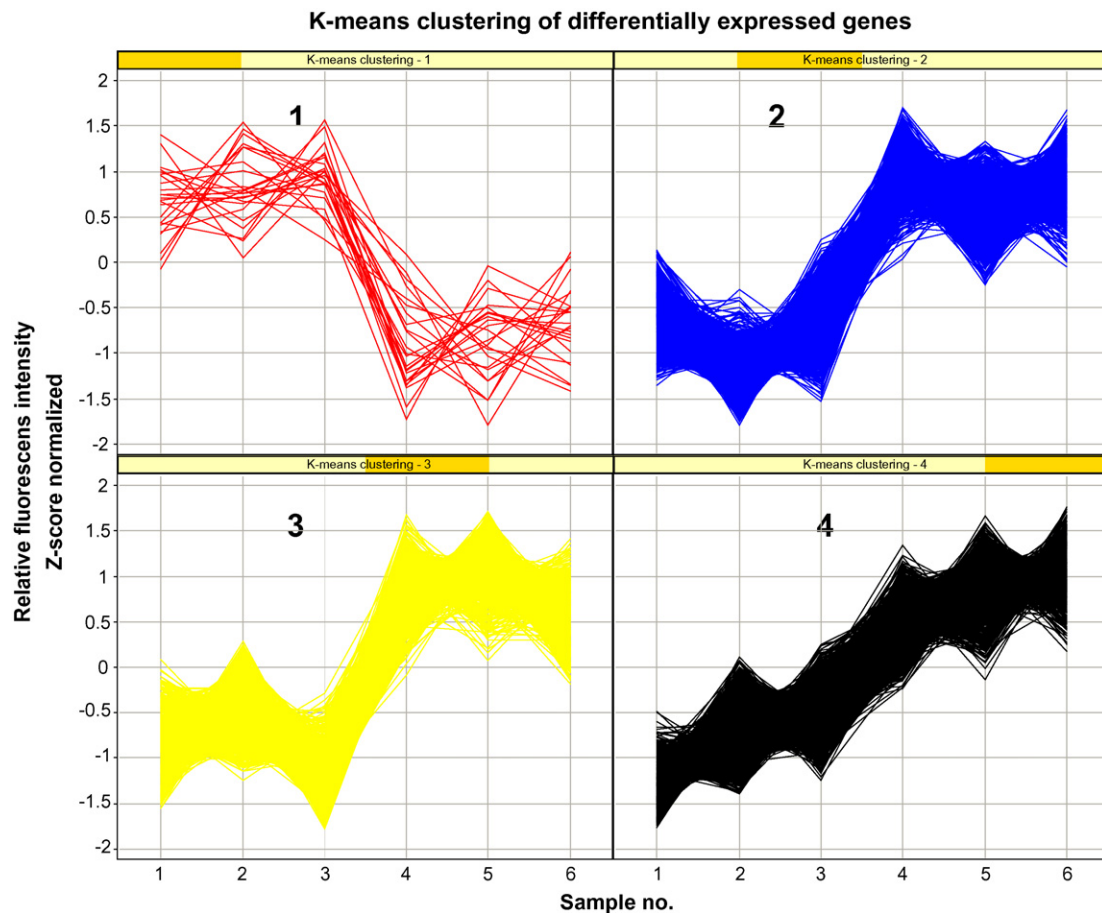
By using the ANOVA function within the Spotfire program it was possible to screen the resulting data-sets for genes which showed significantly different ( $P \leq 0.015$ ) levels of expression

at 15.5 and 2 DPN. This amounted to 1379 genes, 25 of which exhibited a higher level of expression at 15.5 days. All selected genes were clustered into four clusters using K-means clustering, one of which contained the latter 25 genes (Cluster 2, Fig. 1). The remaining genes were distributed amongst the other nine clusters.

### 3.2. Genes with significantly higher levels of expression at 15.5 dpc

Three of these 25 genes have no known functions; 19 of the remaining 22 genes were recognized by WebGestalt (Table 1). The expression of Rbp2 was 12-fold higher at 15.5 dpc, followed by Brd7 (nine-fold), Sox2 (seven-fold). Four-fold increases were found for Ax1, Ibsp and Six2.

GO analysis suggested that most of these genes code for regulatory proteins (Table 1). Three genes code for proteins with functions associated with "Transcription factor activity". These genes, and additional 11 genes, were associated with the molecular function "Binding". Two genes (Krt19 and Tnnt1) were associated with the GO-category "Cytoskeleton" (Table 1). Levels of expression of these genes were 2–13-fold higher at 15.5 dpc compared to 2 DPN. Results for some of these genes, i.e. Brd7, Ibsp, Rbp2, are shown in Fig. 3.



**Fig. 1** – Differences in gene-expression in tooth germs isolated at 15.5 dpc and 2 DPN as visualized by K-means clustering. With the Spotfire program the expression data presented in Fig. 1 have been organized into 10 clusters, labelled 1–10, using K-means clustering. The numbers on the X-axis refers to the microarrays used: 1–3 (Cy3 channel, 15.5 dpc) and 4–6 (Cy5 channel, 2 DPN). The experimental details are otherwise as described in Section 2.

**Table 1 – GO analysis of genes which exhibited significantly higher expression at 15.5 dpc**

GO category	No of genes	Genes (fold higher expression at 15.5 dpc)
Transcription factor activity	3	Prrx1 (2 ± 1), Six2 (4 ± 2), Sox2 (7 ± 5)
Molecular function; binding	14	Axl (4 ± 2), Brd7 (9 ± 1), Ccne1 (1.4 ± 0.2), Cdyl (2 ± 1), Efemp1 (3 ± 1), Grb10 (2 ± 0.1), Hba-a1 (1 ± 0.1), Ibsp (4.0 ± 0.3), Prrx1, Prss12 (2 ± 0.4), Rbp2 (123 ± 5), Six2 (4 ± 2), Slc9a3r2 (2 ± 11), Sox2 (7 ± 5)
Cytoskeleton	2	Krt1-19 (2 ± 1), Tnnt1 (1 ± 0.3)
Twenty-two known genes (of a total of 25) which exhibited significantly higher expression at 15.5 dpc were subjected to GO-analysis using the WebGestalt program, and 19 of these were recognized by the WebGestalt database. Their resulting GO classification of the 19 genes recognized by the WebGestalt database is shown.		

### 3.3. Genes with significantly higher levels of expression at 2 DPN

Of 1379 genes is showing enhanced expression as the tooth germ prepares for, and enters into, the stage when dentin and enamel are synthesized 596 genes were genes with no known function (i.e. having no gene denotations and being identified solely via oligo IDs). Of the remaining 783 genes 738 were recognized by WebGestalt and 758 by DAVID.

Twenty-eight of the more highly expressed genes amongst these 783 genes are presented in Fig. 2. The expression of the enamelin gene (*Enam*) was increased some 450-fold, while that of the ameloblastin gene (*Ambn*) and of the amelogenin gene (*Amlx*) were increased some 180- and 60-fold, respectively. These are likely underestimated values as the fluorescence signals from the *Ambn* and *Amlx* probes were saturated at 2 DPN. The signal from the *Enam*-probe had a magnitude of about 1/20 of the *Ambn*/*Amlx* signals (results not shown). The mRNA levels for the genes coding for integrin  $\alpha 2$  (*Itga2*) and for clusterin (a sulfated glycoprotein) (*Clu*) exhibited about 120-fold increase in expression from day 15.5 dpc to 2 DPN, respectively. The results presented in Fig. 4 show that this powerful up-regulation of expression of *Ambn*, and *Amelx* is initiated around birth (19.5 dpc).

Genes coding for cytoskeletal keratins (*Krt2-6g*, *Krt2-4*, *Krtap5-1*) are also amongst these 28 genes. Additional keratin genes were also several-fold higher expressed (not shown). Regulatory genes also appear amongst these 28 genes, e.g. *Fzd6*, *Wnt10a*, *Shh*, *Irx5* and *Msx2*. Also, mRNA levels for genes coding for several  $Ca^{2+}$ -binding proteins were several-fold higher at 2 DPN compared to 15.5 dpc (e.g. *Tesc*-pending, *Cab45* (*Sdf4*), *Sid470p*, *Ca22*, *S100a9*, the latter four are not shown in Fig. 2).

### 3.4. GO analysis of genes exhibiting significantly higher levels of expression at 2 DPN

GO analysis of the 758 genes recognized by DAVID suggested a significant ( $P < 0.01$ ) association between these genes and the three major GO categories “Biological process”, “Molecular function” and “Cellular compartment”. Each of these major categories contained several sub-categories to which also significant GO associations were found (Table 2). Taken together this functional annotation of these 758 genes provides a description of cellular functions which likely are stimulated at 2 DPN.

Results presented in Table 2 show that in the main category “Biological process” the following sub-categories were

amongst those with the higher number of associated genes: “Cellular physiological process” (372 genes), “Cellular metabolism” (279 genes), “Macromolecule metabolism” (172 genes), “Biosynthesis” (93 genes), “Cell organisation and biogenesis” (82 genes), and “Localization” (81 genes).

In the main category “Cellular compartment” the sub-categories with the higher numbers of genes associated were: “Cell” (433 genes), “Intracellular” (383 genes), “Organelle” (346 genes), “Cytoplasm” (249 genes), “Membrane-bound organelle” (274 genes), “Non-membrane-bound organelle” (125 genes), “Mitochondrion” (94 genes), “Ribonucleoprotein complex” (72 genes), “Ribosome” (53 genes), “Cytosol” (51 genes), “Endoplasmic reticulum” (43 genes), “Organelle membrane” (41 genes), “Mitochondrial membrane” (29 genes), and “Mitochondrial inner membrane” (23 genes) (Table 2).

Results presented in Table 2 also show that in the main category “Molecular function” the sub-category with the higher number of genes were: “Protein binding” (118 genes), “Nucleic acid binding” (119 genes), “Structural molecule activity” (79 genes), “Transporter activity” (75 genes), “RNA binding” (60 genes), “Structural constituent of ribosome” (55 genes), “Carrier activity” (49 genes), “Ion transporter activity” (47 genes) “Oxidoreductase activity” (46 genes), while “Cytoskeletal protein binding” encompassed 18 genes.

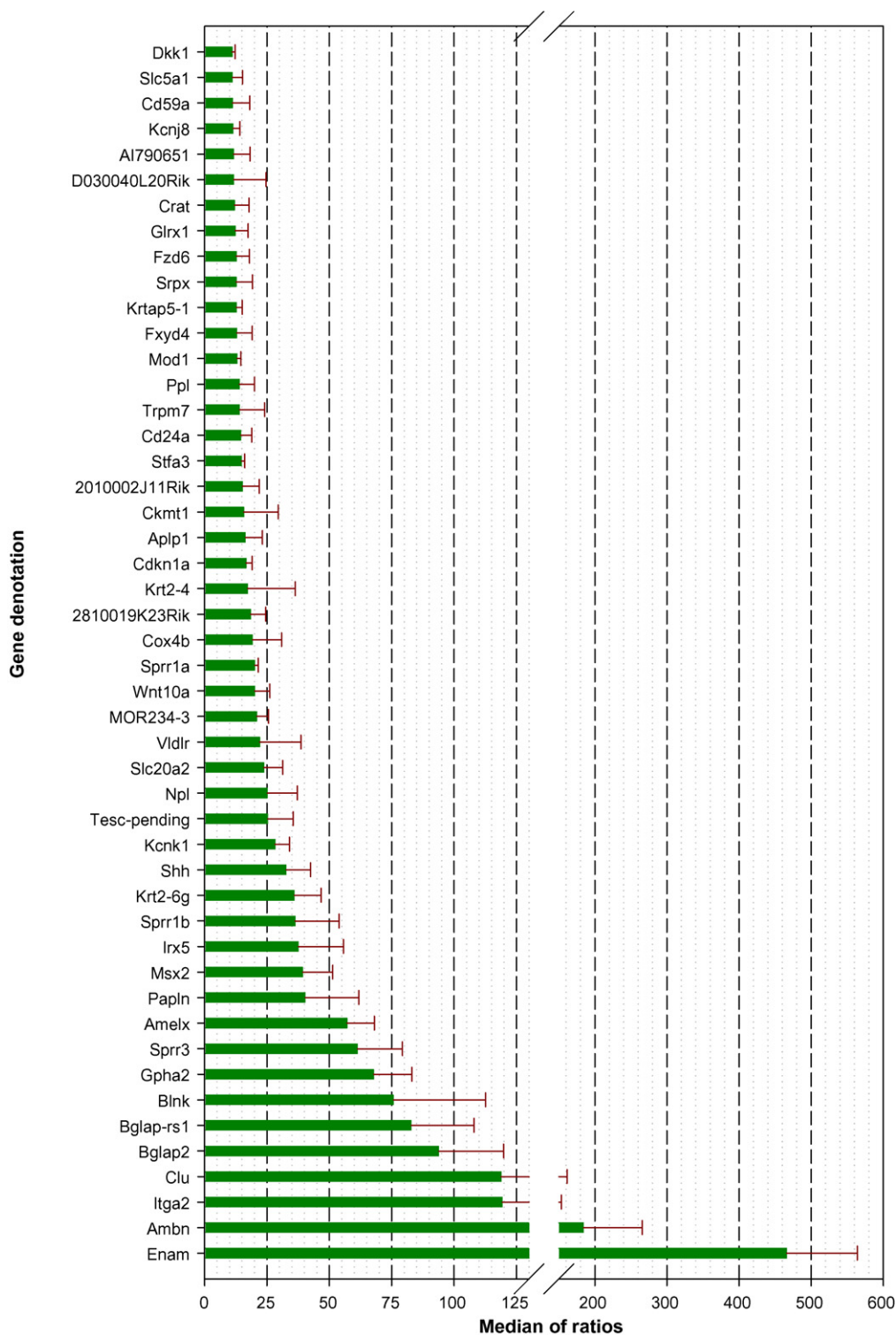
### 3.5. KEGG analysis of genes exhibiting significantly higher levels of expression at 2 DPN

KEGG analysis suggest that the functional significance of the increased expression of the 783 known genes at 2 DPN may be visualized as occurring along three major trends.

One trend emerging is that of stimulated capacity for generation of energy, i.e. for oxidative phosphorylation, fatty acid  $\beta$ -oxidation or TCA-cycle and glycolysis (Tables 2 and 3). KEGG analysis makes this point directly by the 42 genes associated with “Oxidative phosphorylation”, seven genes with “Glycolysis/gluconeogenesis”, eight with “Pyruvate metabolism and TCA cycle” and six with “Fatty acid metabolism” (Table 3).

The second major trend relates to macromolecule metabolism, particularly protein metabolism which appears to be in a highly dynamic state. Forty genes coding for “Ribosome” were significantly higher expressed at 2 DPN, as well as 12 genes coding for “Proteasome” (Table 3). Also, the capacity for “Protein export” appears enhanced (four genes).

Finally, the third trend is represented by numerous genes coding for proteins implicated in pathways regulating cell division, in growth or in organ morphogenesis, which also



**Fig. 2 – Genes which exhibited the higher increase in degree of expression at 2 DPN compared to 15.5 dpc. The 28 genes (of 783) which exhibited the higher increase in expression at 2 DPN compared to 15.5 dpc are presented. The columns show the medians of the ratios of medians (net Cy5 fluorescence intensity/net Cy3 fluorescence intensity) with S.D. indicated, i.e. 2 DPN/15.5 dpc. The genes have been listed in order of increasing median values. All genes presented exhibited a ratio of medians >10. The medians are based on data obtained from triplicate microarrays as described in Section 2.**

exhibited higher expression at 2 DPN. The KEGG results showed association of 12 genes with “Regulation of actin cytoskeleton”, 14 genes with “Cell communication”, 11 genes with “MAPK signalling pathway”, 10 genes with “Cell cycle”, 8

genes with “TGF- $\beta$ -signalling pathway”, and 4 genes with “Notch signalling pathway”. Similarly, genes associated with biosynthesis of cell-cell junctions, i.e. “Tight junction” (10 genes), “Adherens junction” (5 genes), “Gap junction” (7

**Table 2 – Gene ontology of the 783 recognized genes which exhibited significantly higher expression at 2 DPN compared to 15.5 dpc**

	No of genes	Total (%)	P Value
<b>Biological process</b>			
Cellular physiological process	372	61	1.08E-23
Cellular metabolism	279	46	1.63E-17
Cell organization and biogenesis	82	13	4.02E-17
Macromolecule metabolism	172	28	1.18E-17
Biosynthesis	93	15	5.72E-23
Protein biosynthesis	67	11	3.72E-10
Ribosome biogenesis	20	3	1.44E-12
Purine nucleoside triphosphate biosynthesis	17	2	1.91E-9
Nucleoside triphosphate biosynthesis	17	2	5.47E-9
Generation of precursor metabolites and energy	48	7	3.88E-10
Oxidative phosphorylation	19	3	9.48E-10
Hydrogen transport	15	2	2E-5
ATP biosynthesis	15	2	1.08E-4
ATP synthesis coupled proton transport	13	2	2.04E-3
Electron transport	24	3	3.18E-3
Protein localization	41	6	1.31E-7
Protein folding	19	3	4.48E-5
Modification-dependent protein catabolism	12	1	1.38E-3
Macromolecule catabolism	37	6	6.43E-3
Ubiquitin-dependent protein catabolism	11	1	4.02E-3
Apoptotic program	6	0	1.22E-2
Chromosome organization and biogenesis	20	3	3.3E-4
DNA packaging	20	3	3.03E-3
Localization	81	13	1.33E-2
Cellular carbohydrate metabolism	16	2	2.22E-3
Alcohol catabolism	8	1	8.82E-4
Alcohol biosynthesis	5	0	1.17E-2
Coenzyme biosynthesis	15	2	3.73E-7
<b>Cellular compartment</b>			
Cell	433	71	1.69E-6
Cytoplasm	249	41	0E0
Organelle	346	57	0E0
Intracellular	383	63	0E0
Cytosol	51	8	1.91E-21
Membrane-bound organelle	274	45	2.12E-25
Organelle membrane	41	6	8.27E-13
Mitochondrion	94	15	5.31E-11
Mitochondrial membrane	29	4	2.25E-3
Mitochondrial inner membrane	23	3	1.39E-1
Mitochondrial electron transport chain	11	1	3.71E-2
Proton-transporting two-sector ATPase complex	14	2	3.35E-8
Mitochondrial matrix	10	1	9.84E-3
Non-membrane-bound organelle	125	20	3.38E-30
Endoplasmic reticulum	43	7	9.83E-4
Signal recognition particle	3	0	1.67E-1
Nuclear envelope-endoplasmic reticulum network	7	1	1.88E-2
Endomembrane system	12	1	9.23E-2
Vacuolar membrane	3	0	1.44E-1
Lamellipodium	6	0	7.75E-3
Ribonucleoprotein complex	72	11	7.57E-19
Ribosome	53	8	4.74E-13
Cytosolic ribosome (sensu eukaryota)	23	3	1.64E-9
Mitochondrial ribosome	10	1	1.1E-2
Eukaryotic 48s initiation complex	4	0	6.45E-2
Eukaryotic 43s preinitiation complex	5	0	4.14E-2
Proteasome complex (sensu eukaryota)	11	1	4.9E-3
Proteasome core complex (sensu eukaryota)	8	1	8.74E-3
Arp2/3 protein complex	5	0	1.84E-2
Contractile fibre	5	0	8E-2
Keratin filament	3	0	1.56E-1
Actin filament	4	0	1.66E-1
<b>Molecular function</b>			

**Table 2 (Continued)**

	No of genes	Total (%)	P Value
Protein binding	118	19	5.24E-8
Nucleic acid binding	119	19	7.66E-4
Structural molecule activity	79	13	6.21E-27
Structural constituent of ribosome	55	9	4.35E-37
RNA binding	60	9	8.61E-20
mRNA binding	8	1	1.12E-3
rRNA binding	6	0	2.9E-6
Primary active transporter activity	42	6	1.19E-24
Transporter activity	75	12	2.88E-7
Protein transporter activity	15	2	3.74E-5
Carrier activity	49	8	1.38E-15
Cation transporter activity	45	7	8.24E-10
Ion transporter activity	47	7	4.75E-9
Metal ion transporter activity	17	2	1.64E-9
Sodium ion transporter activity	15	2	2.33E-9
Cation-transporting atpase activity	14	2	1.88E-6
Hydrogen ion transporter activity	39	6	2.03E-25
p-p-Bond-hydrolysis-driven transporter activity	17	2	3.28E-8
Hydrogen-transporting ATP synthase activity, rotational mechanism	13	2	5.9E-12
Oxidoreductase activity	46	7	1.55E-6
Electron transporter activity	24	3	4.58E-10
Oxidoreductase activity, acting on NADH or NADPH, quinone or similar compound as acceptor	15	2	8.63E-11
Hydrogen-transporting two-sector ATPase activity	12	1	4.64E-10
Oxidoreductase activity, acting on diphenols and related substances as donors, cytochrome as acceptor	4	0	1.93E-3
Oxidoreductase activity, acting on heme group of donors, oxygen as acceptor	6	0	1.43E-3
Heme-copper terminal oxidase activity	6	0	1.43E-3
AA3-type cytochrome c oxidase	5	0	2.44E-3
Unfolded protein binding	13	2	1.23E-4
Proteasome endopeptidase activity	8	1	3.37E-7
Structural constituent of cytoskeleton	17	2	1.91E-8
Actin binding	16	2	4.06E-4
Cytoskeletal protein binding	18	2	2.28E-3
Oligosaccharyl transferase activity	4	0	1.78E-4

The genes which were significantly ( $P < 0.015$ ) higher expressed at 2 DPN compared to 15.5 dpc were identified using the ANOVA function in the Spotfire program. These genes were subjected to GO analysis using the DAVID software as described in Section 2. The resulting tabulated GO associations between genes and GO categories were significant at  $P < 0.01$ . To conserve space about 30 genes have been listed for categories containing more than about 40 genes. Likewise, most GO categories containing less than five genes have been excluded from the table.

genes), and genes associated with cell-matrix junctions, i.e. "Focal adhesion" (10 genes), were found to be significantly higher expressed at 2 DPN.

For all these groups of genes the levels of significance for the associations between genes and functions were highly significant ( $P < 0.01$ ).

### 3.6. Validation of microarray results using real-time RT-PCR

Some of the genes from the microarray data-set of Fig. 1 were selected to have their expression assayed also by real-time RT-PCR, choosing mostly genes which showed pronounced differences in levels of expression at 15.5 dpc and 2 DPN. The results presented in Fig. 3 and Table 4 show good correlation with results obtained with microarrays. The extent of change was rarely identical, but was invariably in the same direction. As a rule, microarrays measured a greater increase in expression compared to real-time RT-PCR (Fig. 3). *Ambn*, *Amlx* and *Clu* (Fig. 4) exhibited increases in expression orders of magnitude higher than the other genes assayed, and these results are therefore presented in Table 4. The increase in

levels of expression of *Amelx* and *Ambn* between 15.5 dpc and 2 DPN was, however, vastly higher as assayed by RT-PCR (Table 4). This was most likely due to saturation of the fluorescence signals from the 2 DPN samples in the microarray assay (results not shown).

For some of these genes their expression was also assayed at additional stages during tooth germ development using real-time RT-PCR. To this end tooth germs were also collected at 13.5 and 19.5 dpc and 4 DPN and prepared for RT-PCR analysis as described in Section 2. These results (Fig. 4) show that the dramatic increase in expression of *Amelx* and *Ambn* occurs after birth (19.5 dpc), while that of *Clu* is appreciable already around 13.5 dpc. *Ibsp*, *Brd7*, and *Rbp2*, which exhibited a higher expression at 15.5 dpc, exhibited somewhat different patterns of expression (Fig. 4).

## 4. Discussion

The most striking finding emerging from these results is the numerous genes exhibiting significantly increased expression at 2 DPN (1379 genes), contrasting the 25 genes with

**Table 3 – KEGG analysis of genes which were significantly higher expressed at 2 DPN compared to 15.5 dpc**

KEGG pathway	Number of genes	Entrez gene IDs	Enrichment/P value
Oxidative phosphorylation	42	Atp5o, Atp6v0b, Atp5d, Atp5f1, Atp5e, Atp5g3, Atp5h, Atp5j2, Atp5k, Atp5l, Atp6v0b, Atp6v0c, Atp6v1e1, Atp6v1f, Atp6v0e, Cox4i1, Cox7a2l, Cox5a, Cox6c, Cox7a2, Cox8a, Ndufa1, Ndufa3, Ndufa9, Ndufa5, Ndufa11, Ndubf2, Ndubf8, Ndufc1, Ndubf3, Ndufv2, Ndufs5, Ndufa2, Ndufs8, Ndufs2, Ndufs6, Np15, 3110001N18RIK, Sdhd, Ucrb, Uqcr2, Uqcrh,	R = 27.60, P = 1.27E-42
Ribosome	40	Arbp, Fau, Rpl10, Rps2, Rpl10a, Rpl18, Rpl21, Rpl26, Rpl29, Rpl32, Rpl37a, Rpl5, Rps16, Rps18, Rps19, Rps24, Rps29, Rps8, Uba52, Rps15a, Rpl12, Rpl8, Rpl13, Rpl7A, Rps11, Rps28, Rpl35a, Rps23, Rpl15, Rps21, Rpl36a, Rpl36al, Rpl11, Rpl41, 3110001N18RIK, Rpl24, Mrpl13, Rps25, Uba52, Rps27a	R = 33.30, P = 2.87E-43
Regulation of actin cytoskeleton	12	Actb, Arpc2, Arpc3, Arpc4, Cfl1, Cyfip1, F2r, Fgfr1, Map2k1, Pfn1, Pfn2, Vil2	R = 5.96, P = 3.84E-08
Cell communication	14	Actb, Col4a2, Gja1, Krt1-10, Krt1-13, Krt1-14, Krt1-17, Krt2-1, Krt2-4, Krt2-6a, Krt2-6g, Lmna, Lamc1, Lamc2	R = 8.74, P = 3.26E-09
Proteasome	12	Psma2, Psma3, Psma4, Psma6, Psma7, Psmb1, Psmb4, Psmb6, Psmc2, Psmc3, Psmc5, Psmc6	R = 29.97, P = 1.88E-13
Tight junction	10	Actb, Catnb, Csd, Gnai2, Kras2, Nras, Prkci, Ppp2cb, Sec61l1, Vapa	R = 7.56, P = 1.87E-07
Cell cycle	10	Anapc2, Anapc10, Cckn1a, Ccnh, Cul1, Sfn, Smc1l1, Tfpp1, Ywhae, Ywhaz	R = 7.07, P = 3.40E-06
Focal adhesion	10	Actb, Catnb, Col4a2, Grb2, Ilk, Kras2, Nkras, Lamc1, Lamc2, Map2k1	R = 3.67, P = 5.96E-5
TGF- $\beta$ signaling pathway	8	Ikb1, Ikb3, Pitx2, Ppp2cb, Tfdp1, Tgfb2, Cul1	R = 7.22, P = 2.81E-05
Protein export	4	Sec61a1, Sec11l3, Srp9, Srp14	R = 27.25, P = 3.62E-05
Glycolysis/gluconeogenesis	7	Akr1a4, Eno1, Gapd, Ldh2, Pgam1, Pfkfb3, Pfkfb1	R = 9.37, P = 1.94E-05
Gap junction	7	Gja1, Gna11, Gnai2, Grb2, Itrp3, Tuba6, Map2k1	R = 4.63, P = 5.66E-5
Adherens junction	5	Actb, Actb, Catnb, Chd1, Fgfr1	R = 7.18, P = 9.18E-05
MAPK signaling pathway	11	Atf4, Daxx, Fgfr1, Grb2, Hsbp1, Kras2, Nras, Ppm1a, Ppp5c, Map2k1, 1500003O03Rik	R = 2.51, P = 8.46E-4
Notch signaling pathway	4	Jag2, Ncor2, Ncstn, Psenen	R = 5.09, P = 9.85E-4
Fatty acid metabolism	6	Acat1, Hadhsc, Sh3glb1, Nat5, Echs1	R = 7.07, P = 1.01E-4
Pyruvate metabolism and citrate cycle (TCA cycle)	8	Acat1, Cs, Kr1b3, Ldh2, Mdh1, Pfkfb3, Sdhd, Sdh	R = 11.10, P = 2.07E-4

The genes which were significantly ( $P < 0.015$ ) higher expressed at 2 DPN compared to 15.5 dpc were identified using the ANOVA function in the Spotfire program. These genes were subjected to KEGG analysis using the WebGestalt software as described in Section 2. The tabulated KEGG associations were statistically significant ( $P < 0.01$ ), and the level of significance of enrichment (as expressed by the enrichment factor, R) was calculated using the hypergeometric test.<sup>33</sup>

significantly increased expression at 15.5 dpc. Of the genes with increased expression at 2 DPN and with known function (783 genes), 94–96% were recognized by databases searched by DAVID or WebGestalt. Of these 783 genes almost 10% showed an extent of increase in their expression that was more than 10-fold, the majority of which are shown in Fig. 2. Clearly, the increasing differentiation of the developing tooth requires the participation of gene products from many genes which are not active during early stages of tooth development.

#### 4.1. Cellular functions of gene products from genes with significantly higher levels of expression at 15.5 dpc

Of the 22 known genes which exhibited significantly higher levels of expression at 15.5 dpc 19 were recognized by the WebGestalt database. The GO analysis suggested that these 19 genes mostly coded for proteins with regulatory functions (Table 1). This is indicated by the finding that three are involved in “Transcription factor activity”, while 14 are associated with the GO category “Molecular function; binding”, mostly having ligand-binding functions (e.g. retinol-binding protein 2 (Rbp2)). This also applies to proteins coded

for by Brd7 and Ax1 which may be powerful regulators of signalling pathways,<sup>11,12</sup> as well as by Capn6 (results not shown) which may be involved in regulation of signalling pathways.<sup>13,14</sup>

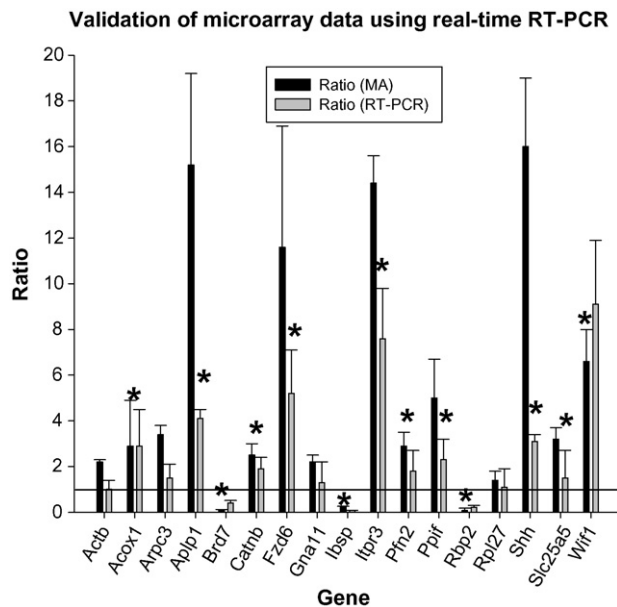
#### 4.2. Functional significance of genes showing increased levels of expression at 2 DPN

GO and KEGG analysis provide a summary glance of possible functional consequences ensuing from the observed increase in expression of 758 known genes at 2 DPN. Three lines of functional implications emerge, i.e. enhanced energy metabolism, a dynamic macromolecule, especially protein, metabolism, and stimulation of pathways regulating cell division and growth.

#### 4.3. Energy metabolism

In the category “Biological process” GO analysis suggested that stimulation had occurred at the level of “Cellular metabolism” (279 genes), “Generation of precursor metabolites and energy” (48 genes), “Oxidative phosphorylation” (19 genes) and “ATP biosynthesis” (15 genes) (Table 2). In the





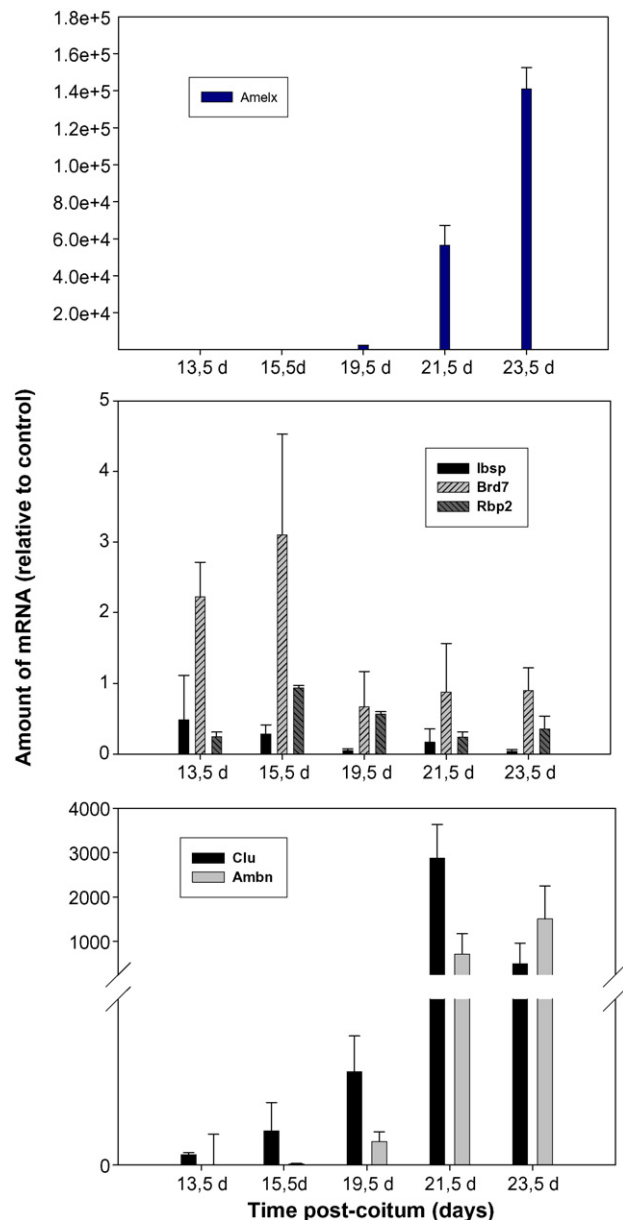
**Fig. 3 – Comparison of expression data—validation of microarray results using real-time RT-PCR.** Real-time RT-PCR was carried out using primers for 17 different genes selected on basis of microarray results. Total RNA was isolated from triplicate tooth germs both at 15.5 dpc and 2 DPN. cDNA synthesis and RT-PCR analysis were carried out on RNA isolated from individual tooth germs as described in Section 2, each RT-PCR assay also carried out in triplicates. The data presented in the figure represents mean ratios (expression at 2 DPN/expression at 15.5 dpc), with S.D. indicated using RT-PCR results and results from triplicate microarrays. Experimental details are otherwise as described in Section 2.

category “Cellular compartment” GO analysis suggested the stimulation to be located to “Mitochondrion” (94 genes), “Mitochondrial inner membrane” (23 genes), “Mitochondrial matrix” (10 genes) and “Cytosol” (51 genes) (Table 2). The GO analysis further suggested that in terms of “Molecular function” the increased activity was associated with “Transporter activity” (75 genes), “Oxidoreductase activity” (46 genes), “Electron transporter activity” (24 genes) and “Hydro-

**Table 4 – Validation of microarray data using real-time RT-PCR-**

Gene	Fold change (2 DPN/15.5 dpc)	
	Microarray-data	RT-PCR data
Ambn	183 ± 167	13.682 ± 1142
Amelx	58 ± 14	1.18 × 10E6 ± 256.000
Clu	118 ± 44	481 ± 115

Expression data from triplicate microarrays (mean ratios, with S.D. indicated) and from RT-PCR-assays for expression of Ambn, Amlx, Brd7, Clu, Ibsp and Rbp2. The tabulated RT-PCR results represent mean ratios (Ct-value at 2 DPN/Ct-value at 15.5 dpc) from at least four separate tooth germs, with S.D. indicated. Experimental details are otherwise described in Section 2.



**Fig. 4 – Time-course of changes in expression of selected genes during development of the tooth germ as assayed using real-time RT-PCR.** Tooth germs were isolated at the various stages shown in the figure. The expression of selected genes (Ambn, Amelx, Brd7, Clu, Ibsp and Rbp2) were measured using real-time RT-PCR as described in Section 2. The plotted values represent means derived from at least four separate tooth germs at each time-point, with S.D. indicated.

gen-transporting ATPase synthase activity, rotational mechanism” (13 genes) (Table 2). Similarly, KEGG analysis showed highly significant association between 42 genes and “Oxidative phosphorylation” (Table 3), suggestive of enhanced energy metabolism, as is eight, seven and six genes associated with “Pyruvate metabolism and TCA cycle”, “Glycolysis/gluconeogenesis” and “Fatty acid metabolism”, respectively (Table 3).

#### 4.4. Cell division and growth

The third major functional trend relates to families of genes coding for proteins implicated in pathways regulating cell division and growth. This conclusion is based on the enhanced expression of numerous genes coding for ribosomal RNAs and for components of the transcriptional and translational machinery (Tables 2 and 3), while biosynthesis of ribosomal RNA, via ribosomal DNA transcription, has been suggested to drive cell growth.<sup>15</sup> Similarly, the KEGG finding of 10 genes being associated with "Cell cycle" is directly suggestive of stimulated cell divisions (Table 3). Two genes coding for markers of cell-proliferation, *Pcna* and *Mki67*, were about 1.5-fold higher expressed at 2 DPN (results not shown).

This line of reasoning is further substantiated by GO and KEGG analysis (Tables 2 and 3) suggesting that protein metabolism is in a highly dynamic state as both biosynthetic (e.g. "Ribonucleoprotein complex", "Eucaryotic 48s initiation complex", "43s pre-initiation complex") and proteolytic (e.g. "Proteasome complex", "Apoptotic program") capacity may be enhanced. In sum, these findings are in line with the first molar tooth germ growing exponentially between 15.5 dpc and 2 DPN,<sup>16,17</sup> and with apoptosis being prevalent in the epithelial compartment of the tooth germ.<sup>18-20</sup>

#### 4.5. Cytoskeleton, focal adhesion, adherens and tight and gap junctions

Results from both GO and KEGG analysis suggested that the cytoskeleton is in a highly dynamic state between 15.5 and 2 DPN, e.g. GO results showing 18 genes associated with "Cytoskeletal protein binding" ("Molecular function", Table 2), KEGG analysis showing 14 and 12 genes being associated with "Cell communication" and "Regulation of actin cytoskeleton", respectively (Table 3). Included amongst these genes were *Actb*, and genes coding for proteins having regulatory functions with respect to F-actin synthesis (e.g. *Cfl1*, *Fgfr1*, *Pfn1*, *Pfn2*, *Arpc2*, *Arpc3* and *Arpc4*) and several genes coding for cytokeratins ("Cell communication", Table 3). *Fgfr1* has previously been studied in tooth germs.<sup>21-23</sup>

In secretory ameloblasts the inter-cellular junctions, and associated actin cytoskeleton, must be in a particularly dynamic state as the ameloblasts are required to move in a complex manner in order to create the complex enamel structures found in mouse and rat molars.<sup>24-26</sup> In this respect our conclusions appear in accordance with previous findings.

#### 4.6. Mineralization and Ca<sup>2+</sup>-homeostasis

The several 100-fold increases in expression of *Amelx*, *Ambn*, and *Enam* genes illustrate that powerful up-regulation of protein biosynthesis is likely to have occurred at this stage (Figs. 2 and 4, Table 4). This is in agreement with the known onset around birth of synthesis of dentin and enamel in the mandibular first molar,<sup>4,27</sup> involving biosynthesis and secretion of extra-cellular matrix proteins.<sup>28-30</sup> *Tuftelin* is another enamel protein (coded by the *Tuft1* gene), but may exhibit wider temporo-spatial expression.<sup>31,32</sup> Our microarray data showed no significant expression of *Tuft1* suggesting that the

oligo-probe for this gene is not functioning, or that this gene was not appreciably expressed at 15.5 dpc and 2 DPN in our mice.

#### 4.7. Time-course of changes in expression of selected genes as assayed by real-time RT-PCR

The onset of an appreciable expression of *Amelx* and *Ambn* at 19.5 dpc (Fig. 4) coincides with initiation of enamel formation.<sup>4,27,29</sup> The subsequent increase in expression is probably due to an increased number, both absolutely and relatively, of cells in the inner enamel epithelium differentiating into secretory ameloblasts.

The expression of the *Clu* (clusterin) gene appears to follow a different time-course compared to that of *Ambn* and *Amelx*. Unlike *Ambn* and *Amelx* the level of the mRNA for clusterin is appreciable already at 13.5 dpc and appears to peak about 21.5 dpc.

#### 4.8. Concluding remarks

Results from our bioinformatic analysis appear in agreement with already established findings. Microarray surveys, combined with GO and KEGG analysis, may therefore provide a valuable tool in assessing the functional profile of gene expression in tooth germs.

The choice of time-points is important. In this study time-points relatively far apart were chosen to highlight differences in gene expression, hence facilitating the bioinformatic interpretation of expression data. Considerable variation in gene expression is likely to occur between these time-points. Therefore, in future studies comparison of gene expression at time-points less widely spaced may prove worthwhile.

### Acknowledgements

The expert technical assistance of Mrs. B. Gehrken, Mrs. T. Woldene and Mr. B. Geronimo is gratefully acknowledged.

### REFERENCES

1. Thesleff I, Sharpe P. Signalling networks regulating dental development. *Mech Dev* 1997;67:111-23.
2. Thesleff I, Vaahtokari A, Kettunen P, Aberg T. Epithelial-mesenchymal signaling during tooth development. *Connect Tissue Res* 1995;32:9-15.
3. <http://bite-it.helsinki.fi/>.
4. Cohn SA. Development of the molar teeth in the albino mouse. *Am J Anat* 1957;101:295-319.
5. Kaufman MH. *The atlas of mouse development (2nd printing)*. London: Academic Press; 1994.
6. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 2000;132:365-86.
7. Pfaffl MW, Horgan GW, Dempfle L. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* 2002;30:e36.

8. Cheadle C, Vawter MP, Freed WJ, Becker KG. Analysis of microarray data using Z score transformation. *J Mol Diagn* 2003;5:73–81.
9. Dennis Jr G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 2003;4:P3.
10. Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res* 2005;33:W741–8.
11. Zhou J, Ma J, Zhang BC, Li XL, Shen SR, Zhu SG, et al. BRD7, a novel bromodomain gene, inhibits G1-S progression by transcriptionally regulating some important molecules involved in ras/MEK/ERK and Rb/E2F pathways. *J Cell Physiol* 2004;200:89–98.
12. Zhou M, Peng C, Nie XM, Zhang BC, Zhu SG, Yu Y, et al. Expression of BRD7-interacting proteins, BRD2 and BRD3, in nasopharyngeal carcinoma tissues. *Ai Zheng* 2003;22:123–7.
13. Dear TN, Boehm T. Diverse mRNA expression patterns of the mouse calpain genes *Capn5*, *Capn6* and *Capn11* during development. *Mech Dev* 1999;89:201–9.
- [14] Numajiri T, Mitsui S, Hisa Y, Ishida T, Nishino K, Yamaguchi N. The expression of a motoneuron-specific serine protease, motopsin (PRSS12), after facial nerve axotomy in mice. *Br J Plast Surg* 2005;59:393–7.
15. Russell J, Zomerdijk JC. RNA-polymerase-I-directed rDNA transcription, life and works. *Trends Biochem Sci* 2005;30:87–96.
16. Gaunt WA. An analysis of the growth of the cheek teeth of the mouse. *Acta Anat (Basel)* 1963;54:220–59.
17. Peterka M, Vonesch JL, Ruch JV, Cam Y, Peterkova R, Lesot H. Position and growth of upper and lower tooth primordia in prenatal mouse—3D study. *J Craniofac Genet Dev Biol* 2000;20:35–43.
18. Shigemura N, Kiyoshima T, Sakai T, Matsuo K, Momoi T, Yamaza H, et al. Localization of activated caspase-3-positive and apoptotic cells in the developing tooth germ of the mouse lower first molar. *Histochem J* 2001;33:253–8.
19. Viriot L, Peterkova R, Vonesch JL, Peterka M, Ruch JV, Lesot H. Mouse molar morphogenesis revisited by three-dimensional reconstruction. III: Spatial distribution of mitoses and apoptoses up to bell-staged first lower molar teeth. *Int J Dev Biol* 1997;41:679–90.
20. Vaahokari A, Aberg T, Thesleff I. Apoptosis in the developing tooth: association with an embryonic signaling center and suppression by EGF and FGF-4. *Development* 1996;122:121–9.
21. Kettunen P, Karavanova I, Thesleff I. Responsiveness of developing dental tissues to fibroblast growth factors: expression of splicing alternatives of FGFR1, -2, -3, and of FGFR4; and stimulation of cell proliferation by FGF-2, -4, -8, and -9. *Dev Genet* 1998;22:374–85.
22. Peters KG, Werner S, Chen G, Williams LT. Two FGF receptor genes are differentially expressed in epithelial and mesenchymal tissues during limb formation and organogenesis in the mouse. *Development* 1992;114:233–43.
23. Salmivirta K, Gullberg D, Hirsch E, Altruda F, Ekblom P. Integrin subunit expression associated with epithelial-mesenchymal interactions during murine tooth development. *Dev Dyn* 1996;205:104–13.
24. Risnes S. A method of calculating the speed of movement of ameloblasts during rat incisor amelogenesis. *Arch Oral Biol* 1979;24:299–306.
25. Lyngstadaas SP, Moinichen CB, Risnes S. Crown morphology, enamel distribution, and enamel structure in mouse molars. *Anat Rec* 1998;250:268–80.
26. Nishikawa S, Kitamura H. Localization of actin during differentiation of the ameloblast, its related epithelial cells and odontoblasts in the rat incisor using NBD-phalloidin. *Differentiation* 1986;30:237–43.
27. Gaete M, Lobos N, Torres-Quintana MA. Mouse tooth development time sequence determination for the ICR/Jcl strain. *J Oral Sci* 2004;46:135–41.
28. Begue-Kirn C, Krebsbach PH, Bartlett JD, Butler WT. Dentin sialoprotein, dentin phosphoprotein, enamelysin and ameloblastin: tooth-specific molecules that are distinctively expressed during murine dental differentiation. *Eur J Oral Sci* 1998;106:963–70.
29. Lyngstadaas SP. On the molecular biology of tooth formation. Ph.D., University of Oslo; 1995.
30. Snead ML, Luo W, Lau EC, Slavkin HC. Spatial- and temporal-restricted pattern for amelogenin gene expression during mouse molar tooth organogenesis. *Development* 1988;104:77–85.
31. Deutsch D, Leiser Y, Shay B, Fermon E, Taylor A, Rosenfeld E, et al. The human tuftelin gene and the expression of tuftelin in mineralizing and nonmineralizing tissues. *Connect Tissue Res* 2002;43:425–34.
32. Zeichner-David M, Vo H, Tan H, Diekwisch T, Berman B, Thiemann F, et al. Timing of the expression of enamel gene products during mouse tooth development. *Int J Dev Biol* 1997;41:27–38.
33. Adams WT, Skopek TR. Statistical test for the comparison of samples from mutational spectra. *J Mol Biol* 1987;194:391–6.