Original Article

Systems genetics analysis of mouse chondrocyte differentiation

Jaijam Suwanwela[1], Charles R. Farber[2], Bau-lin Haung[1], Buer Song[3], Calvin Pan[4], Karen M. Lyons[3] and Aldons J. Lusis[4,5,6,7]

[1]Department of Oral Biology, School of Dentistry at UCLA, Los Angeles, CA 90095, [2]Center for Public Health Genomics, Departments of Medicine and Biochemistry and Molecular Biology, University of Virginia, VA 22908, [3] Department of Molecular, Cell and Developmental Biology and Department of Orthopedic Surgery, [4]Department of Medicine, [5]Department of Human Genetics, [6]Department of Microbiology, Immunology, and Molecular Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095, [7]Molecular Biology Institute, UCLA, Los Angeles, CA 90095

Running Title: Systems genetics analysis of mouse chondrocyte differentiation

Correspondence should be addressed to:
Jaijam Suwanwela, Department of Oral Biology & Medicine, UCLA School of Dentistry, Los Angeles, CA 90095
Tel: 310-825-1955; Fax: 310-794-7109 Email: jaijam1220@gmail.com

E-mail addresses for other authors:
CRF: crf2s@eservices.virginia.edu
BLH: hbl61801@ucla.edu
KML: KLyons@mednet.ucla.edu
AJL: JLusis@mednet.ucla.edu

Additional Supporting Information may be found in the online version of this article.

Number of words in manuscript: 8,597

Number of words in abstract: 194

Number of Tables: 5

Number of Figures: (black and white = 2), (color = 7)

Conflict of Interest: All authors have no conflicts of interest.

ABSTRACT:

One of the goals of systems genetics is the reconstruction of gene networks that underlie key processes in development and disease. To identify cartilage gene networks that play an important role in bone development, we used systems genetics approach that integrated microarray gene expression profiles from cartilage and bone phenotypic data from two sets of recombinant inbred strains. Microarray profiles generated from isolated chondrocytes were used to generate weighted gene co-expression networks. This analysis resulted in the identification of subnetworks (modules) of co-expressed genes, which were then examined for relationships with bone geometry and density. One module exhibited significant correlation with femur length (r=0.416), anterio-posterior diameter (r=0.418), medial-lateral diameter (r=0.576) and bone mineral density (r=0.475). Highly connected genes (n=28) from this and other modules were tested in vitro using pre-chondrocyte ATDC5 cells and RNA interference. Five of the 28 genes were found to play a role in chondrocyte differentiation. Two of these, *Hspd1* and *Cdkn1a*, were previously known to function in chondrocyte development, while the other three, *Bhlhb9, Cugbp1* and *Spcs3*, are novel genes. Our integrative analysis provided a systems-level view of cartilage development and identified genes that may be involved in bone development.

Introduction

Mesenchymal cell condensation is the starting point of skeletal patterning, outlining the shapes of future bones[1]. In endochondral ossification, the main type of bone formation, mesenchymal cells condense and differentiate into chondrocytes to shape the early skeleton. The hyaline cartilage then becomes infiltrated with blood vessels and osteoblasts, forms periosteum, and finally becomes mineralized[2-4].

The limbs are formed from buds of mesoderm surrounded by ectodermal cells. The apical ectodermal cells regulate the distal outgrowth of the developing limb, and the dorsal ectoderm regulates dorsal–ventral patterning. The mesodermal cells subsequently differentiate into chondrocytes to form the templates of the long bones. Chondrocytes in the growth plates of long bones undergo a coordinated differentiation process. They begin to proliferate, hypertrophy, and undergo terminal differentiation. Chondrocyte differentiation is a multi-step process characterized by successive changes in cell morphology and gene expression. This multi-step process requires the coordinated expression of many genes, including genes encoding proteins for extracellular matrix (ECM), morphogenesis, proliferation, angiogenesis, and apoptosis[5,6]. Disturbances of these processes can lead to skeletal dysplasias[6,7].

We now report the application of a systems genetics approach to identify cartilage gene networks and to examine their involvement in bone development. Most of the knowledge regarding the function of genes involved in bone developmental processes has been derived from studies in animal models and cell lines as well as from the identification of disease genes in skeletal disorders. However, such studies provide insights only into the roles of individual genes or developmental pathways. In contrast, integrative genomics is an approach that links different levels of a biological system, such as the genome, transcriptome and phenome, to understand

their relationships. It analyzes all genes in the genome simultaneously. This approach organizes gene expression data into a functionally relevant framework in order to explore the physiology of skeletal development from a systems perspective[8,9].

A central concept of systems biology is that of networks consisting of nodes and edges. In co-expression networks, the nodes represent gene transcripts and the edges represent correlations between transcripts. The networks in our study were generated by co-expression analysis of mouse rib cartilage cells in 27 different recombinant inbred strains of mice. One highly coregulated subnetwork, or module, was significantly correlated with bone development parameters among strains. Genes from this module were used to identify the potential regulators of chondrocyte development[10]. Highly connected genes in this and other modules were then studied with respect to expression profiles during differentiation of a chondrocyte cell line, and those exhibiting significant developmental regulation were further examined by siRNA knockdown for effects on chondrocyte differentiation. These studies resulted in the identification of 2 genes previously shown be important in bone development as well as 3 novel genes.

Materials and methods

Animals and chondrocyte isolation

The recombinant inbred strains that have been proved to have significant different in bone traits and were available from the Jackson Laboratory (Bar Harbor, Maine, United States) were used. Those strains were C3H/HeJ, C57BL/6J, DBA/2J, B6Cc3-1/KccJ, BXH2/TyJ, BXH4/TyJ, BXH6/TyJ, BXH7/TyJ, BXH8/TyJ, BXH9/TyJ, BXH14/TyJ, BXH19/TyJ, BXH22/KccJ, BXD1/TyJ, BXD2/TyJ, BXD6/TyJ, BXD 12/TyJ, BXD16/TyJ, BXD 19/TyJ, BXD21/TyJ, BXD 24a/TyJ, BXD27/TyJ BXD 28/TyJ, BXD 32/TyJ, BXD 39/TyJ, BXD 40/TyJ and BXD 42/TyJ recombinant inbred (RI) strains. All mouse protocols were performed according to the guidelines

of the American Association for Accreditation of Laboratory Animal Care (AAALAC). Cartilage from the rib cage of one to two day-old male mice was dissected to remove bone and any adherent non-cartilage tissue for microarray analysis. The cartilage was digested in 0.3% bacterial collagenase (Invitrogen) for 10 hours and the cells were collected by centrifugation. RNA was isolated and purified using the Rneasy kit (Qiagen). It was then quantified and assessed for purity using a NanoDrop spectrophotometer (Rockland). RNA integrity was verified with a BioAnalyzer 2100 (Agilent). Due to low amount of RNA we could isolate from the rib cages, we had to pool the RNA from 3 mice from the same strain. All 27 strains were applied separately to the Illumina arrays.

Microarray experiments

Illumina Mouse-6 V1 BeadChip mouse whole-genome expression arrays were used in this study. Of the 27 RNA samples from 27 strains, 3 were hybridized twice and were used as technical replicates. A total of 200 ng of DNA-free, quality-checked RNA was amplified using the Ambion Illumina RNA amplification kit with biotin UTP (Enzo) labeling. The Ambion Illumina RNA amplification kit uses T7 oligo(dT) primers to generate single stranded cDNA followed by a second strand synthesis to generate double-stranded cDNA, which is then column purified. *In vitro* transcription was done to synthesize biotin-labeled cRNA using T7 RNA polymerase. The cRNA was then column purified. The cRNA was then checked for size and yield using the Molecular Probes Invitrogen Quant-iT RiboGreen assay. A total of 1.5 µg of cRNA was hybridized to each array using standard Illumina protocols with streptavidin-Cy3 (Amersham) being used for detection. Slides were scanned on an Illumina Beadstation and processed using BeadStudio (Illumina, Inc).

Data extraction and normalization

The R software (http://cran.r-project.org/), a system for statistical computation and graphics, was used to analyze the data[11]. Data were normalized using Lumi[12], a Bioconductor package designed to analyze Illumina microarray data which includes data input, quality control, variance stabilization, normalization and gene annotation. The function "lumiExpresso", which uses a variance-stabilizing transformation (VST) algorithm, was used to take advantage of the technical replicates available on every Illumina microarray.

Construction of a weighted gene co-expression network and functional categorization of genes

Network methods have been applied to identify and characterize various biological interactions[13-15]. Weighted gene co-expression network analysis (WGCNA) was performed on variably expressed genes in this study as previously described [8, 16, 17]. Briefly the data were filtered to minimize noise due to biologically irrelevant genes. First, the array detection values were used to select the 9,623 genes with statistical evidence of expression in at least 50% of the strains. The 3600 most varying genes were then selected for network construction. A correlation matrix was obtained by calculating the Pearson correlations between all variable probe sets across all strains. In weighted networks, soft thresholding of the Pearson correlation matrix is used for determining the connection strengths between two genes. A soft power adjacency function $a_{ij} = |cor(x_i, x_j)|^\beta$, where a_{ij} represents the resulting adjacency, was then used to construct an adjacency matrix. We chose a power of $\beta = 6$ based on the scale-free topology criterion proposed in Zhang and Horvath (2005). This power was chosen such the resulting network exhibited approximate scale-free topology and a high mean number of connections[12,17-19]. For module identification, topological overlap (TO), a biologically meaningful measure of node similarity, was calculated. A pair of genes was said to have high topological overlap if they were both strongly connected to the same group of genes. Next, the probe sets were hierarchically

clustered using dissimilarity, 1-TO, as the distance measure, and modules were determined using a dynamic tree-cutting algorithm (http://www.genetics.ucla.edu/labs/horvath/binzhang/ DynamicTreeCut). Modules were defined as sets of genes with high ''topological overlap'' [8,21]. The algorithm for dynamic tree cut was based on an adaptive process of cluster decomposition and combination, and the process was iterated until the number of clusters became stable. As a result, modules corresponded to branches of the dendogram (Fig. 1A). The whole network connectivity ($k$all) for each gene was determined by summing the connectivities of that gene with each of the other genes in the network and intramodular connectivity ($k$in) for each gene was determined by summing the connectivities of that gene with all other genes in that module. Genes with the highest kin were defined as module hub genes. Networks were graphically depicted using WebQTL (www.genenetwork.org). Group analyses were performed on all generated gene lists using the Database for Annotation, Visualization and Integrated Discovery (DAVID) online analysis program (the sixth version http://david.niaid.nih.gov/david/ ease.htm). Genes from each module in the co-expression network were submitted separately to the DAVID database, which clusters genes according to a series of common gene ontology and pathway categories. The proportion of module genes in each category was then compared with the proportion in the whole genome to compute enrichment scores and corresponding P values[22,23].

Integrating network analysis, bone geometry and BMD

Adult mice from the same strains that were used for microarray were sacrificed at 16 weeks of age. Femurs from 5 mice of each strain were dissected free of soft tissue, and length, anterio-posterior diameter, and medial-lateral diameter at the mid-shaft were measured to the nearest 0.01 mm with digital calipers[24]. BMD scans were performed using a Lunar PIXImus II Densitometer (GE Healthcare, Piscataway, NJ). The averages of trait measurements from 5 mice

from each strain were used. To examine the overall correlations of module genes with each of the phenotypes, we made use of gene significance (GS) measures. The GS was calculated as the absolute value of the correlation between each trait and gene expression values. For example, GS femur length (i) = |cor(x (i), femur length)| where x (i) was the gene expression profile of the ith gene. We could assume that the higher the absolute value of GS (i), the more biologically significant was the i-th gene. Module significance (MS) was determined as the average GS measured for all genes in a given module. Therefore, MS provided a measurement for overall correlation between the trait and the module. To assign a p-value to the module significance measure, we used the Fisher transformation as implemented in the R function ''cor.test''.

Genotype and module QTL analysis

Genotype data for the BXD and BXH strains were obtained from Gene Network (www.genenetwork.org)[25]. The BXD genotype file contained genotypes for 3795 markers (mean spacing of 0.6 centimorgan or 0.7 megabase [Mb]). This genotype file included all markers, both SNPs and microsatellites, with unique strain distribution patterns (SDPs). Further information can be found at www.genenetwork.org/dbdoc/BXDGeno.html and www.genenetwork.org/ dbdoc/BXHGeno.html. QTL linkage mapping was carried out in a combined dataset from BXD and BXH strains with the QTL Reaper software package (qtlreaper.sourceforge.net). A genotype file that was a combination of all the BXD and BXH strains was created. For all alleles that were D or H, we renamed it to N. Essentially, we treated this cross as "BXN" so the BXHs and BXDs were all the same cross. One thousand permutations of the strain labels were performed to estimate the genome-wide *P* values[26]. QTL graphics were generated using the multiple mapping tool from GeneNetwork. Interval mapping was used to map genetic loci associated with mRNA abundance [27,28]. A thousand permutations were used to

calculate the Likelihood Ratio Statistic (LRS) thresholds to assess genome-wide significance in linkage analysis. Logarithms of odds (LOD) values were obtained by dividing the LRS by 4.6. To identify module quantitative trait loci (module QTL), expression data from the 41 genes in salmon module were mapped using the QTL Cluster function. Module QTL were defined as loci with a significant enrichment for expression quantitative trait locus (eQTL) of the genes within a predetermined gene module. Empirical p-values and hypergeometric p-values were used to determine whether the proportion of module genes that map to the module QTL was significantly higher than that of the 3,600 network genes. Empirical p-values were calculated by using 10,000 permutations of 41 randomly sampled genes from the whole genome.

Cells and culture conditions

Mouse ATDC5 pre-chondrocyte cells[29] were maintained in medium consisting of a 1:1 mixture of Dulbecco's modified Eagle's medium and Ham's F-12 medium (DMEM/Ham's F12) (Gibco) containing 5% Fetal Bovine Serum (FBS) (Atlanta Biologicals), 10 mg/ml human transferrin (Roche) and 3×10^{-8} M sodium selenite (Sigma) (the maintenance medium) at 37° C in a humidified atmosphere consisting of 5% CO2. These cells were passaged every 4 days[30-32].Chondrogenesis and cartilage nodule formation were induced only in a postconfluent phase when cells were cultured in the maintenance medium supplemented further with 10 μg/ml bovine insulin (BD Biosciences) (the differentiation medium). ATDC5 cells were plated in 6-multiwell plastic plates at an initial cell density of 1×10^{4} cells/well. Four days after plating (day 0) chondrogenesis was induced by growth in differentiation medium. Cells were grown for another 28 days, with the medium being replaced every other day. The synthesis of ECM was used to evaluate chondrocyte differentiation. Sulfated glycosaminoglycans were visualized by staining with alcian blue. Cells were washed twice with PBS, fixed with methanol at -20°C for 2 min,

stained with 0.1% alcian blue 8GX (Sigma) in 0.1N HCl overnight, and rinsed repeatedly with distilled water. The preparations were evaluated by phase-contrast microscopy.

RNA isolation and real-time RT-PCR

Total RNA was extracted from ATDC5 cells at day 0, 7, 14, 21 and 28 using RNeasy Mini Kit (Qiagen) and treated with DNase I. First-strand cDNA was synthesized from total RNA (500 ng) using Superscript II enzyme (Invitrogen) and random hexamer primers. A total of 50 ng of cDNA was used as a template for real-time PCR. Sequences for primers were chosen using the Primer 3 program (http://fokker.wi.mit.edu/ primer3/input.html). The sequences of primers are presented in Supplementary Table 1. Primers were used at 5 µM with 10 µl of SYBR Green Master Mix (Qiagen) in a final volume of 20 µl. SYBR Green PCR amplification and real-time fluorescence detection were performed using the iCycler instrument (Bio-Rad) PCR and cycling conditions were as follows: 95°C for 15 minutes (denaturation), followed by 45 cycles at 94°C for 15 s, 60°C for 20 s (annealing), and 72°C for 20 s (amplification).

siRNA construction and transfection

Small interfering RNAs (siRNAs) directed against target genes were designed using Ambion Cenix software, and analyzed by a BLAST search to ensure gene specificity. Two of siRNAs were used per one target gene. A siRNA with a nonsilencing oligonucleotide sequence; 5'-UUCUCCGAACGUGUCACGUTT-3' and 5'-ACGUGACA CGUUCGGAGAATT-3' showing no known homology to mammalian genes was used as a negative control (Qiagen). The ATDC5 cells were plated in 6-multiwell plastic plates at an initial cell density of $1x10^4$cells/well. One day later, cells were transfected with siRNA (260 nM) using the RNAiFect transfection reagent (Qiagen) according to manufacturer's instructions. The differentiation medium was replaced one day after each siRNA transfection. At day 0, 7, 14, 21 and 28 after transfection mRNA was

harvested, treated with DNase I, subjected to reverse transcription and real-time quantitative PCR.

In situ *hybridization*

Non-radioactive images of frozen sections of E14.5 embryos were extracted from the GenePaint.org database (http://www.genepaint.org)[33]. Images were exported to Photoshop, and the tonal range adjusted using the Levels command to extend across the full range.

Results

Mouse cartilage co-expression network construction and their relationship to underlying biological processes

The WGCNA network algorithim was applied to a subset of cartilage gene expression data from 27 male BXH and BXD RI strain mice (see Materials and Methods). The 3,600 most varying genes were used for network construction. Fourteen gene modules were identified (Fig. 1A). A full list of genes by module appears in Supplemental Table 1. Genes in each module shared expression patterns that were more similar to one another than to the expression patterns of genes in other modules. We designated each module by an arbitrary color in order to distinguish between modules. The number of genes included in the modules ranged from 38 (cyan) to 569 (turquoise), and their mean overall connectivity (kall) ranged from 6.14 (cyan) to 79.30 (turquoise).

We were able to biologically characterize the modules using DAVID, a program that measures over-representation of genes with specific GO, KEGG, and BioCarta terms relative to a reference list. We examined the functional significance of all modules by testing for enrichment, for molecular function and for biological pathway memebership for genes in each module. For molecular function, 11 out of 14 modules not including yellow, cyan and tan had significant

enrichment in protein binding function (p<0.05). The yellow and cyan modules were enriched for catalytic activity (p=4.6 x 10^{-8} and p=4.8 x 10^{-3} respectively) and the tan module was enriched for peroxidase activity (p=1.5 x 10^{-2}). Although the expression patterns in each module were different, many of the modules shared similar GO categorizations, suggesting that some modules may be functionally related. For the biological pathway analysis, we found that eight modules were significantly enriched. Overall, each module was enriched with distinct gene sets belonging to separate biological pathways (Table 1).

Identification of hub genes in cartilage co-expression network

By definition, "hub" genes are genes that have high intramodular connectivity (kin). Previous studies have shown that the most highly connected genes tend to have critical cellular functions[10,13-17]. We defined the kin for each gene based on the sum of adjacencies with all other module genes[10,16,34,35] (see Materials and Methods) (Supplementary Table 2). Here, we selected the three genes with the highest kin in each module as hub genes (Table 4). Many modules contained hub genes important in chondrocyte differentiation such as SRY-box containing gene 5 (Sox5)[36-38], Lamin A (Lmna)[39], C1q and tumor necrosis factor related protein 3 (C1qtnf3)[40,41], Disabled homolog 2 (Dab2) and tumor protein, translationally-controlled 1 (Tpt1).

Salmon module genes have high GS with BMD and femoral geometry traits

The BXH and BXD mice showed a wide range of values for BMD and femoral geometry traits (Supplemental Table 2). Correlation analyses performed from all bone traits showed that only femur length and anterio-posterior diameter were significantly correlated (p<0.0001) (Pearson's r=0.53). To determine the physiological relevance of each module, we measured overall correlations between the traits and the module, using the module significance scores (MS) (see

Materials and Methods). A module with a high MS value for a particular trait is on average composed of genes highly correlated with that trait. Among all modules, the most significant MS scores were observed for the salmon module with all traits, including femur length (MS= 0.42), anterio-posterior diameter (MS=0.42), medial-lateral diameter (MS=0.58) and BMD (MS=0.48) (p<0.05), except anterio-posterior diameter, with which the purple module had the highest MS score (MS=0.52) (Fig. 1B). The first principal component of the modules (module eigengene) also used to measure the correlation between the traits and the modules and this yielded similar, although weaker correlation (data not shown). Since the salmon module was the most significantly related to bone traits, our follow-up analysis focused on this module. A visualization of the salmon module subnetwork is presented in Fig. 2. General information for all salmon module genes is presented in Supplementary Table 2.

Genetic control of module gene expression

We performed QTL mapping to identify genomic regions influencing the expression of modules [42-44]. We used the web-based eQTL mapping tool, WebQTL, to visualize putative regulatory loci controlling each transcript. Genomic regions with a significant enrichment for salmon module gene eQTLs were defined as module QTL. The LRS values for the forty one transcripts from the salmon module were plotted against the strength of LRS values. A QTL clustering of the LRS values for the 41 salmon module genes shows patterns of genetic control of salmon module gene expression (Fig. 3A). Chromosome 3, 5, 9, 11 and 17 contain clusters that consist of a set of transcripts with significant QTL (Fig. 3B). The details of the peak markers for each chromosome are given in Table 2. These loci are likely to contain genes which regulate the chondrocyte subnetwork related to bone development. Their identification will require fine mapping using congenic strain or other strategies.

Identification of ATDC5 cell culture stages using chondrogenic markers

ATDC5 cells that represent a chondroprogenitor clone were maintained as a growing population of undifferentiated cells and were then induced through an insulin-dependent pathway to differentiate along a pathway corresponding to cartilage development *in vivo* [32]. As shown in a representative experiment (Fig. 4A), the cellular condensations first appeared 5 days after induction, and the formation, growth and maturation of cartilaginous nodules occured thereafter on days 11–31. The timely pattern of expression of cartilage marker genes characteristic of early and late stages, the collagen type II (*Col II*), collagen type X (*Col X*) and *Runx2*, indicated an orderly progression of chondrogenic differentiation (Fig. 4B). *Col II*, which is known to be expressed in proliferating cells, was most highly induced on day 14, and gradually declined from day 21 to 28, while *Col X* which is known to be expressed in hypertrophic cells, was highly induced from day 21 to 28. *Runx2* is known to be expressed in precartilagenous condensations, downregulated in proliferating chondrocytes, and then strongly upregulated in prehypertrophic and hypertrophic chondrocytes. Thus, cells at day 7 represented precartilagenous chondrocytes, whereas cells at day 14 are predominantly proliferating chondrocytes. Cells at days 21 and 28 were enriched for pre-hypertrophic and hypertrophic chondrocytes, respectively. These initial investigations validated our experimental system with regard to the previously established parameters[32].

Comparative profile of 28 candidate gene expressions during differentiation of ATDC5

From the results of our network analysis we selected 28 genes for further investigation. These candidates included the 15 most connected genes in the salmon module and one most connected genes from the other 13 modules. The expression patterns of all 28 genes were monitored during the differentiation of ATDC5 cells using quantitative real time RT-PCR. The aim was to exclude

genes that did not exhibit a change in expression during differentiation in order to identify those genes whose function can be limited to cell proliferation and differentiation. Twelve out of 28 genes demonstrated at least a three-fold difference in expression during differentiation. Four distinct expression patterns were found among the genes assayed. The first expression pattern was observed for basic helix-loop-helix domain containing, class B9 (*Bhlhb9*) and heparin-binding EGF-like growth factor (*Ovca*) in which the expression levels increased from day one with a peak on day 14, and decreased on days 21-28 (Fig. 5A). This expression pattern corresponds to *Col II* which is expressed in the proliferating growth plate chondrocyte, but not in post mitotic hypertrophic chondrocytes. The second pattern, in which expression progressively increased from day one to 28, was observed for CUG triplet repeat, RNA binding protein 1 (*Cugbp1*) and cyclin-dependent kinase inhibitor 1A (*Cdkn1a*) (Fig. 5B). This expression pattern was seen for *Col X*, a marker for hypertrophic chondrocyte. The third expression pattern was observed for macrophage migration inhibitory factor (*Mif*) and yippee-like 5 (*Ypel5*), in which expression was induced on day 7, gradually reduced on days 14-21 and increased again on day 28 (Fig. 5C). This expression pattern was similar to the expression pattern of *Runx2*, a gene characteristic of early and late stage chondrocytes. The fourth expression pattern, characterized by decreasing expression from the beginning to the end of culture period, was found in nuclease sensitive element binding protein 1 (Nsep1), heat shock 90kDa protein 1, beta (Hspcb), signal peptidase complex subunit 3 homolog (Spcs3), ArfGAP with FG repeats 2 (Hrbl), heat shock protein 1 (Hspd1) and peptidylprolyl isomerase A (Ppia). The expression of Nsep1, Hspcb and Spcs3 gradually decreased from day one to 28 (Fig. 5D) while Hrbl, Hspd1 and Ppia decreased from days 1-21 with a small increase at day 28 (Fig. 5E). The expression patterns of these genes did not follow the patterns of our marker genes, Col II, Col X, and Runx2. However, many of

them have functions associated with cellular stress, degradation, apoptosis, etc. Nsep1 is a member of the EFIA/NSEP1/YB-1 family of DNA-binding proteins. Gene targeting of Nsep1 induces an early lethal phenotype in embryos, due to either haploinsufficiency of Nsep1 or formation of a dominant negative form of the protein[45]. Hence, its specific role in chondrogenesis is unknown. Spcs3 has not been studied in mammalian cells, but is essential for protein secretion. Several studies have reported that Hspd1, a chaperone protein of molecular weight of 60 kDa, is involved in carcinogenesis and apoptosis. It has been reported to be a ligand of toll-like receptor (TLR)-4 and induce apoptosis via the TLRs[46,47]. It is also expressed in chondrocytes in response to endoplasmic reticulum (ER) stress[48]. Ppia, also as known as Cyclophilin A (CyPA), is a 20-kDa chaperone protein secreted from vascular smooth muscle cells (VSMCs) in response to reactive oxygen species[49]. It is also produced by chondrocytes in response to activation of stress pathways[50].

We also used previously published microarray data to find the most highly variable gene expression patterns in the in vitro ATDC5 cell model [51]. The 200 most variably expressed genes are shown in Supplementary Table 3 and Supplementary Figures 1 and 2. We did not detect our candidate genes among those identified in these previous data, but it should be noted that major cartilage-specific genes such as Col2a1, Col10a1, aggrecan, Sox9, and Runx2 are also not represented. Hence our systems approach identifies a unique set of genes.

The role of candidate cartilage-related genes in chondrocyte differentiation

To test whether any of the candidate genes we identified are involved in chondrocyte differentiation, the genes which exhibited altered expression throughout the time course of ATDC5 differentiation were knocked down with small interfering RNA (siRNA). After inhibiting expression, the expression of markers was monitored to determine if reduced

expression of these candidate genes affected chondrocyte differentiation. ATDC5 cells treated with nonspecific siRNA had no effect on gene expression during chondrocyte differentiation. In contrast, cells treated with candidate gene-specific siRNAs exhibited reduced expression of the target gene in both siRNAs on average of 60-70%. This inhibition was specific, as *Gapdh* RNA expression was not affected by candidate gene-specific siRNA. Of the 12 candidates, five were found to reduce the expression of key chondrocyte differentiation genes. *Col X* mRNA levels were significantly reduced in week 3 and week 4 when the cells were transfected with *Bhlhb9*, *Cdkn1a*, *Spcs3* and *Cugbp1* siRNA (Fig. 6A). *Runx2* mRNA levels were significantly reduced from week 1 to week 4 when the cells were transfected with *Spcs3* siRNA and were significantly reduced in week 3 and week 4 when the cells were transfected with *Bhlhb9* and *Cdkn1a* siRNA (Fig. 6B). In contrast, the level of *Col II* was not significantly affected by these genes. *Col II* expression was significantly reduced from week 2 to week 4 when the cells were transfected with *Hspd1* siRNA (Fig. 6C).

Expression of candidate cartilage-related genes during chondrogenesis.

We examined the patterns of expression of those genes shown to impact chondrogenesis *in vitro* in order to determine whether or not they may have a related function *in vivo* (Fig. 7). As discussed above, expression of *bhlhb9* folows the same pattern as that of *Col2a1* during ATDC5 cell differentiation, and siRNA against *bhlhb9* inhibits differentiation. Consistent with an essential function in chondrogenesis, *bhlhb9* is highly expressed in proliferating growth plate chondrocytes (Fig. 7A). The expression of *Cdkn1a* and *Cugbp1* increases throughout differentiation, and siRNA against these genes blocks terminal maturation, as monitored by *Col10a1* and *Runx2* expression. *In situ* hybridization of E14.5 embryos reveals that *Cdkn1* is expressed prominently in perichondrium and in osteoblasts, but not in proliferating chondrocytes

(Fig.7B). Within the fully formed growth plate, *Cdkn1* expression has been shown to be restricted to hypertrophic chondrocytes[52, 53]. This pattern of expression is consitent with the block in terminal maturation seen when *Cdkn1a* is inhibited by siRNA. *Cugbp1*, which shows a related pattern of expression during ATDC5 differentiation, shows a similar pattern of expression *in vivo* (Fig. 7C). *Spcs3* levels decline throughout ATDC5 differentiation. Consistent with this timecourse, siRNA against *Spcs3* leads to impaired ATDC5 function in progenitor cells prior to the onset of chondrocyte differentiation, as monitored by the onset of its effects at week1[54-57]. Expression of *Spcs3* is seen at highest levels in perichondrium, where progenitor cells reside. Thus, the timecourses of expression of genes identified in our microarray that show an effect in ATDC5 cells correlate with patterns of expression *in vivo*.

Discussion

We utilized a systems-based approach to identify candidate pathways and genes involved in chondrocyte differentiation. We then tested a subset of these using analyses in a tissue culture model of differentiation. Five genes were shown to significantly influence chondrocyte differentiation as judged by siRNA knockdown, two of which have been implicated in bone biology in previous studies.

Our strategy was to utilize a systems genetics experimental design to organize global expression patterns into biologic subnetworks relevant to chondrocyte and bone biology. Thus, we isolated chondrocytes from two sets of recombinant inbred strains known to segregate for various bone phenotypes, including size and density, for global expression quantification. The genes were then subjected to weighted gene co-expression network (WGCNA) analysis to identify functional modules. These were then related to bone traits among the recombinant inbred strains using

correlation. One module (salmon) was highly correlated with several bone traits, in some cases (bone diameter) explaining over 25% of the trait variance.

There are several lines of evidence suggesting that our cartilage co-expression networks are biologically significant. First, WGCNA sorts variable genes and delineates genes into modules enriched for specific molecular functions and biological pathways. With respect to molecular function, we found that many of the modules share similar GO categorizations. Eleven of 14 modules showed significant enrichment in protein binding function, even though the expression patterns in each module were different. This suggests that some modules may be functionally related. For biological pathways, each module was enriched with distinct gene sets belonging to separate biological pathways. It is noteworthy that many of the resulting enrichment pathways, such as cyclin and cell cycle regulation as well as the Wnt signaling pathway, are crucial for chondrocyte development. Second, many modules contained hub genes important in chondrocyte differentiation. For example, the blue module had a hub gene that is important for cartilage development (*Sox5*). *Sox5* is a critical effector of chondroblast differentiation[36]. It is activated in prechondrocytes and highly expressed in chondroblasts in all developing cartilage elements of the mouse embryo[37]. $Sox5^{-/-}$ mutants die *in utero* with rudimentary and poorly developed cartilage anlagen, while *Sox5* heterozygous knockout mice are born with limited skeletal abnormalities[38]. One of the hub genes in the green yellow module (*Lmna*) has an effect on bone strength. Heterozygous knockout of *Lmna* increases the frequency of fracture in mice on certain genetic backgrounds[39]. *C1qtnf3*, a hub gene in the yellow module has been reported to be predominantly expressed in cartilage[40]. Because its product is a secretory protein produced by cartilage tissue, this gene was later named "cartducin". In situ hybridization analysis showed that cartducin transcripts were restricted to the proliferating chondrocytes in the growth plate

cartilage. Chondrogenic differentiation stimulated by combined treatment with bone morphogenetic protein-2 and insulin induced cartducin expression along with type II and IX collagen expression in chondrogenic progenitor N1511 cells[40, 41]. The hub genes in the magenta (*Dab2*) and turquoise (*Tpt1*) modules both have effects on cell growth. *Dab2* is involved in mesenchymal cell differentiation, while *Tpt1* is involved in cell apoptosis. Overall, we found many hub genes that have known biological functions related to skeletal development. Therefore, other unknown hub genes within the network may have important roles in skeletogenesis. The advantage of our analysis is that these results would never have been uncovered by using traditional microarray analysis methods. Moreover, hub genes from many modules are key players in the biological processes suggested by DAVID analysis. Finally, most of the modules are conserved when BXH and BXD are analyzed separately (data not shown). These data support the notion that gene co-expression networks can be used to organize gene expression data into a functionally relevant framework to explore skeletal development from a systems perspective.

Because size, shape and mineral content are all critical components of bone strength, the phenotype under study becomes critical for attempting to understand the biomechanical mechanisms that ultimately define bone morphology and strength. Overall correlation between relevant bone traits and module expression showed that the salmon module is directly related to all bone traits with significant MS scores. This relationship was also evident when examining the list of genes within this module. Three out of the 10 most connected genes in the salmon module have known functions related to bone or cartilage: cyclin-dependent kinase inhibitor 1A (*Cdkn1a*), heat shock 60kDa protein 1 (*Hspd1*) and Ngfi-A binding protein 1 (*Nab1*). *Cdkn1a* was found to affect the proliferation of chondrocytes. It has been shown to be expressed in postmitotic, hypertrophic chondrocytes but not in proliferating chondrocytes[58]. *Cdkn1a* induces

the expression of fibroblast growth factor 2 (*Fgf2*), a negative regulator of chondrocytic growth, and the inhibitory effect of *Fgf2* on chondrocytic proliferation was partially reduced in *Cdkn1a*-null limbs[59]. Increased expression of *Cdkn1a* in cartilage has been demonstrated in thanatophoric dysplasia human embryos[60]. The expression of *Hspd1* was found to be increased in secondary hyperparathyroidism in which the secretion of parathyroid hormone (*PTHrP*), the hormone that regulates the rate of growth plate chondrocyte development, is excessive[61]. It is thus reasonable to expect that the highly related hormone, *PTHrP*, exerts a similar inductive effect on *Hspd1* expression. *Nab1* has a function in bone remodeling. *Nab1* is an endogenous repressor of C-8 sterol isomerase (*Erg2*), a gene important for osteoclast survival. Inhibition of *Egr2* increases osteoclast apoptosis. Wild-type *Egr2* or *Egr2* point mutants unable to bind *Nab1/2* suppress basal osteoclast apoptosis and rescue osteoclasts from apoptosis induced by *Egr2* inhibition[62].

The yellow module is enriched with genes involved in *Wnt* signaling, which is one of the most important pathways for bone development. For example, genes in this module include connective tissue growth factor (*Ctgf*), nephroblastoma overexpressed (*Nov* or *Ccn3*), and catenin (cadherin associated protein) beta 1 (*Ctnnb1*). *Ctgf*, belongs to a larger *CCN* gene family, and acts downstream of Wnt in cartilage[63]. It is required for endochondral ossification[64]. *Nov*, also a member of *Ccn* family, is also expressed in growth plate chondrocytes, and is essential for chondrogenesis[65]. *Ctnnb1* or beta-catenin is as an integral component in the *Wnt* signaling pathway. Through several cytoplasmic relay components, *Wnt* signal is transduced to beta-catenin, which enters the nucleus and forms a complex with TCF to activate transcription of *Wnt* target genes. Overall, yellow module seems to have important function in chondrogenesis and bone development. Therefore, this module would be an interesting module to study in the future.

However, the yellow module had low correlation with bone traits. This might be because the *Wnt* pathway has such significant functions at early stages of skeleetal development. We used cartilage from neonatal (P2) mice (day-2 mice) for microarray and enrichment analysis, and the correlation with the phenotypes from later stage (16-week) mice may be low. Clearly, additional phenotypic correlation studies are warranted in the future.

We also used the Modulated Modularity Clustering (MMC) method[66] to analyze our 3,600 gene expression data to see how the modules fractionate with other method (supplementary figure 3 and 4). Some interesting overlap groups of genes are found. Data from MMC Pearson correlation shows that twenty nine out of 41 genes in our salmon module which has high correlation to all bone traits are clustered in one module in MMC. The list of these overlaping 28 genes is in supplementary table 4. Data from Spearman correlation shows more scattering group of genes. However, the genes in salmon module are still clustered in some modules as shown in supplementary table 4.

As far as the global transcriptional profiles are concerned, only a few investigations have focused on monitoring changes in expression levels of chondrocyte genes in a temporal fashion[67,68]. This study provides a complete determination of expression patterns that can make accurate assessments of candidate gene function during the chondrogenic differentiation pathway. Time-course expression profiles were described for 12 highly regulated genes, including *Bhlhb9*, *Ovca*, *Cugbp1*, *Cdkn1a*, *Mif1*, *Ypel5*, *Ybx1*, *Hspcb*, *Spcs3*, *Hrbl*, *Hspd1* and *Ppia*. Ten of these had not been associated with chondrogenesis previously. Our knock down experiments indicated that five of the 12 candidate genes regulate the expression of chondrocyte marker genes and all these five genes are in salmon module. Interestingly, when *Bhlhb9*, the gene which has an expression pattern corresponding to *Col II,* was knocked down, the expression levels of *Col X* and *Runx2*

were decreased. This suggests that *Bhlhb9* may play a role in chondrocyte differentiation. When *Cugbp1*, the gene which has an expression pattern similar to *Col X*, was knocked down, the expression level of *Col X* was decreased, raising the possibility that *Cugbp1* is also required for differentiation. The findings that *Cdkn1a* and *Hspd1*, previously known skeletogenesis-associated genes, regulated the expression of chondrogenesis marker genes also validates the approach and analysis. A limitation of this study is that it is a board study to define a network of genes involved in chondrocyte development with some validations and predictions. However, the detailed mechansims of actions of the genes identified remain to be determined. We have identified a requirement for Hspd1, Cdkn1a, Bhlhb9, Cugbp1 and Spcs3 in vitro. The in vivo functions of these genes will be investigated in the future studies. Overall, this study suggests that systems-level analyses can provide significant insights not easily reached through conventional approaches.

References

1. Hall BK, Miyake T 2000 All for one and one for all: condensations and the initiation of skeletal development. Bioessays 22: 138–147.

2. Kronenberg HM 2003 Developmental regulation of the growth plate. Nature: 423: 332–336.

3. Zelzer E, Olsen BR 2003 The genetic basis for skeletal diseases. Nature 423: 343-348.

4. Olsen BR, Reginato AM, Wang W 2000 Bone development. Annu. Rev. Cell Dev. Biol 16: 191–220.

5. Alman BA 2008 Skeletal dysplasias and the growth plate.Clin Genet 73: 24–30.

6. Ng JK, Tamura K, Buscher D, Izpisua-Belmonte JC 1999 Molecular and cellular basis of pattern formation during vertebrate limb development. Curr Top Dev Biol 41: 37–66.

7. Tickle C 2002 Molecular basis of vertebrate limb patterning. Am J Med Genet 112 (3): 250–255.

8. Zhang B, Horvath S 2005 A General Framework for Weighted Gene Co-Expression Network Analysis. Stat Appl Genet Mol Biol 4: Article17.

9. Dong J, Horvath S 2007 Understanding Network Concepts in Modules. BMC Systems Biology 1:24.

10. Ghazalpour A, Doss S, Zhang B, Wang S, Plaisier C, Castellanos R, Brozell A, Schadt EE, Drake TA, Lusis AJ, Horvath S 2006 Integrating genetic and network analysis to characterize genes related to mouse weight. PLoS Genet 2(8):e130.

11. Ihaka, R. and R. Gentleman 1996 R: a language for data analysis and graphics J. Comp. Graph. Stat 5: 299-314.

12. Du P, Kibbe WA, Lin SM 2008 lumi: a pipeline for processing Illumina microarray. Bioinformatics 1:24(13):1547-8.

13. Barabasi AL, Oltvai ZN 2004 Network biology: Understanding the cell's functional organization. Nat Rev Genet 5: 101–113.

14. Featherstone DE, Broadie K 2002 Wrestling with pleiotropy: Genomic and topological analysis of the yeast gene expression network. Bioessays 24: 267–274.

15. Basso K, Margolin AA, Stolovitzky G, Klein U, Dalla-Favera R 2005 Reverse engineering of regulatory networks in human B cells. Nat Genet 37: 382–390.

16. Horvath S, Zhang B, Carlson M, Lu KV, Zhu S, Felciano RM, Laurance MF, Zhao W, Qi S, Chen Z, Lee Y, Scheck AC, Liau LM, Wu H, Geschwind DH, Febbo PG, Kornblum HI, Cloughesy TF, Nelson SF, Mischel PS 2006 Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. Proc Natl Acad Sci U S A. 14;103(46):17402-7.

17. Oldham MC, Horvath S, Geschwind DH 2006 Conservation and evolution of gene coexpression networks in human and chimpanzee brains. Proc Natl Acad Sci U S A. 21;103(47):17973-8.

18. Vazquez A, Dobrin R, Sergi D, Eckmann JP, Oltvai ZN, Barabasi AL 2004 The topological relationship between the large-scale attributes and local interaction patterns of complex networks. Proc Antl Acad Sci U S A. 28;101(52): 17940-5.

19. Jeong H, Mason S, Barabasi A, Oltvai Z 2001 Lethality and centrality in protein networks. Nature 411:41– 42.

20. Wagner A, Fell DA 2001 The small world inside large metabolic networks. Proc Biol Sci 268: 1803–1810.

21. Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabasi AL 2002 Hierarchical organization of modularity in metabolic networks. Science 297: 1551–1555.

22. Sherman BT, Huang da W, Tan Q, Guo Y, Bour S, Liu D, Stephens R, Baseler MW, Lane HC, Lempicki RA 2007 DAVID Knowledgebase: a gene-centered database integrating heterogeneous gene annotation resources to facilitate high-throughput gene functional analysis. BMC Bioinformatics 2;8:426.

23. Huang da W, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, Stephens R, Baseler MW, Lane HC, Lempicki RA 2007 The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol 8(9):R183.

24. Donahue L, Beamer WG, Bogue MA, Churchill GA with Tsaih S-W, Sears AR, MacKenzie GW Models of skeletal geometry and bone strength. Mouse Phenome Database, the Jackson Laboratory. Available at http://phenome.jax.org/pub-cgi/phenome/mpdcgi?rtn=projects/docstatic&doc=Donahue2/Donahue2_Protocol. Accessed 2004.

25. Lopez Franco GE, O'Neil TK, Litscher SJ, Urban-Piette M, Blank RD 2004 Accuracy and precision of PIXImus densitometry for ex vivo mouse long bones: comparison of technique and software version. J Clin Densitom **7**(3)**:**326-33.

26. Churchill GA, Doerge RW 1994 Empirical threshold values for quantitative trait mapping. Genetics 138:963-971.

27. Doerge RW 2002 Mapping and analysis of quantitative trait loci in experimental populations. Nat.Rev.Genet 3:43-52.

28. Wu RL, Ma CX, Casella G 2007 Statistical Genetics of Quantitative Traits: Linkage, Maps, and QTL. Springer.

29. Atsumi T, Miwa Y, Kimata K, Ikawa Y 1990 A chondrogenic cell line derived from a differentiating culture of AT805 teratocarcinoma cells. Cell Differ Dev 30:109–116.

30. Akiyama H, Shigeno C, Hiraki Y, Shukunami C, Kohno H, Akagi M, Konishi J, Nakamura T 1997 Cloning of a mouse smoothened cDNA and expression patterns of hedgehog signalling molecules during chondrogenesis and cartilage differentiation in clonal mouse EC cells, ATDC5, Biochem. Biophys. Res. Commun 235: 142-147.

31. Shukunami C, Shigeno C, Atsumi T, Ishizeki K, Suzuki F, Hiraki Y 1996 Chondrogenic differentiation of clonal mouse embryonic cell line ATDC5 in vitro: differentiation-dependent gene expression of parathyroid hormone (PTH)/PTH-related peptide receptor. J. Cell Biol 133: 457-468.

32. Shukunami C, Ishizeki K, Atsumi T, Ohta Y, Suzuki F, Hiraki Y 1997 Cellular hypertrophy and calcification of embryonal carcinoma-derived chondrogenic cell line ATDC5 in vitro. J. Bone Miner. Res 12:1174-1188.

33. Visel A, Thaller C, and Eichele G 2004 GenePaint.org: an atlas of gene expression patterns in the mouse embryo. Nucleic Acids Res. 32: D552-D556.

34. Carlson MR, Zhang B, Fang Z, Mischel PS, Horvath S, Nelson SF 2006 Gene connectivity, function, and sequence conservation: Predictions from modular yeast co-expression networks. BMC Genomics 7: 40.

35. Carter SL, Brechbuhler CM, Griffin M, Bond AT 2004 Gene co-expression network topology provides a framework for molecular characterization of cellular state. Bioinformatics 20: 2242–2250.

36. Lefebvre V 2002 Toward understanding the functions of the two highly related Sox5 and Sox6 genes. J Bone Miner Metab 20: 121-130.

37. Lefebvre V, Li P, de Crombrugghe B. 1998. A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. EMBO J 17: 5718-5733.

38. Smits P, Dy P, Mitra S, Lefebvre V. 2004. Sox5 and Sox6 are needed to develop and maintain source, columnar, and hypertrophic chondrocytes in the cartilage growth plate. J Cell Biol 164: 747-758.

39. Fong LG, Ng JK, Meta M, Coté N, Yang SH, Stewart CL, Sullivan T, Burghardt A, Majumdar S, Reue K, Bergo MO, Young SG 2004 Heterozygosity for Lmna deficiency eliminates the progeria-like phenotypes in Zmpste24-deficient mice. Proc Natl Acad Sci U S A 28;101(52):18111-6.

40. Maeda T, Abe M, Kurisu K, Jikko A, Furukawa S 2002 Molecular cloning and characterization of a novel gene, CORS26, encoding a putative secretory protein and its possible involvement in skeletal development. J Biol Chem 2;276(5):3628-34.

41. Maeda T, Jikko A, Abe M, Yokohama-Tamaki T, Akiyama H, Furukawa S, Takigawa M, Wakisaka S 2006 Cartducin, a paralog of Acrp30/adiponectin, is induced during chondrogenic differentiation and promotes proliferation of chondrogenic precursors and chondrocytes. J Cell Physiol 206(2):537-44.

42. Doss S, Schadt EE, Drake TA, Lusis AJ 2005 Cis-acting expression quantitative trait loci in mice. Genome Res 15:681-691.

43. Peirce JL, Li H, Wang J, Manly KF, Hitzemann RJ, Belknap JK, Rosen GD, Goodwin S, Sutter TR, Williams RW, Lu L 2006 How replicable are mRNA expression QTL? Mamm Genome 17:643-656.

44. Rockman MV, Kruglyak L 2006 Genetics of global gene expression. Nat Rev Genet 7:862-872.

45. Fan L, Jones SN, Padden C, Shen Q, Newburger PE 2006 Nuclease sensitive element binding protein 1 gene disruption results in early embryonic lethality. J Cell Biochem 1; 99(1):140-5.

46. Hwang YJ, Lee SP, Kim SY, Choi YH, Kim MJ, Lee CH, Lee JY, Kim DY 2009 Expression of heat shock protein 60 kDa is upregulated in cervical cancer. Yonsei Med J. 30;50(3):399-406.

47. Kim SC, Stice JP, Chen L, Jung JS, Gupta S, Wang Y, Baumgarten G, Trial J, Knowlton AA 2009 Extracellular heat shock protein 60, cardiac myocytes, and apoptosis. Circ Res. 4;105(12):1186-95.

48. Cheng TC, Benton HP. 1994 The intracellular Ca(2+)-pump inhibitors thapsigargin and cyclopiazonic acid induce stress proteins in mammalian chondrocytes. Biochem J. 301:563-8.

49. Satoh K, Matoba T, Suzuki J, O'Dell MR, Nigro P, Cui Z, Mohan A, Pan S, Li L, Jin ZG, Yan C, Abe J, Berk BC 2008 Cyclophilin A mediates vascular remodeling by promoting inflammation and vascular smooth muscle cell proliferation. Circulation 2008 17;117(24):3088-98.

50. Catterall JB, Rowan AD, Sarsfield S, Saklatvala J, Wait R, Cawston TE. 2006 Development of a novel 2D proteomics approach for the identification of proteins

secreted by primary chondrocytes after stimulation by IL-1 and oncostatin M. Rheumatology 45:1101-9.

51. Microarray analyses of gene expression during chondrocyte differentiation identify novel regulators of hypertrophy. James CG, Appleton CT, Ulici V, Underhill TM, Beier F. Mol Biol Cell. 2005 Nov;16(11):5316-33. Epub 2005 Aug 31.

52. Zenmyo M, Komiya S, Hamada T, Hiraoka K, Suzuki R and Inoue A 2000 p21 and Parathyroid Hormone-Related Peptide in the Growth Plate. Calcif. Tiss. Int. 67: 378-81.

53. Xhang, R., Murakami, S., Coustry, F., Wang, Y. and de Crombrugghe, B. 2005 Constitutive activation of MKK6 in chondrocytes of transgenic mice inhibits proliferation and delays endochondral bone formation. PNAS 103: 365-370.

54. Enomoto H, Enomoto-Iwamoto, Iwamoto M, Nomura M, Himeno S, Kitamura M, Kishimoto Y, Komori T. 2000 J Biol Chem 275, 8695-8702.

55. Brown A. J., Alicknavitch M, D'Souza S, Daikoku T, Kirn-Safran C. B, Marchetti D, Carson D D, Farach-Carson M. C. 2008 Bone 43, 689-699.

56. Kawasaki Y, Kugimiya F, Chikuda H, Kamekura S, Ikeda T, Kawamura N, Saito T, Shinoda Y, Higashikawa A, Yano F, Ogasawara T, Ogata N, Hoshi K, Hofmann F, Woodgett J. R, Nakamura K, Chung U. I, Kawaguchi H. 2008 J Clin Invest 118, 2506-2515

57. Usui M, Xing L, Drissi H, Zuscik M, O'Keefe R, Chen D, Boyce B. F.2008 J Bone Miner Res 23, 314-325.

58. Aikawa T, Segre GV, Lee K 2001 Fibroblast growth factor inhibits chondrocytic growth through induction of p21 and subsequent inactivation of cyclin E-Cdk2. J Biol Chem 3;276(31):29347-52.

59. Parker SB, Eichele G, Zhang P, Rawls A, Sands AT, Bradley A, Olson EN, Harper JW, Elledge SJ 1995 p53-independent expression of p21Cip1 in muscle and other terminally differentiating cells. Science 17;267(5200):1024-7.

60. Su WC, Kitagawa M, Xue N, Xie B, Garofalo S, Cho J, Deng C, Horton WA, Fu XY 1997 Activation of Stat1 by mutant fibroblast growth-factor receptor in thanatophoric dysplasia type II dwarfism. Nature 86, 288–292.

61. Santamaría I, Alvarez-Hernández D, Jofré R, Polo JR, Menárguez J, Cannata-Andía JB 2005 Progression of secondary hyperparathyroidism involves deregulation of genes related to DNA and RNA stability. Kidney Int 67(6):2267-79.

62. Bradley EW, Ruan MM, Oursler MJ 2008 Novel pro-survival functions of the Kruppel-like transcription factor Egr2 in promotion of macrophage colony-stimulating factor-mediated osteoclast survival downstream of the MEK/ERK pathway.J Biol Chem 21;283(12):8055-64. Epub 2008 Jan 15.

63. Huang B-L, Brugger S, Lyons K.M. 2010 Stage-specific control of connective tissue growth factor (Ctgf/Ccn2) expression in chondrocytes by Sox9 and beta catenin. J. Biol Chem. Submitted

64. Ivkovic S, Yoon BS, Popoff SN, Safadi FF, Libuda DE, Stephenson RC, Daluiski A, Lyons KM 2003 Connective tissue growth factor coordinates chondrogenesis and angiogenesis during skeletal development. Development 130: 2779-2791

65. Heath E, Tahri D, Andermarcher E, Schofield P, Fleming S, Boulter CA. 2008 Abnormal skeletal and cardiac development, cardiomyopathy, muscle atrophy and cataracts in mice with a targeted disruption of the Nov (Ccn3) gene. BMC Dev Biol. 8:18.

66. Stone EA, Ayroles JF. 2009 Modulated modularity clustering as an exploratory tool for functional genomic inference. PLoS Genet. 5(5):e1000479.

67. Sekiya I, Vuoristo JT, Larson BL, Prockop DJ 2002 In vitro cartilage formation by human adult stem cells from bone marrow stroma defines the sequence of cellular and molecular events during chondrogenesis. Proc Natl Acad Sci USA 99: 4397–4402.

68. Clancy BM, Johnson JD, Lambert AJ, Rezvankhah S, Wong A, Resmini C, Feldman JL, Leppanen S, Pittman DD 2003 A gene expression profile for endochondral bone formation: oligonucleotide microarrays establish novel connections between known genes and BMP-2-induced bone formation in mouse quadriceps. Bone 33:46–63.

Fig 1.

Fig 2.

Fig 3.

Insulin +

Day | 7 | 14 | 21 | 28

Col II

Col X

Runx2

GAPDH
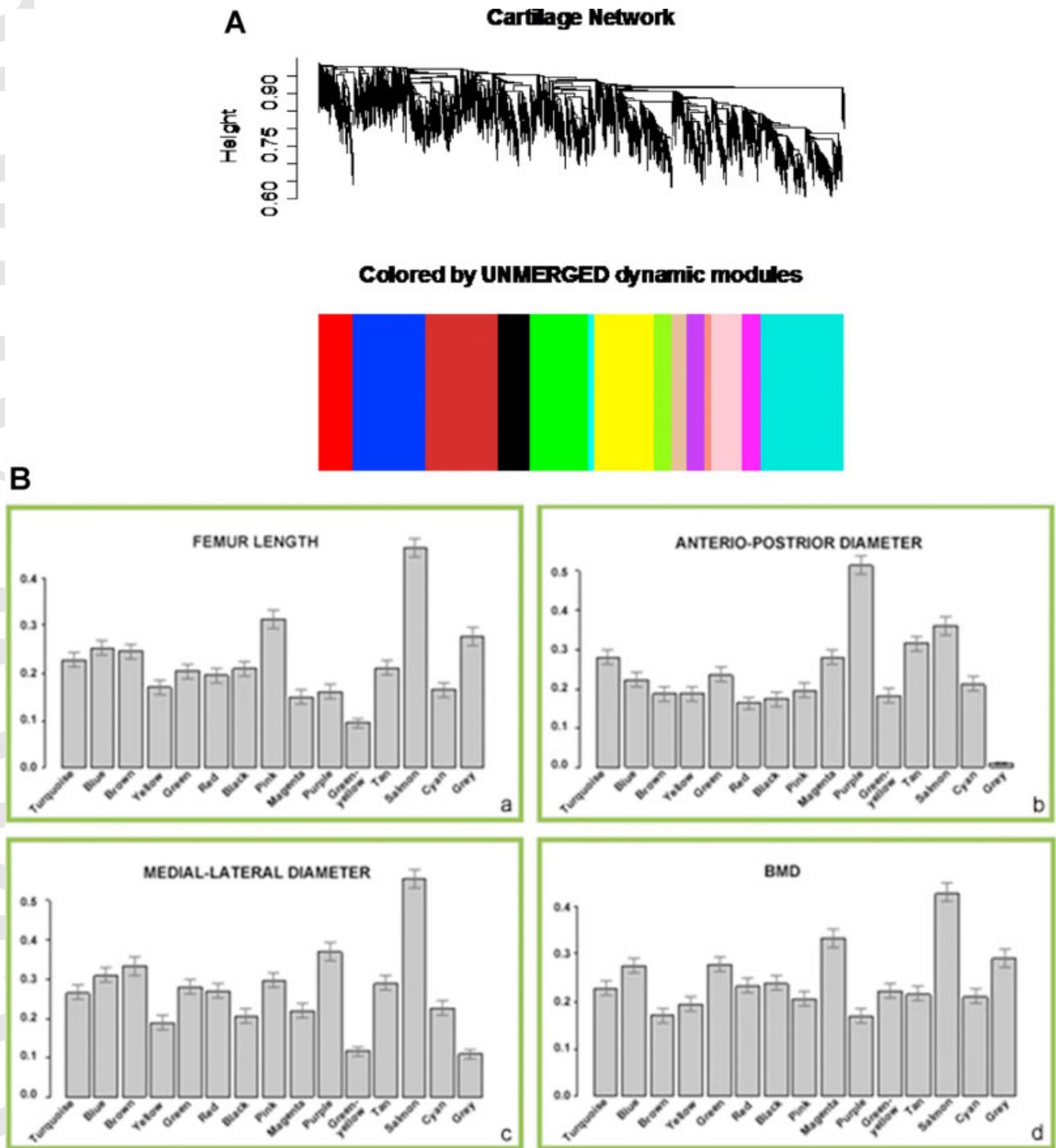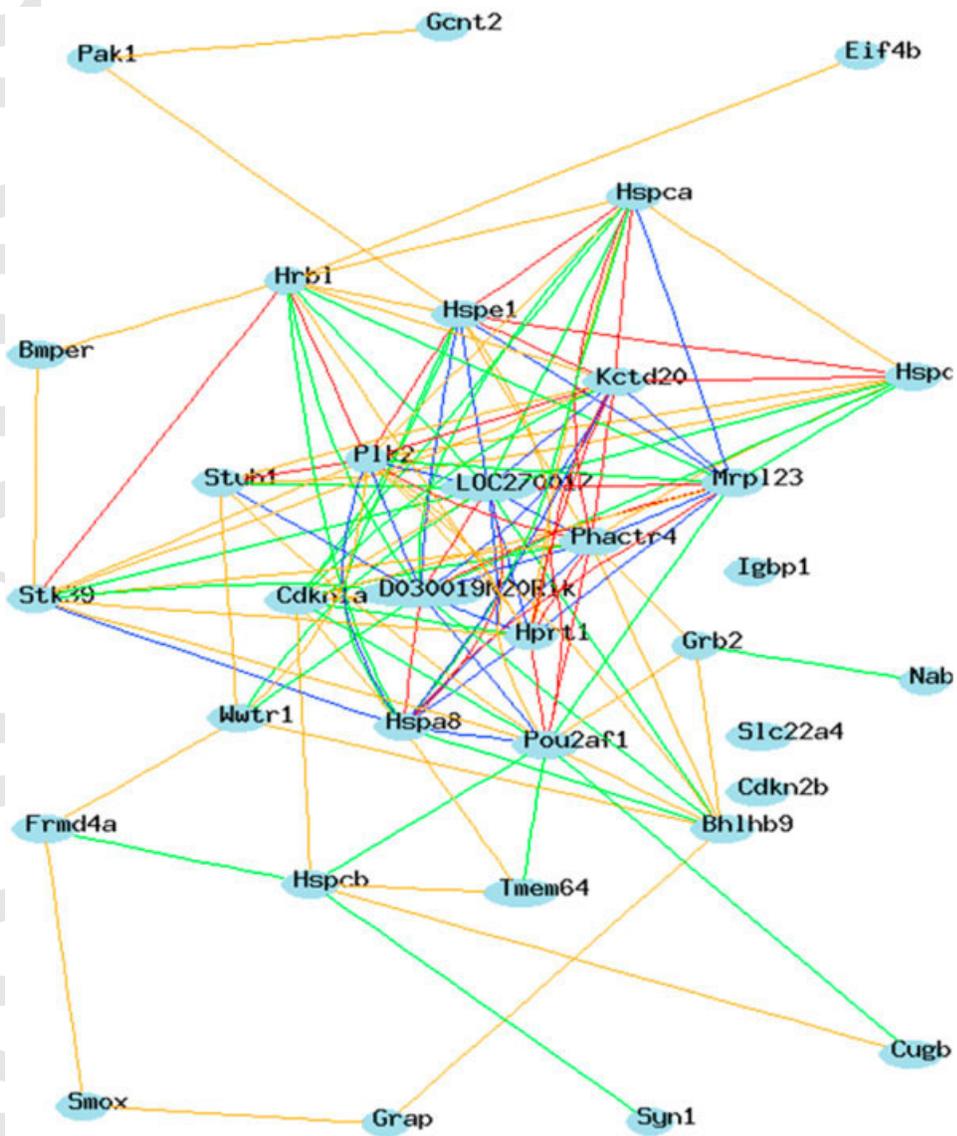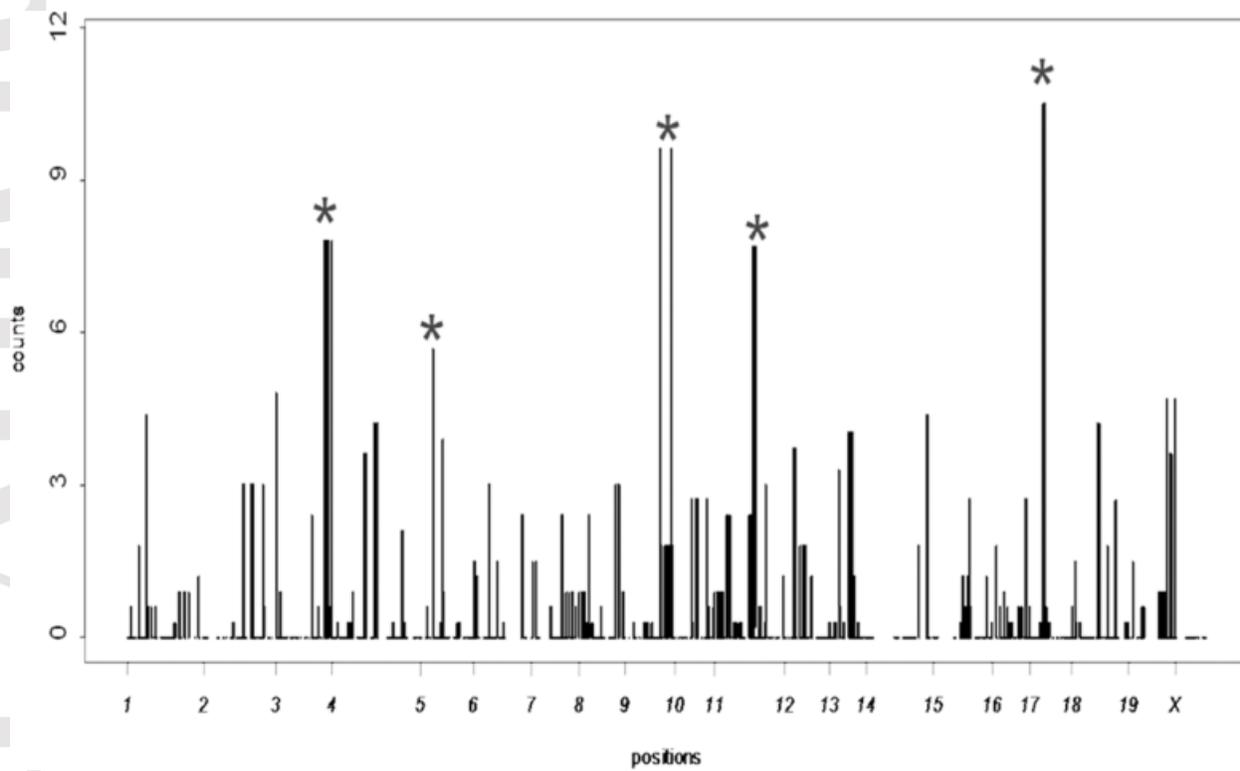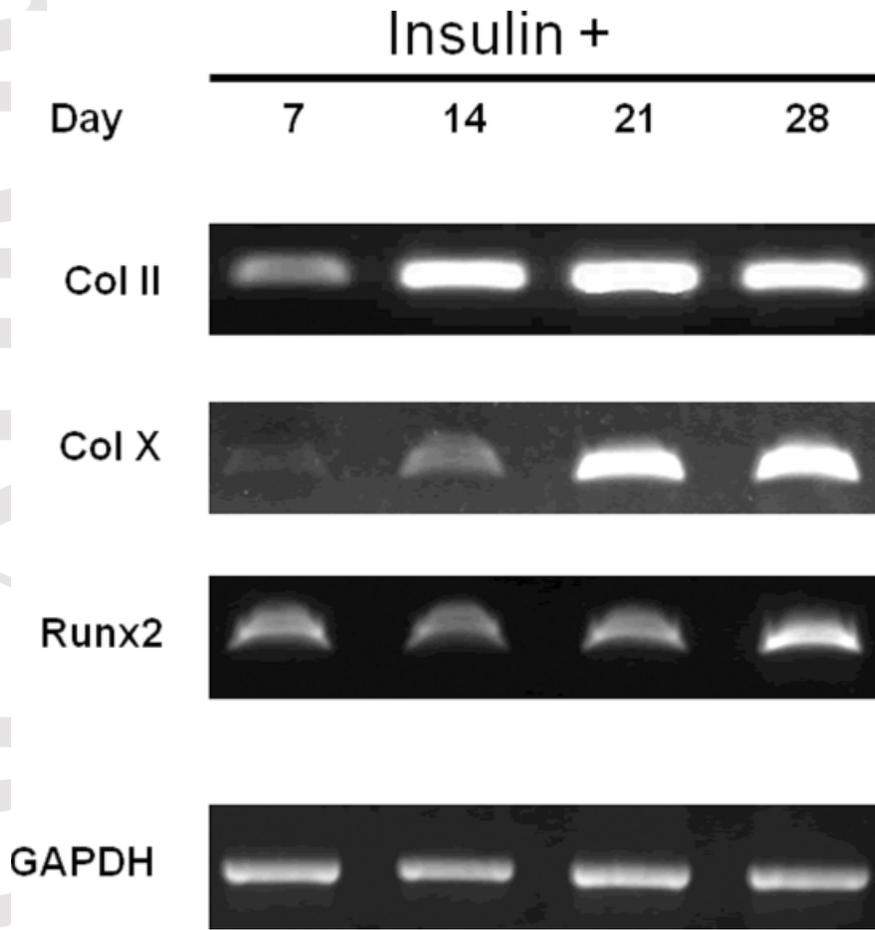
Fig 4.
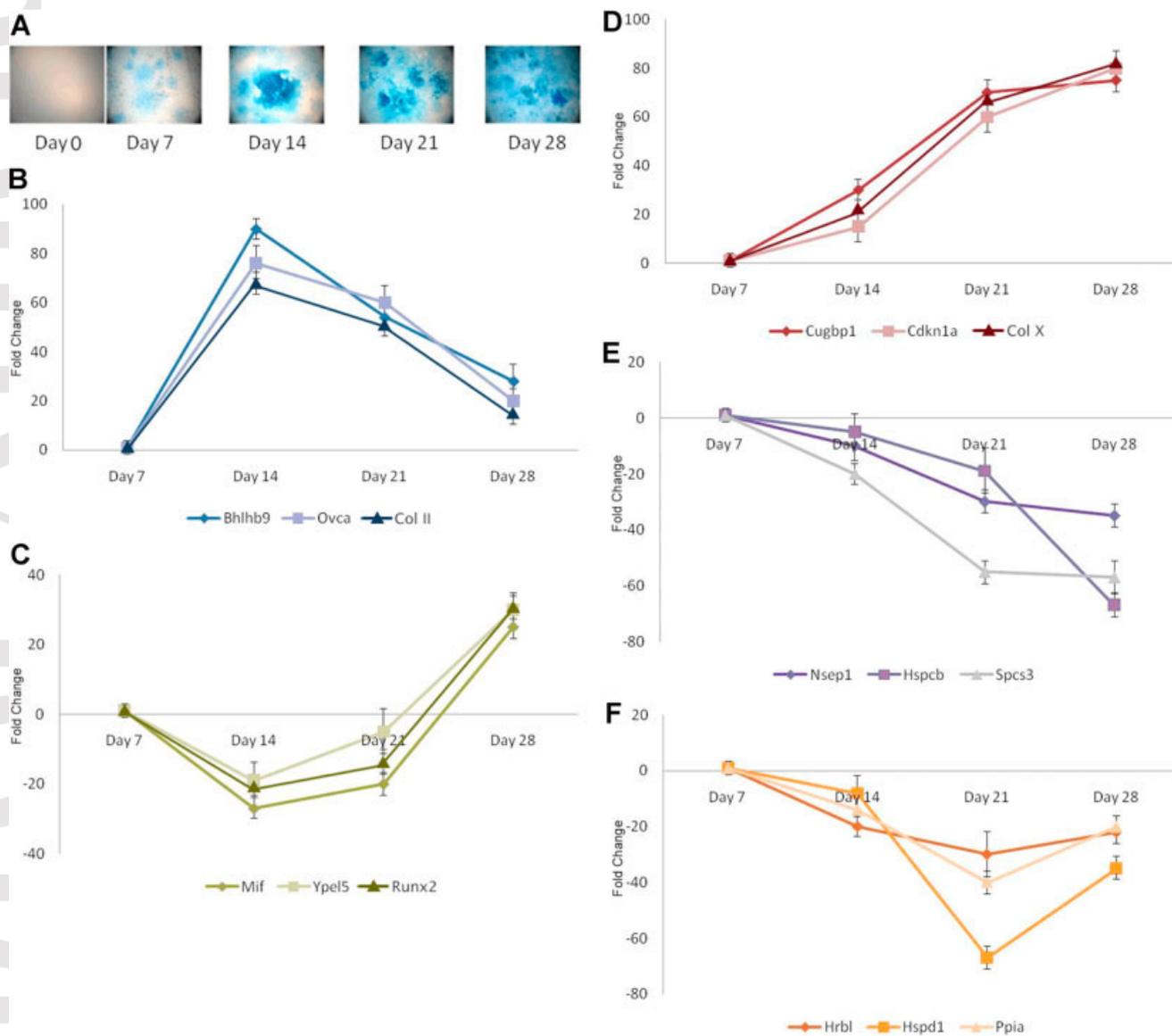
Fig 5.

A Col X
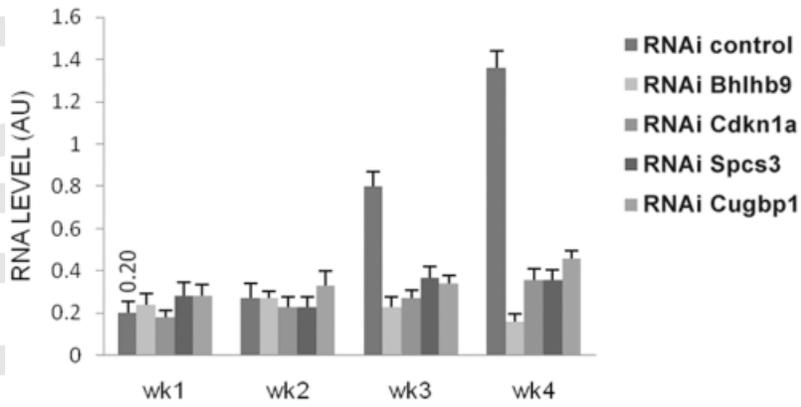
**B** Runx2
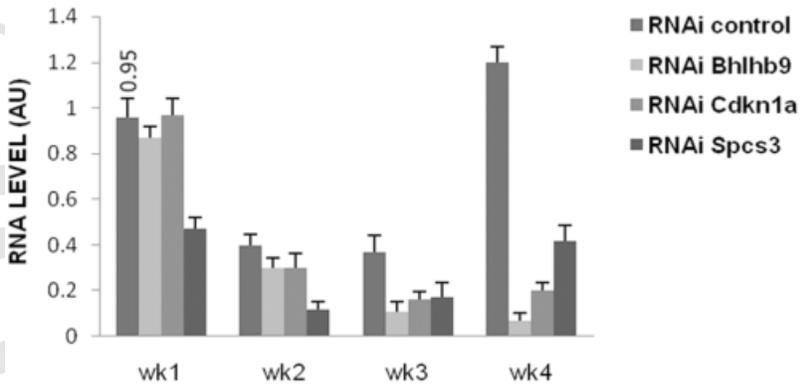
C Col II
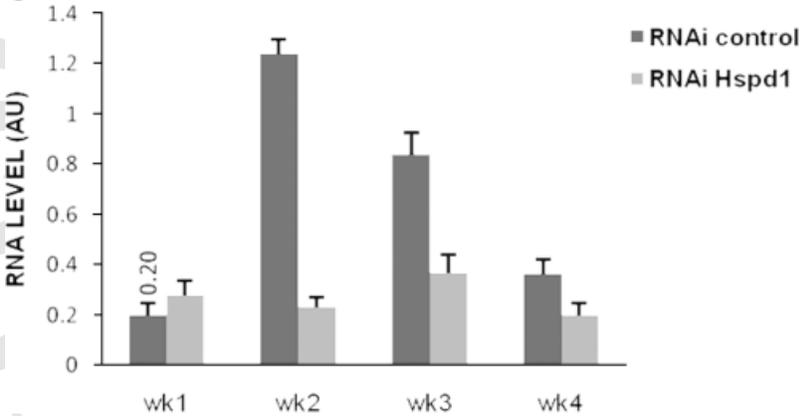
Fig 6.

Table 1. Analysis of enrichment of chondrocyte module for known pathways using DAVID.

Module	Biological pathway	Bonferroni corrected P-value
Purple	Glycolysis/gluconeogenesis	1×10^{-9}
Salmon	Cyclin and cell cycle regulation	4.2×10^{-4}
Magenta	MAPK signaling	4.3×10^{-4}
Green yellow	Proteasome	1.2×10^{-2}
Brown	Ribosome	1.3×10^{-2}
Blue	Adherens junction	1.3×10^{-2}
Turquoise	Focal adhesion	1.5×10^{-2}
Yellow	Wnt signaling	2.2×10^{-2}

Table 2. Salmon genes module QTL

Module QTL	SNP ID	SNP Physical Position	Number of transcripts	Mean LRS	Empirical p-value	Hypergeometric p-value
Chr 3	rs13477528	chr3:159,474,770-159,475,270	8	23.2	1.04E-02	2.04E-03
Chr 5	rs13478595	chr5:151,148,287-151,148,787	5	26.3	4.10E-03	1.10E-04
Chr 9	rs3705403	chr9:97,259,504-97,260,004	10	20.6	5.30E-03	6.80E-04
Chr 11	rs13481262	chr11:119,215,875-119,216,375	7	27.4	1.30E-04	7.30E-05
Chr 17	rs3706382	chr17:94,530,102-94,530,602	10	30.4	4.30E-05	2.80E-06