Newcastle University e-prints

Date deposited: 28th March 2013

Version of file: Author final

Peer Review Status: Peer reviewed

Citation for item:

Reppe S, Sachse D, Olstad OK, Gautvik VT, Sanderson P, Datta HK, Berg JP, Gautvik KM. [Identification of transcriptional macromolecular associations in human bone using browser based in silico analysis in a giant correlation matrix](#). *Bone* 2013, **53**(1), 69-78.

Further information on publisher website:

<http://www.elsevier.com>

Publisher's copyright statement:

The definitive version of this article, published by Elsevier, 2013, is available at:

<http://dx.doi.org/10.1016/j.bone.2012.11.015>

Always use the definitive version when citing.

Use Policy:

The full-text may be used and/or reproduced and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not for profit purposes provided that:

- A full bibliographic reference is made to the original source
- A link is made to the metadata record in Newcastle E-prints
- The full text is not changed in any way.

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

<p>Robinson Library, University of Newcastle upon Tyne, Newcastle upon Tyne. NE1 7RU. Tel. 0191 222 6000</p>
--

Title page

Identification of transcriptional macromolecular associations in human bone using browser based *in silico* analysis in a giant correlation matrix

Sjur Reppe^a, Daniel Sachse^{a,b}, Ole K. Olstad^a, Vigdis Gautvik^c, Paul Sanderson^d, Harish K. Datta^e, Jens Petter Berg^{a,b} and Kaare M. Gautvik^{a,c,f}

^aDepartment of Medical Biochemistry, Oslo University Hospital, Ullevaal, Norway

^bDepartment of Medical biochemistry, University of Oslo, Norway

^cInstitute of Basic Medical Sciences University of Oslo, Oslo, Norway

^dThe Royal Victoria Infirmary, Newcastle Upon Tyne, UK

^eInstitute of Cellular Medicine, Newcastle University, Newcastle Upon Tyne, UK

^fDepartment of Medical Biochemistry, Lovisenberg Deacon Hospital, Oslo, Norway

Funding sources

This work was supported by South-Eastern Norway Regional Health Authority, and Oslo University Hospital, Ullevaal, project #29750104, the Norwegian Osteoporosis Society, Anders Jahre's Foundation for Promotion of Science and the Research Council of Norway. This work is part of the European Union project OSTEOGENE (no. FP6-502491).

Running title: Human bone transcriptional interactome

Degrees and e-mail addresses

Sjur Reppe:	Ph.D.	sjur.reppe@medisin.uio.no
Daniel Sachse:	M.Sc.	daniel.sachse@medisin.uio.no
Ole K. Olstad:	Ph.D.	o.k.olstad@medisin.uio.no
Vigdis T. Gautvik:		v.t.gautvik@medisin.uio.no
Paul Sanderson :	M.D.	paul.sanderson@nuth.nhs.uk
Harish K. Datta:	M.D. Ph.D.	H.K.Datta@newcastle.ac.uk
Jens P. Berg:	M.D. Ph.D.	j.p.berg@medisin.uio.no
Kaare M Gautvik:	M.D. Ph.D.	k.m.gautvik@medisin.uio.no

Corresponding author:

Sjur Reppe
Department of Medical Biochemistry
Oslo University Hospital, Ullevaal
Kirkeveien 166
0450 Oslo, Norway
Phone: +47 40614313
Fax: +47 22851058
e-mail: sjur.reppe@medisin.uio.no

Abstract

Intracellular signalling is critically dependent on gene regulatory networks comprising physical molecular interactions. Presently, there is a lack of comprehensive databases for most human tissue types to verify such macromolecular interactions. We present a user friendly browser which helps to identify functional macromolecular interactions in human bone as significant correlations at the transcriptional level. The molecular skeletal phenotype has been characterized by transcriptome analysis of iliac crest bone biopsies from 84 postmenopausal women through quantifications of ~23000 mRNA species. When the signal levels were inter-correlated, an array containing >260 million correlations were generated, thus recognizing the human bone interactome at the RNA level. The matrix correlation and p values were made easily accessible by a freely available online browser. We show that significant correlations within the giant matrix are reproduced in a replica set of 13 male vertebral biopsies. The identified correlations differ somewhat from transcriptional interactions identified in cell culture experiments and transgenic mice, thus demonstrating that care should be taken in extrapolating such results to the *in vivo* situation in human bone.

The current giant matrix and web browser is a valuable tool for easy access to the human bone transcriptome and molecular interactions represented as significant correlations at the RNA-level. The browser and matrix should be a valuable hypothesis generating tool for identification of regulatory mechanisms and serve as a library of transcript relationships in human bone, a relatively inaccessible tissue.

Keywords: human, bone biopsies, expression profiling, browser, correlation,

1. Introduction

The regulation of gene expression in complex organisms relies on functional cohorts of macromolecules, including transcription factors (TFs), chromatin modulators, RNA enzymes/splicing RNAs, microRNAs and other non-coding RNAs. A part of this complex machinery consists of inducer and suppressive TFs whose actions are mediated by enhancer elements containing multiple, but distinct TF binding sites scattered over distances of roughly 100 kb in mammals [1]. Moreover, TFs interact with other signaling molecules that inhibit or induce transcription. The molecular mechanisms involved in the over-all regulation of transcription remain as one of the most challenging problems in molecular biology.

The identification of highly correlated transcripts in bone is likely a valuable hypothesis generating tool, since they are expected to be similarly regulated and/or involved in the same signaling pathways [2,3]. These associated transcripts may help to identify novel TFs, and their gene targets of importance for bone metabolism. A previous study of 3 datasets of gene expression in immortalized B cells from normal individuals showed that correlations in expression levels could be used to predict gene function [4]. Software analyzing molecular interactions have been made to explore the complex biology of expression data as well as signifying highly correlating transcripts, e.g. [5-7]. However, they have various shortcomings, such as not being user friendly as the dataset intended for study must be normalized, uploaded and then specifically adapted to the program [5]. Another program enables online identification of positively or negatively correlating pairs of microRNAs and mRNAs in large datasets, but does not detect correlating pairs of mRNAs [6]. In a third alternative program the correlations can not be restricted to bone and neither p-values nor adjusted p-values are calculated [7].

We have performed global correlation analysis on transcription profiles from trans-iliacal bone biopsies of post menopausal women aiming at detecting biological meaningful interactions. Trans-iliacal bone biopsies is of particular value as they represent a complex and rather inaccessible tissue comprising precursor and mature normal bone cells; osteocytes, osteoblasts and osteoclasts, as well as populations of multipotent immunological and hematological mesenchymal stem cells at different stages of differentiation. Based on the global correlation matrix for bone, we developed a web-based online browser that can be used to explore correlations for specific genes of interest in human bone.

We show that the current giant matrix and web browser appear to be a valuable tool for exploring various transcriptional regulations experimentally documented as interactions at the macromolecular levels and provides easy and unique access to the human bone transcriptome.

2. Materials and methods

2.1. Bone donors and biopsies

Norwegian women (50–86 years) representing a cohort with varying bone mineral densities (BMDs) and free of primary diseases known to affect the skeleton were consecutively recruited at the Lovisenberg Deacon Hospital, the Out-patient Clinic, Oslo [8]. Briefly, out of 301 registered women two thirds were rejected due to medication or secondary diseases other than primary postmenopausal osteoporosis that could affect bone metabolism. Thus, 84 trans-iliacal bone biopsies were obtained from os ileum at the same location (2.0 cm from crista iliaca and 2 cm from spina iliaca). All bone biopsies were performed in the morning on fasting individuals [9]. The biopsies were cylindrical, 0.8 cm in diameter with average length of about 1.5 cm. Connective tissue and muscle were removed before freezing in liquid nitrogen. For the replica study, 13 bone biopsies from vertebral lamina were obtained from British Caucasian males as described [10].

2.2. RNA purification and gene expression analysis

Total RNA was purified and subjected to analysis on HG-U133 plus 2.0 chips (Affymetrix) as described [8].

Reproducibility of the analysis on HG-U133 plus 2.0 chips was controlled by TaqMan Gene Expression analysis (real-time RT-PCR validation) (Applied Biosystems) of 21 selected transcripts according to the manufacturer's instructions with the ribosomal protein L41 (RPL41) as internal standard [8]. The data for postmenopausal biopsies have been submitted to the European Bioinformatics Institute (EMBL-EBI) ArrayExpress repository, ID: E-MEXP-1618, while the data on male biopsies are accessible from the same site through accession number: E-MEXP-2219.

2.3. Ethics

Prior to the study, validation and recommendation were obtained by the Norwegian and Northumberland Local Research Ethics Committees (REK 2010/2539 and REC 04/Q0902/29, respectively). All sampling and procedures followed the law in both countries, and all subjects gave their informed written consent.

2.4. Calculations, statistics and the web application

Pearson product-moment correlation coefficients (r) were computed between expressions of all genes (~23 000 probesets) across 84 women using log transformed signal values and saved in a database along with their corresponding p-values. A web application publicly available at <http://app.uio.no/med/klinmed/correlation-browser/iliac-v1.1/> was programmed in order to access the database and flexibly search for correlations of interest. Search results are returned together with raw and Bonferroni-corrected p-values and a measure of the false discovery rate (FDR) as estimated by the Benjamini & Hochberg procedure [11]. This procedure has been shown to control FDR when the tests are independent or positively correlated [12], a reasonable assumption when identifying differentially expressed genes. The browser program is further described in the Supplementary.

3. Results

3.1. Expressional profiling and web application

Using HG-U133 plus 2.0 microarrays and the Affymetrix MAS 5.0 software, we first identified the global human iliac bone transcriptome in 84 postmenopausal women as described previously [8]. We then generated matrixes of all ~23000 by ~23000 pairwise Pearson correlations and their corresponding p-values using the statistics package R. A web application, the correlation browser, was created to facilitate the assessment of specific and significant associations within this matrix. It is publicly available at <http://app.uio.no/med/klinmed/correlation-browser/iliac-v1.1/> and also calculates Bonferroni corrected p-values or FDR q-values as described in Materials and Methods and Supplementary. The user interphase of the browser is depicted in Fig. 1.

3.2. Verification of correlations to GR, SMAD4, CREB1

To test whether gene interactions identified by *in-vitro* cell culture studies reflected correlations *in vivo* in human bone, we looked for well documented reference studies identifying genes directly regulated by specific TFs. We then studied whether the relationship between the TFs and their regulated genes were reproduced as significant correlations in our giant matrix and in an independent set of male vertebral bone biopsies. Thus, reference studies had to be selected that identified TF binding promoter regions and corresponding genes. Such studies involve chromosome immunoprecipitation and often additional verification by cloning of the identified promoter elements into reporter plasmids, transfection of mammalian cells and activation of the regulating TF.

We selected 3 reference studies that identified a relatively large number of TF target genes, and that also were of relevance to bone. These studies identified target genes of the TFs glucocorticoid receptor (GR), SMAD4 and CREB1 [13-15].

3.3. Verification of correlations to *GR* mRNA

Seventy different genes were identified as GR (encoded by the *NR3C1* gene) targets by chromosome immunoprecipitation from A549 adenocarcinomic human alveolar basal epithelial cells, followed by luciferase assays of the cloned response elements in reporter plasmids [13]. Of the 70 genes, promoter elements from 50 and 20 were activated or repressed, respectively, upon stimulation of the transfected cells with dexamethasone for GR induction. In the iliac bone biopsies, transcripts representing 40 out of the 70 genes reached detection level. Our correlation analysis showed that 31 of them (>77 % of those present) were significantly ($p < 0.05$) correlated to glucocorticoid receptor mRNA expression levels after Bonferroni correction for multiple testing (Table 1). Interestingly, more than half of the transcripts ($n=21$) showed inverse associations (marked as *) as compared to the reference study [13]. The iliacal correlation r -values were well reproduced in the male vertebral biopsies, with somewhat less significance, probably due to the more limited number of samples.

Furthermore, we made a search for the 50 topmost correlated transcripts to *GR* in our giant matrix. We used the manually curated database of genes containing *GR* binding elements (GREs) from BIOBASE (Biobase, Beverly, MA) (<http://biobase-international.com/>) to identify that 22 out of the 50 topmost *GR* correlated genes contained GREs (Supplementary Table S1).

3.4. Verification of correlations to *SMAD4* mRNA

Qin et al. [14] identified *SMAD4* targets in human ovarian surface epithelium by first identifying *SMAD4* binding promoter elements by Chip-Chip analysis. 150 TGF β /*SMAD* responsive *SMAD4* binding genes were then verified by expression profiling of immortalized ovarian surface epithelial (IOSE) cells treated with TGF β for 0, 3, 6 or 12 hrs. Out of 47

genes regulated at two or more different time periods, 38 (> 80%) were correlated to *SMAD4* mRNA levels in the iliac bone biopsies as judged by significant p-values after Bonferroni correction for multiple testing (Table 2). The correlation values of the replica male vertebral biopsies were in good agreement with those from postmenopausal iliac crest.

3.5. Verification of correlations to *CREB1* mRNA

Many genes contain also CREB1 binding elements in their promoter regions, and thus are activated by forskolin via CREB1. A well documented study by Impey et al. [15] was selected for evaluation versus the human bone data. The study identified 70 cAMP regulated genes that bound CREB1 in their promoter regions. Of these, 37 reached detection level in the iliac bone biopsies and 33 (> 89%) were significantly correlated to *CREB1* mRNA levels in human bone after Bonferroni correction for multiple testing (Table 3). Also here the correlation values of the replica male vertebral biopsies were in good agreement with those from postmenopausal bone.

3.6. Further controls of the relevance of giant matrix correlations

More than 95 % of bone cells are osteocytes [16], and analysis of expressional correlations after osteocyte specific gene ablation was used as one control. We tested whether transcripts that changed expression in 6 week old mice femurs upon osteocyte specific ablation of the *Pkd1* (polycystic kidney disease 1) gene [17] were reproduced as significant correlations in human iliac bone. Significant correlations were obtained for all of the 7 mRNAs that reached detection levels in the human biopsies (those encoded by *SOST*, *SPP1*, *DMP1*, *MMP13*, *IBSP*, *BGLAP* and *RUNX2*). Of these *SOST* and *SPP1* showed similar correlations as found in mice, while the rest was inversely correlated (data not shown).

Fig. 1 identifies the transcripts most significantly correlating to *SOST* mRNA having “transcription” as GO keyword. The resulting list of topmost *SOST* associations represents mostly mRNAs of genes involved in transcriptional regulation of *SOST*. Several of these are known to regulate *SOST* transcription. *RUNX2* binds to response elements in the *SOST* promoter [18]. *SATB2* may also be listed as a *SOST* regulating transcription factor as it interacts with and enhance the activity of *RUNX2* [19]. Furthermore, all but 3 of the 9 different transcripts most significantly correlated to *SOST mRNA* (Fig. 1) are known to be related to the Wnt signaling pathway (*SFRP4*, *DVL3*) and/or osteoblast development/differentiation (*FOXC2*, *SATB2*, *DLX5*, *RUNX2*) [20, 21].

4. Discussion

The vast majority of the *in vitro* experimentally documented transcriptional regulatory effects of ubiquitously distributed and /or generally important gene regulators (GR, SMAD4 and CREB1) could be reproduced in bone biopsies as highly significant associations between transcription factor mRNAs and mRNAs of regulated genes, and were reproduced in male vertebral biopsies. The reference studies were performed in bone unrelated tissue and in cell cultures, thus giving a general validity to the results [13-15].

In the replication set we found a very high similarity between the direction in correlations between male spine and female iliac biopsies. Of 38 significant correlations in males (sum of Tables 1 - 3), only 1 showed inverse correlation as compared to the situation in females.

We found that some of the giant matrix correlations were inversely correlated when compared to the published experimental data (Table 1, 2 and 3). However, this is not altogether surprising given the complex nature of the cell milieu of trabecular bone in terms of variety of cell types at various stages of differentiation, and known paradoxical action of growth factors and hormones. This may indicate cell or tissue specific transcriptional regulation due to use of different *in vivo/in vitro* systems. This suggestion is supported by our finding from similar blood mRNA correlation studies using global gene expression analysis, in which several genes showed inverse correlations compared to the ones described in Table 1, 2 and 3 (not shown). It is hardly surprising that gene induction or repression as the net effect of binding of a TF to DNA and/or protein-protein interactions may depend on cell type and/or cell differentiation state. The dual functions of TFs as repressor or activator for the same gene have in fact been shown experimentally e.g. for the *Drosophila* homeodomain proteins, Vnd [22], and NkX6 [23]. The contextual interactions determine whether they act as repressors or activators. Also the estrogen receptor α can function as activator or repressor

depending on differential splicing [24]. The data presented in this paper indicate that GR, SMAD4 and CREB1 may also have dual function.

Identification of GREs in 22 out of the top 50 *GR* correlated genes (Supplementary Table S1) is in agreement with a recent study [25] combining identification of genome wide GR binding elements with identification of dexamethasone induced transcripts in A549 cells. The study revealed that 62% of dexamethasone-induced GR binding sites contain GRE [25]. GR makes heterodimers with other TFs, e.g. those binding to AP-1 sites, and thereby by protein-protein interaction indirectly provide DNA binding for the GR [26], thus explaining while not all GR binding promoters contain GREs.

Like the GR, also SMAD4 can indirectly bind to DNA. SMAD4 can bind through its DNA-binding R-SMAD partner and/or other TFs like RUNX2 or Schnurri [27], thereby providing a plausible explanation why some sites were not picked up in the reference study [14].

Some regulatory interactions found in cell based systems could not be reproduced as significant p-values in the matrix, possibly due to transcriptional regulators or posttranslational modifications present or not present within the human bone biopsies. Thus, the present data also show the necessity to study human bone tissue for understanding species specific gene transcription and macromolecular interactions leading to specific physiological and pathophysiological system effects.

Fig. 1, identifying transcripts of genes most significantly correlated to *SOST* mRNA as genes that are known to be of crucial importance for bone metabolism, should indicate that significant correlating transcripts represent a snapshot of the human bone gene interactome at the transcript level.

Our results from female iliac bone, and confirmed in male vertebrae demonstrate that a global transcriptional correlation matrix, combined with an easy-access browser application,

should be generally useful as a first identification of candidate interacting genes in bone, and in fact be of relevance in both genders. In conclusion, we have demonstrated that the *in silico* browser is useful to refine microchips data through correlations to experimental results and thus discover important macromolecular interactions in human bone.

Acknowledgements

The present work is dedicated to all the female and male bone donors who voluntarily participated and made this study possible. We acknowledge the help from Drs Lise Sofie Nissen-Meyer, University of Oslo, and Ingegerd Frøyshov Larsen and Hilde Refvem at the Deacon Hospital Oslo and the Newcastle upon Tyne Hospital NHS Trust in the patient registration. We are indebted to the Medical Out-patient Clinic, the Deacon Hospital Oslo, and NGH, Newcastle upon Tyne for access to surgical facility and patient care assistance. The Osteoporosis Laboratory, the Deacon Hospital, by Bioengineer Anne Runningen is especially gratefully acknowledged. The authors have no conflicting financial interests. This work was supported by the 6th EU framework program to “OSTEOGENE” [LSHM-CT-2003-502941]; the South-Eastern Norway Regional Health Authority; and Oslo University Hospital, Ullevaal.

Figure legend

Fig. 1. Interphase of the correlation browser. A typical search starts with inserting an identifier in the first window of “search option”, either EntrezGene ID (e.g. "1234"), an Accession Number (e.g. "BE644809" or "NM_005715"), a Gene Symbol (e.g. "*SOST1*" or "*DKK1*", not case-sensitive) or an Affymetrix probeset ID (e.g. "1562529_s_at"). Then, a specific transcript can be searched for by inserting a second identifier in the second window under “search option”. Alternatively, the window may be left open to obtain a list of the transcripts most significantly correlating to the identifier in the first window. Filling in boxes in the “output options” fields enable restriction of output to e.g. transcripts having specific keywords in Gene ontologies (GO), TFs (genes having “transcription” as part of the Gene Title) or only positive or negative correlations.

Tables

Table 1. Correlations of Glucocorticoid receptor (*NR3C1*) mRNA with Glucocorticoid receptor regulated transcripts expressed in bone biopsies

Gene symbol	Name	Iliac biopsies (n=84)			Vertebral biopsies (n=13)	
		r	p	P [#]	r	p
PDE4B	phosphodiesterase 4B, cAMP-specific	0.80*	6.2E-20	5.0E-19	0.68*	9.8E-03
SDPR	serum deprivation response	0.69	2.3E-13	1.9E-12	0.38	2.0E-01
FGD4	FYVE, RhoGEF and PH domain containing 4	0.63	1.3E-10	1.1E-09	NA	NA
AMIGO2	adhesion molecule with Ig-like domain 2	0.62*	1.3E-9	1.3E-09	0.70*	7.5E-03
C19orf42 (Stomatin)	chromosome 19 open reading frame 42	0.60	1.7E-9	4.2E-08	0.79	1.3E-03
FKBP5	FK506 binding protein 5	0.59	2.7E-9	2.2E-08	0.72	5.6E-03
S100P	S100 calcium binding protein P	-0.56*	3.3E-8	1.3E-07	-0.75*	3.0E-03
CPEB4	cytoplasmic polyadenylation element binding protein 4	-0.55*	6.5E-8	1.0E-06	-0.40*	1.8E-01
ARRDC3 (KIAA1376)	arrestin domain containing 3	0.53*	2.8E-7	1.1E-06	0.54*	5.6E-02
CXCR7 (RDC1)	chemokine (C-X-C motif) receptor 7	0.52*	3.5E-7	1.4E-06	0.83*	4.0E-04
CDC42EP3	CDC42 effector protein (Rho GTPase binding) 3	0.52	4.0E-7	4.7E-06	0.66	1.4E-02
SLC19A2	solute carrier family 19 (thiamine transporter), member 2	0.52	5.0E-7	2.0E-06	0.04	9.0E-01
TGFBR3	transforming growth factor, beta receptor III	0.50	1.4E-6	1.1E-05	0.57	4.4E-02
AKAP13	A kinase (PRKA) anchor protein 13	-0.47*	6.0E-6	1.7E-04	0.39	1.9E-01
GADD45b	growth arrest and DNA-damage-inducible, beta	-0.46*	8.9E-6	1.1E-04	-0.30*	3.2E-01
CDKN1C (Kip2/p57)	cyclin-dependent kinase inhibitor 1C (p57, Kip2)	0.46	1.1E-5	1.3E-04	0.81	8.0E-04
IRS2	insulin receptor substrate 2	0.46	1.4E-5	1.1E-04	0.45	1.2E-01
DNAJC15 (MCJ)	DnaJ (Hsp40) homolog, subfamily C, member 15	0.46	1.4E-5	1.1E-04	0.76	2.8E-03
SERPINB9(CAP3/IP9)	serpin peptidase inhibitor, clade B (ovalbumin), member 9	0.45*	1.7E-5	1.4E-04	0.14*	6.5E-01
PMP22 (PMP2)	peripheral myelin protein 22	0.45*	1.7E-5	6.7E-05	0.85*	2.0E-04
TSC22D3 (GILZ)	TSC22 domain family, member 3	0.40	2.0E-4	7.9E-04	0.46	1.1E-01
ABHD2	abhydrolase domain containing 2	-0.39*	2.3E-4	5.5E-03	-0.46*	1.1E-01
EPB41L4B (EHM2)	erythrocyte membrane protein band 4.1 like 4B	0.38	3.4E-4	2.7E-03	0.35	2.5E-01
GEM	GTP binding protein overexpressed in skeletal muscle	0.37*	5.3E-4	2.1E-03	0.53*	6.4E-02
MT1X (MT-1I)	metallothionein 1X	0.35	1.2E-3	9.7E-03	0.57	4.2E-02
ENC1	ectodermal-neural cortex 1 (with BTB-like domain)	0.30*	1.9E-3	7.5E-03	0.63*	2.0E-02
BHLHE40 (BHLHB2)	basic helix-loop-helix family, member e40	0.33*	2.3E-3	9.1E-03	0.68*	1.1E-02
PRRG4 (TMG4)	proline rich Gla (G-carboxyglutamic acid) 4 (transmembrane)	0.32	3.0E-3	2.4E-02	0.06	8.3E-01
PLK2 (SNK)	polo-like kinase 2	0.31*	4.7E-3	1.9E-02	0.71*	6.8E-03
THBD	Thrombomodulin	0.30	5.7E-3	2.3E-02	0.66	1.5E-02
B3GNT5	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 5	0.29	1.9E-3	3.2E-02	-0.58*	4.0E-02
SEC14L1	SEC14-like 1 (<i>S. cerevisiae</i>)	-0.30*	6.3E-3	7.5E-02	-0.54*	5.4E-02
TNFAIP3	tumor necrosis factor, alpha-induced protein 3	0.28	9.0E-3	7.2E-02	0.40	1.8E-01
SLC26A2	solute carrier family 26 (sulfate transporter), member 2	-0.28*	1.1E-2	1.3E-01	0.23	4.5E-01
MGAM	maltase-glucoamylase (alpha-glucosidase)	-0.23*	3.5E-2	1.4E-01	-0.49*	8.9E-02
RGS2	regulator of G-protein signaling 2, 24kDa	-0.22*	4.0E-2	1.6E-01	-0.32*	2.9E-01
FLVCR2 (FLJ20371)	feline leukemia virus subgroup C cellular receptor family, member 2	0.22	4.9E-2	2.0E-01	0.15	6.3E-01
LRRC8A (LRRC8)	leucine rich repeat containing 8 family, member A	-0.20*	6.9E-2	2.7E-01	0.07	8.3E-01
PTGS2 (COX-2, CG1/XP28)	prostaglandin-endoperoxide synthase 2	0.11*	3.3E-1	1.3E+0	0.11*	7.2E-01
CCL2	chemokine (C-C motif) ligand 2	-0.08	4.6E-1	1.9E+0	0.01*	9.7E-01

“*” Inverse correlations as compared to regulation in the reference study [13]. Parentheses denotes synonymous gene names used in the reference study. In cases where multiple probesets represent the same gene, only the most significant correlation is shown. Identical probesets were compared in the main and replica studies. Shaded cells represent significant (p or p[#] <0.05) correlations. p[#]: Bonferroni corrected p-values. NA: not applicable (undetected in the dataset)

Table 2. Correlations of *SMAD4* mRNA with *SMAD4* regulated transcripts expressed in bone biopsies

Gene symbol	Name	Iliac biopsies (n=84)			Vertebral biopsies (n=13)	
		r	p	p [#]	r	p
CSNK1A1	casein kinase 1, alpha 1	0.81	6.3E-21	2.1E-19	0.83	5.0E-04
STK38L	serine/threonine kinase 38 like	0.75	1.4E-16	8.3E-16	0.62	2.3E-02
UBE2D3	ubiquitin-conjugating enzyme E2D 3 (UBC4/5 homolog, yeast)	0.73	4.4E-15	5.3E-14	0.35	2.4E-01
C6orf62	chromosome 6 open reading frame 62	0.71	7.0E-14	8.3E-13	0.66	1.5E-02
GPHN	Gephyrin	0.66*	5.7E-12	3.4E-11	0.71*	6.3E-03
SRPK2	SFRS protein kinase 2	0.64	5.4E-11	6.5E-10	0.42	1.5E-01
ARHGAP5	Rho GTPase activating protein 5	0.63	1.6E-10	1.5E-09	0.78	1.6E-03
NFAT5	nuclear factor of activated T-cells 5, tonicity-responsive	0.60	2.1E-09	1.3E-08	0.36	2.3E-01
ZNF638	zinc finger protein 638	0.60	1.3E-09	1.9E-08	0.61	2.6E-02
PPP4R2	protein phosphatase 4, regulatory subunit 2	0.59	4.1E-09	2.5E-08	0.45	1.3E-01
HERC1	hect domain and RCC1 (CHC1)	0.57	1.3E-08	3.8E-08	0.28	3.6E-01
MTUS1	microtubule associated tumor suppressor 1	0.58*	1.1E-08	6.3E-08	0.75*	3.2E-03
FNTB	Farnesyltransferase, CAAX box, beta	0.56	2.5E-08	1.5E-07	0.70	7.9E-03
KLF9	Kruppel-like factor 9	0.56*	3.1E-08	1.8E-07	0.53*	6.1E-02
ARL2BP	ADP-ribosylation factor-like 2 binding protein	0.52	3.9E-07	1.2E-06	0.82	6.0E-04
BDNF	brain-derived neurotrophic factor	-0.46	9.2E-06	2.8E-05	NA	NA
KLF12	Kruppel-like factor 12	0.48*	5.0E-06	3.0E-05	0.70*	7.7E-03
TGFB2	transforming growth factor, beta receptor II (70/80kDa)	0.45*	1.4E-05	4.2E-05	0.26*	3.8E-01
IL1B	interleukin 1, beta	0.43	4.6E-05	1.4E-04	0.26	3.9E-01
GTF2H2	general transcription factor IIH, polypeptide 2, 44kDa	0.43	4.8E-05	1.4E-04	0.29	3.4E-01
PELI2	pellino homolog 2 (Drosophila)	0.42	5.8E-05	1.7E-04	0.75	3.0E-03
FBXO32	F-box protein 32	0.43	3.8E-05	2.3E-04	0.58	3.7E-02
LTBP2	latent transforming growth factor beta binding protein 2	0.43	5.3E-05	3.2E-04	0.30	3.1E-01
MICAL2	microtubule associated monooxygenase, calponin and LIM domain containing 2	-0.42*	7.3E-05	4.4E-04	-0.48*	9.6E-02
ADAM19	ADAM metalloproteinase domain 19 (meltrin beta)	0.40	1.7E-04	5.0E-04	0.28	3.5E-01
SOX4	SRY (sex determining region Y)-box 4	0.39	2.7E-04	8.0E-04	0.49	9.1E-02
TTC14	tetratricopeptide repeat domain 14	-0.40*	1.6E-04	9.8E-04	-0.22*	4.8E-01
SYTL2	synaptotagmin-like 2	0.39*	2.2E-04	1.3E-03	0.73*	4.4E-03
TNFSF13B	tumor necrosis factor (ligand) superfamily, member 13b	-0.37	4.6E-04	2.7E-03	-0.30	3.2E-01
DMBT1	deleted in malignant brain tumors 1	-0.35	1.2E-03	3.7E-03	-0.23	4.6E-01
SMAD3	SMAD family member 3	0.30*	5.1E-03	1.5E-02	0.89*	2.3E-04
PREX1	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 1	0.31*	3.7E-03	2.2E-02	0.35*	2.4E-01
SP110	SP110 nuclear body protein	0.34*	1.5E-03	2.3E-02	0.36*	2.2E-01
TNC	Tenascin C	0.31	4.0E-03	2.4E-02	0.05	8.6E-01
DNM3	dynammin 3	-0.30*	4.8E-03	2.9E-02	0.22	4.8E-01
UGDH	UDP-glucose 6-dehydrogenase	0.28	1.1E-02	3.2E-02	0.59	3.5E-02
LRRC20	leucine rich repeat containing 20	0.27*	1.2E-02	3.5E-02	0.31*	3.0E-01
NASP	nuclear autoantigenic sperm protein (histone-binding)	-0.30*	5.3E-03	4.8E-02	0.41	1.6E-01
STARD4	StAR-related lipid transfer (START) domain containing 4	0.25	2.4E-02	7.1E-02	0.78	1.7E-03
ABAT	4-aminobutyrate aminotransferase	0.26	1.9E-02	1.1E-01	0.25	4.1E-01
EPB41L4B	erythrocyte membrane protein band 4.1 like 4B	0.24*	2.5E-02	1.5E-01	NA	NA
OAS1	2',5'-oligoadenylate synthetase 1, 40/46kDa	0.24*	2.6E-02	1.6E-01	-0.22	4.8E-01
APOL6	apolipoprotein L, 6	0.20*	6.2E-02	1.9E-01	0.50*	8.1E-02
MSN	Moesin	-0.18*	9.9E-02	3.0E-01	-0.22*	4.6E-01
C6orf105	chromosome 6 open reading frame 105	0.15	1.7E-01	5.0E-01	0.01	9.7E-01
EPSTI1	epithelial stromal interaction 1 (breast)	-0.18	9.6E-02	5.8E-01	0.17*	5.8E-01
SEMA4B	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM)	-0.06	5.8E-01	1.7E+00	-0.36	2.2E-01

“*” Inverse correlations as compared to regulation in the reference study [14]. In cases where multiple probesets represent the same gene, only the most significant correlation is shown. Identical probesets were compared in the main and replica studies. Shaded cells represent significant (p or p[#] < 0.05) correlations. p[#]: Bonferroni corrected p-values NA: not applicable (undetected in the dataset)

Table 3. Correlations of *CREB1* mRNA with forskolin regulated transcripts expressed in bone biopsies

Gene symbol	Name	Iliac biopsies (n=84)			Vertebral biopsies (n=13)	
		r	p	p [#]	r	p
GSK3B (GSK3β)	glycogen synthase kinase 3 beta	0.79	2.78E-19	5.85E-18	0.88	2.5E-05
MAT2A (MAT IIA)	methionine adenosyltransferase II, alpha	0.75	1.37E-16	2.87E-15	0.94	5.0E-05
PSEN1 (presenilin-1)	presenilin 1	0.75	1.78E-16	4.98E-15	0.96	2.5E-05
CIRBP	cold inducible RNA binding protein	-0.69*	2.36E-13	8.25E-12	-0.54*	5.7E-02
AZIN1 (Oazi)	antizyme inhibitor 1	0.68	8.21E-13	1.15E-11	0.65	1.7E-02
ARRB1 (β arrestin 1)	arrestin, beta 1	-0.59*	2.50E-09	3.50E-08	-0.53*	6.4E-02
JUND (JunD)	jun D proto-oncogene	-0.57*	1.48E-08	1.03E-07	-0.07*	8.2E-01
MAF (c-maf)	v-maf musculoaponeurotic fibrosarcoma oncogene homolog (avian)	0.55	5.81E-08	1.22E-06	0.03	9.2E-01
NUP98 (Nup98)	nucleoporin 98kDa	0.51	5.81E-07	1.22E-05	0.87	3.9 E-04
PPP1R15A (Myd116)	protein phosphatase 1, regulatory (inhibitor) subunit 15A	-0.45*	1.96E-05	2.75E-04	0.04	8.9E-01
BMP2	bone morphogenetic protein 2	0.44	2.25E-05	3.15E-04	0.45	1.3E-01
EGLN1 (Egl 9 homolog)	egl nine homolog 1	-0.43*	3.71E-05	7.78E-04	-0.25*	4.2E-01
Id3	inhibitor of DNA binding 3	-0.43*	3.84E-05	2.69E-04	-0.49*	8.8E-02
MEF2D (Mef2D)	myocyte enhancer factor 2D	0.43	3.89E-05	2.72E-04	-0.35*	2.4E-01
NDEL1 (Nudel)	nudE nuclear distribution gene E homolog (A. nidulans)-like 1	0.42	6.63E-05	4.64E-04	-0.29*	3.4E-01
TUBB3 (βIII tubulin)	tubulin, beta 3	-0.42*	7.25E-05	1.02E-03	0.22	8.9E-01
Hes1	hairly and enhancer of split 1, (Drosophila)	-0.42*	7.90E-05	1.11E-03	-0.24*	4.4E-01
Pim1	pim-1 oncogene	0.41	1.07E-04	7.49E-04	0.15	6.2E-01
JunB	jun B proto-oncogene	-0.40*	1.38E-04	9.67E-04	-0.07*	8.1E-01
VEGFA (VEGF)	vascular endothelial growth factor A	0.40	1.43E-04	2.00E-03	0.16	6.1E-01
EGR1	early growth response 1	-0.40*	1.82E-04	3.82E-03	0.47	1.1E-01
CAV1 (Caveolin)	caveolin 1, caveolae protein, 22kDa	-0.39*	2.36E-04	3.30E-03	0.09	7.7E-01
NFIL3 (Nfil3)	nuclear factor, interleukin 3 regulated	0.39	2.50E-04	1.75E-03	0.42	1.5E-01
Dyrk1a	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A	0.39	2.95E-04	2.07E-03	0.49	8.7E-02
RhoB	ras homolog gene family, member B	-0.38*	3.03E-04	2.12E-03	-0.27*	3.7E-01
Id2	inhibitor of DNA binding 2	0.37	4.73E-04	6.62E-03	0.36	2.3E-01
DLGAP4 (DAP-4)	discs, large (Drosophila) homolog-associated protein 4	0.34	1.49E-03	3.13E-02	-0.36*	2.3E-01
IRS1	insulin receptor substrate 1	-0.34*	1.61E-03	1.12E-02	-0.13*	6.7E-01
CEBPB (C/EBPβ)	CCAAT/enhancer binding protein (C/EBP), beta	-0.34*	1.74E-03	1.22E-02	-0.19*	5.4E-01
PER1 (Per1)	period homolog 1 (Drosophila)	-0.33*	2.22E-03	1.55E-02	0.25	4.1E-01
KLF10 (TIEG1)	Kruppel-like factor 10	0.32	2.79E-03	1.95E-02	0.38	4.7E-01
PDE4D	phosphodiesterase 4D, cAMP-specific	0.31	3.72E-03	7.81E-02	0.38	2.0E-01
BHLHE40 (Bhlhb2)	basic helix-loop-helix family, member e40	-0.31*	4.53E-03	3.17E-02	-0.32*	2.8E-01
ID1 (Id1)	inhibitor of DNA binding 1	-0.29*	6.55E-03	4.58E-02	-0.49*	8.8E-02
DUSP10 (MKP5)	dual specificity phosphatase 10	0.27	1.16E-02	8.13E-02	0.44	1.3E-01
KLF4 (Klf4)	Kruppel-like factor 4 (gut)	-0.27*	1.18E-02	8.23E-02	0.37	2.2E-01
FOS (c-fos)	FBJ murine osteosarcoma viral oncogene homolog	-0.25*	2.29E-02	1.60E-01	-0.15*	6.3E-01
BTG2 (Btg2)	BTG family, member 2	-0.12*	2.97E-01	2.08E+00	0.16	6.0E-01

“*” Inverse correlations as compared to regulation in the reference study [15]. In cases where multiple probesets represent the same gene, only the most significant correlation is shown. Identical probesets were compared in the main and replica studies. Shaded cells represent significant (p or p[#] <0.05) correlations. p[#]: Bonferroni corrected p-values

References

- [1] Levine M, Tjian R. Transcription regulation and animal diversity. *Nature* 424 (2003) 147-151.
- [2] Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D et al. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat Genet.* 34 (2003) 166-176.
- [3] Zeng T, Li J. Maximization of negative correlations in time-course gene expression data for enhancing understanding of molecular pathways. *Nucleic Acids Res.* 38 (2010) e1.
- [4] Nayak RR, Kearns M, Spielman RS, Cheung VG. Coexpression network based on natural variation in human gene expression reveals gene interactions and functions. *Genome Res.* 19 (2009) 1953-1962.
- [5] Zhang M, Ouyang Q, Stephenson A, Kane MD, Salt DE, Prabhakar S et al. Interactive analysis of systems biology molecular expression data. *BMC Syst Biol.* 2 (2008) 23.
- [6] Ritchie W, Flamant S, Rasko JE. mimiRNA: a microRNA expression profiler and classification resource designed to identify functional correlations between microRNAs and their targets. *Bioinformatics* 26 (2010) 223-227.
- [7] Jupiter D, Chen H, VanBuren V. STARNET 2: a web-based tool for accelerating discovery of gene regulatory networks using microarray co-expression data. *BMC Bioinformatics* 10 (2009) 332.
- [8] Reppe S, Refvem H, Gautvik VT, Olstad OK, Hovring PI, Reinholt FP et al. Eight genes are highly associated with BMD variation in postmenopausal Caucasian women. *Bone* 46 (2010) 604-612.
- [9] Reppe S, Stilgren L, Olstad OK, Brixen K, Nissen-Meyer LS, Gautvik KM et al. Gene expression profiles give insight into the molecular pathology of bone in primary hyperparathyroidism. *Bone* 39 (2006) 189-198.
- [10] Kalogeropoulos M, Varanasi SS, Olstad OK, Sanderson P, Gautvik VT, Reppe S et al. Zic1 transcription factor in bone: neural developmental protein regulates mechanotransduction in osteocytes. *FASEB J.* 24 (2010) 2893-2903.
- [11] Y.Benjamini and Y.Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B.* 57 (2008) 289-300.
- [12] Benjamini Y, Yekutieli D. The Control of the False Discovery Rate in Multiple Testing Under Dependency. *Ann Statist.* 29 (2001) 1165-1188.
- [13] Wang JC, Derynck MK, Nonaka DF, Khodabakhsh DB, Haqq C, Yamamoto KR. Chromatin immunoprecipitation (ChIP) scanning identifies primary glucocorticoid receptor target genes. *Proc Natl Acad Sci U S A* 101 (2004) 15603-15608.

- [14] Qin H, Chan MW, Liyanarachchi S, Balch C, Potter D, Souriraj IJ et al. An integrative ChIP-chip and gene expression profiling to model SMAD regulatory modules. *BMC Syst Biol.* 3 (2009) 73.
- [15] Impey S, McCorkle SR, Cha-Molstad H, Dwyer JM, Yochum GS, Boss JM et al. Defining the CREB regulon: a genome-wide analysis of transcription factor regulatory regions. *Cell* 119 (2004) 1041-1054.
- [16] Rochefort GY, Pallu S, Benhamou CL. Osteocyte: the unrecognized side of bone tissue. *Osteoporos Int.* 21 (2010) 1457-1469.
- [17] Xiao Z, Dallas M, Qiu N, Nicolella D, Cao L, Johnson M et al. Conditional deletion of *Pkd1* in osteocytes disrupts skeletal mechanosensing in mice. *FASEB J.* 25 (2011) 2418-2432.
- [18] Severson B, Taylor S, Pan Y. *Cbfa1/RUNX2* directs specific expression of the sclerostosis gene (*SOST*). *J Biol Chem.* 279 (2004) 13849-13858.
- [19] Dobrev G, Chahrour M, Dautzenberg M, Chirivella L, Kanzler B, Farinas I et al. *SATB2* is a multifunctional determinant of craniofacial patterning and osteoblast differentiation. *Cell* 125 (2006) 971-786.
- [20] Long F. Building strong bones: molecular regulation of the osteoblast lineage. *Nat Rev Mol Cell Biol.* 13 (2012) 27-38.
- [21] Monroe DG, McGee-Lawrence ME, Oursler MJ, Westendorf JJ. Update on Wnt signaling in bone cell biology and bone disease. *Gene* 492 (2012) 1-18.
- [22] Yu Z, Syu LJ, Mellerick DM. Contextual interactions determine whether the *Drosophila* homeodomain protein, *Vnd*, acts as a repressor or activator. *Nucleic Acids Res.* 33 (2005) 1-12.
- [23] Syu LJ, Uhler J, Zhang H, Mellerick DM. The *Drosophila* *Nkx6* homeodomain protein has both activation and repression domains and can activate target gene expression. *Brain Res.* 1266 (2009) 8-17.
- [24] Bollig A, Miksicek RJ. An estrogen receptor-alpha splicing variant mediates both positive and negative effects on gene transcription. *Mol Endocrinol.* 14 (2000) 634-649.
- [25] Reddy TE, Pauli F, Sprouse RO, Neff NF, Newberry KM, Garabedian MJ et al. Genomic determination of the glucocorticoid response reveals unexpected mechanisms of gene regulation. *Genome Res.* 19 (2009) 2163-2171.
- [26] Herrlich P. Cross-talk between glucocorticoid receptor and AP-1. *Oncogene* 20 (2001) 2465-2475.
- [27] Ross S, Hill CS. How the Smads regulate transcription. *Int J Biochem Cell Biol.* 40 (2008) 383-408.

- [28] Hawse JR, Cicek M, Grygo SB, Bruinsma ES, Rajamannan NM, van Wijnen AJ et al. TIEG1/KLF10 modulates Runx2 expression and activity in osteoblasts. PLoS One 6 (2011) e19429.
- [29] Powell WF, Jr., Barry KJ, Tulum I, Kobayashi T, Harris SE, Bringhurst FR et al. Targeted ablation of the PTH/PTHrP receptor in osteocytes impairs bone structure and homeostatic calcemic responses. J Endocrinol. 209 (2011) 21-32.

Figure(s)
Click here to download high resolution image

Correlations in the bone dataset

Use the form to search for:

- an EntrezGene ID (e.g. "1234")
- an Accession Number (e.g. "BX644809" or "NM_005715")
- a Gene Symbol (e.g. "RMOCA" or "ZNF532", not case-sensitive)
- a variable label (e.g. "1562525_s_at" or "hsa-miR-29c")
- a variable number (e.g. "#1234" for the 1234th variable in the dataset)

(View the full list of available key terms [here]).

Search Options: Find correlations between...

Gene/Label/Number: and (Empty also means "all others")

Output Options

Restrict output to ☐ miRs, ☐ Transcription factors? (Empty: No restrictions.)

with keyword:

Limit output to rows.

Remove ☒ "self correlations", i.e. correlations of a variable with itself

Sort results by ☐ positive, ☐ negative or ☒ absolute value of correlation

Starting timer.
Connected to database on localhost via TCP/IP
Searching for user input in variable labels, Entrez IDs, Accession Numbers...
Evaluating targets, filters and restrictions...
(User has set a text filter on GO terms; may take a little while.)

Finding correlations with 1 variable:					
id	Label1	AccNum	GeneSym	Info	Gene Name
14559	223869_at	AF331844	SOST		sclerostin

Fetching correlation data, please wait...
Getting list of correlations for id 14559.
Getting list of p-values for id 14559.
1 of 1 done.
Calculating corrections for multiple testing (n=3358), please wait...
Done; whole program took 5.605 seconds. Transmitting 50 results:

Rank	Label1	Label2	AccNum	GeneSym	Info	Pearson r	Raw p	Bonferroni	FDR (BH)
1	223869_at	206115_at	NM_004430	EGR3		0.705611	6.652139e-14	2.233788e-10	2.233788e-10
2	223869_at	239058_at	BF446723	FOXC2		0.677734	1.436148e-12	4.822586e-9	2.411293e-9
3	223869_at	204051_s_at	AW089415	SFRP4		0.670923	2.892873e-12	9.714269e-9	3.238090e-9
4	223869_at	213435_at	AB028957	SATB2		0.629525	1.414949e-10	4.751398e-7	1.187850e-7
5	223869_at	204052_s_at	NM_003014	SFRP4		0.619184	3.423032e-10	1.149454e-6	2.298908e-7
6	223869_at	213707_s_at	NM_005221	DLX5		0.611629	6.396285e-10	2.147872e-6	3.579787e-7
7	223869_at	220617_s_at	NM_018181	ZNF532		0.557445	3.633108e-8	1.219998e-4	1.742854e-5
8	223869_at	201908_at	NM_004423	DVL3		0.539087	1.220065e-7	4.096979e-4	5.121223e-5
9	223869_at	232231_at	AL353944	RUNX2		0.531726	1.943944e-7	6.527763e-4	7.253070e-5
10	223869_at	211675_s_at	AF054589	MOFIC		0.530762	2.064561e-7	6.932795e-4	6.932795e-5

Supplementary Material

[Click here to download Supplementary Material: Supplementary.doc](#)