

RESEARCH

Open Access

# An integrated proteomics analysis of bone tissues in response to mechanical stimulation

Jiliang Li<sup>1</sup>, Fan Zhang<sup>2,3</sup>, Jake Y Chen<sup>2,3,4\*</sup>

From BIOCOMP 2010 - The 2010 International Conference on Bioinformatics and Computational Biology Las Vegas, NV, USA. 12-15 July 2010

## Abstract

Bone cells can sense physical forces and convert mechanical stimulation conditions into biochemical signals that lead to expression of mechanically sensitive genes and proteins. However, it is still poorly understood how genes and proteins in bone cells are orchestrated to respond to mechanical stimulations. In this research, we applied integrated proteomics, statistical, and network biology techniques to study proteome-level changes to bone tissue cells in response to two different conditions, *normal loading* and *fatigue loading*. We harvested ulna midshafts and isolated proteins from the control, loaded, and fatigue loaded Rats. Using a label-free liquid chromatography tandem mass spectrometry (LC-MS/MS) experimental proteomics technique, we derived a comprehensive list of 1,058 proteins that are differentially expressed among normal loading, fatigue loading, and controls. By carefully developing protein selection filters and statistical models, we were able to identify 42 proteins representing 21 Rat genes that were significantly associated with bone cells' response to quantitative changes between normal loading and fatigue loading conditions. We further applied network biology techniques by building a fatigue loading activated protein-protein interaction subnetwork involving 9 of the human-homolog counterpart of the 21 rat genes in a large connected network component. Our study shows that the combination of decreased anti-apoptotic factor, Raf1, and increased pro-apoptotic factor, PDCD8, results in significant increase in the number of apoptotic osteocytes following fatigue loading. We believe controlling osteoblast differentiation/proliferation and osteocyte apoptosis could be promising directions for developing future therapeutic solutions for related bone diseases.

## Introduction

Bone tissues are sensitive to its mechanical environment [1]. It is well accepted that the presence of a reasonable level of mechanical stress on bones (known as *normal loading*) could enhance bone formation and maintain a healthy bone mass [2]. Prolonged absence of normal loading on bones—usually associated with extended physical inactivity due to injuries—could decrease bone formation and increase bone resorption, eventually leading to bone loss and disuse osteoporosis. When the level of mechanical stimulations exceeds the normal amount for an extended period of time, a stress condition known as *fatigue loading* could occur. In fatigue loading, *micro-damage* such as small cracks in bone tissues may

appear, triggering a cascade of bone remodeling processes that attempt to repair damaged bone tissues via sequential bone resorption and formation [3]. When fatigue loading conditions are not recognized early and addressed, the risks for bone injuries and bone diseases will increase. Therefore, understanding the constituents and functions of molecular repertoires involved in fatigue loading has been a central focus of study in molecular biology of the bone.

It still remains unknown what all the mechanically-sensitive genes and proteins in bone cells under mechanical stress are and how their differential expressions are regulated [4]. Past research identified osteoblast as being recruited to bone surfaces to form new bones in response to loading [5]. In fatigue loading conditions, the migration of osteoblast to the bone surface is known to co-occur with migrations of osteoblast

\* Correspondence: [jakechen@iupui.edu](mailto:jakechen@iupui.edu)

<sup>2</sup>Indiana University School of Informatics, Indianapolis, IN 46202, USA  
Full list of author information is available at the end of the article

progenitors and osteoblast to bone damaged areas, thus activating bone remodeling process and damage repairs [6-11]. This process requires temporal coordination of osteoblast and osteoblast to repair damaged bone tissues. Therefore, osteoblast-associated genes were reported and presumed to be involved with different levels of mechanical stimulation signals [12]. Several biochemical studies have also suggested that anabolic mechanical stimulation may increase the expression of *c-fos*, *osteopontin*, *COX-2*, *guanosine triphosphatases (GTPases)*, *adenylate cyclase*, *phospholipase C (PLC)*, and *mitogen-activated protein kinases (MAPKs)*, which can further lead to elevated expression of bone anabolic factors such as *prostaglandins* and *Nitric oxide* (See reference [13] for a review).

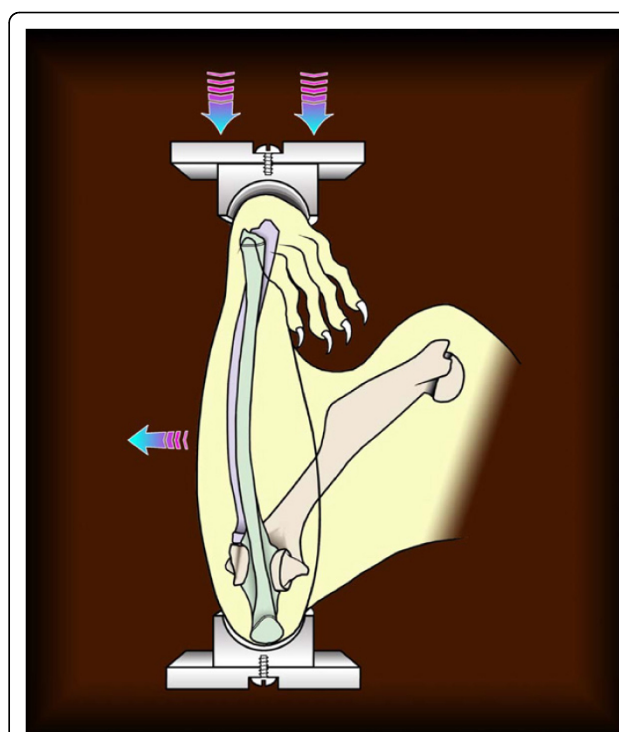
In this work, we performed the first proteomic study of mechanical loading of bone tissues using Rat as an animal model. Prior to our study, large-scale functional genomics analysis of the activation of bone remodeling process were performed in a few microarray studies [14,15]. While these earlier studies suggested osteocyte apoptosis and Wnt signaling pathways were two critical biological processes involved, proper controls against normal loading conditions were not performed in those experimental studies. It was not clear what mRNA level changes observed in fatigue loading were shared in common with normal loading. Nor is it clear whether the biological processes observed at the mRNA expression level could overlook critical protein changes, since many recent studies revealed that large-scale gene expression and proteomics tend to complement (instead of significantly overlap) with each other [16,17]. Elucidating proteomics level changes, particularly when integrated with prior findings of genes and new models developed at the molecular signaling network/pathway level, can lead to new insights on bone mechanical stress and development of novel molecular biomarkers.

## Experimental procedures

### Design of bone loading experiments using rat models

In order to study proteomics profile differences in living bone tissues, an ulnar axial compression loading system was chosen (see illustration in Figure 1). The system allows loading experimentation at different stress levels for animal models [6,10,11].

Female Sprague-Dawley Rat (age: 6 months; weight: 250-300 grams) were purchased from Harlan (Indianapolis, Indiana, USA). Animals were acclimatized for two weeks and housed in environmentally controlled rooms in Laboratory Animal Resource Center (LARC) of Indiana University School of Medicine and fed standard Rat chow and water *ad libitum*. All the procedures performed in this study were in accordance with the



**Figure 1** An illustration of the ulnar axial compression loading system to study the effects of different levels of mechanical stress on bones in animal models.

Indiana University Animal Care and Use Committee Guideline.

Nine animals were divided randomly into 3 groups: control (CTRL), loading (L) and fatigue loading (FL) groups. All the animals were anesthetized with an intraperitoneal injection of ketamine (60 mg/kg; Ketaset<sup>®</sup> – Fort Dodge Animal Health, Fort Dodge, IA) and xylazine (7.5 mg/kg; Sedazine<sup>®</sup> – Fort Dodge Animal Health, Fort Dodge, IA). The animals in the control group were sacrificed 96 hours post-injection without being subject to mechanical loading. The right ulnae of the remained animals were loaded or overloaded based on treatment groups. The animals in the loading group were loaded with a peak force of 20 N for 360 cycles and then sacrificed at 96 hours after the loading session. For the animals in fatigue loading group, one bout of loading with a peak force of 20 N at 2 Hz was not stopped until 10-15% stiffness loss. The overloaded animals were also sacrificed at 96 hours after the loading session.

Load was applied using a load-controlled, electromagnetic loading device. Total loading cycles was adjusted through the connected load controller. Stiffness loss during the loading procedure was observed through continuous monitoring of displacement of the arm on the loading device using a CCD Laser Displacement Sensor (LK Series, Keyence Corp. Osaka, Japan).

### Liquid chromatography coupled tandem mass spectrometry proteomics analysis

The ulnae were dissected out immediately and cleaned of all muscle and connective tissue after all the Rats were sacrificed. Both of 5-mm proximal and distal ends of the ulnae were removed. The remaining ulna midshafts were snap frozen in liquid nitrogen and stored at -80°C until protein isolation. For total protein isolation, Rat ulna midshafts were shattered and ground to a fine powder under liquid nitrogen using mortars and pestles. There were three groups (The control, loading and fatigue loading groups), three samples per group, and two HPLC injections per sample (Table 1).

Label-free protein identification and protein quantitative analysis services were performed by professionals at the Protein Analysis and Research Center/Proteomics Core of Indiana University School of Medicine, co-located at Monarch Life Sciences, Inc, Indianapolis. For a thorough review of the principle and method developed at Monarch, refer to the review by Wang *et al* [18].

The protein identification tasks were analyzed using standard commercial-strength protocols and commercial software packages developed at Monarch, which have supported many scientific research case studies in areas including proteomics studies, biomarker discovery, and bioinformatics analysis, e.g., [19-21]. Briefly, Tryptic peptides were analyzed using Thermo-Finnigan linear ion-trap mass spectrometer (LTQ) coupled with a HPLC system. Peptides were eluted with a gradient from 5 to 45% Acetonitrile developed over 120 minutes and data were collected in the *triple-play* mode (MS Scan, zoom scan, and MS/MS scan). The acquired raw peak list data were generated by XCalibur (version 2.0) using default parameters and further analyzed by an algorithm using default parameters described by Higgs *et al* [22]. MS database searches were performed against the combined protein data set from International Protein Index (IPI; version 1.2) [23] and the non-redundant NCBI-nr human protein database (2005 version), which totaled 22,180 protein records. The resulting MS/MS data were searched using SEQUEST Cluster from Thermo Scientific (bundled with BioWorks software suite version 2.70 based on the original SEQUEST algorithm [24]). During

search, we set the number of missed cleavages permitted to be 2. We search fixed modifications to be Iodoethanol on Cys and variable modifications to be Oxidation on Met. The mass tolerance for precursor ions were set at 2 Da and the mass tolerance for fragment ions were set at 0.7 Da. For novel protein that could not be positively identified by SEQUEST, we used the de novo sequencing function of the BioWorks software to obtain peptide sequence information for the collision-induced dissociation (CID) spectra. Various data processing filters for protein identification were applied to keep only peptides with the XCorr score above 1.5 for singly charged peptides, 2.5 for doubly charged peptides, and 3.5 for triply charged peptides. These XCorr scores were set according to linear discriminant analysis similar to that described in DTASelect (version 2.0) to control false-positive rate at below 5% levels. These empirical thresholds were validated in large data sets processed by Monarch in similar conditions and peptide identification parameters. The false positive rates of these large-scale studies under the used parameters were estimated from the number and quality of spectral matches to the decoy database.

Protein quantification tasks were also conducted using software developed at Monarch Life Sciences, Inc. First, all extracted ion chromatograms (XICs) were aligned by retention time. Each aligned peak were matched by precursor ion, charge state, fragment ions from MS/MS data, and retention time within a one-minute window. Then, after alignment, the area-under-the-curve (AUC) for each individually aligned peak from each sample was measured, normalized, and compared for relative abundance—all as described in [22]. The normalization methods by Higgs *et al* [22] were used, and the data were then transformed back to the original scale. Here, a linear mixed model generalized from individual ANOVA (Analysis of Variance) was used to quantify protein intensities and calculate statistical significance. In principle, the linear mixed model considers three types of effects when deriving protein intensities based on weighted average of quantile-normalized peptide intensities: 1) *group effect*, which refers to the fixed non-random effects caused by the experimental conditions or treatments that are being compared; 2) *sample effect*, which refers to the random effects (including those arising from sample preparations) from individual biological samples within a group; 3) *replicate effect*, which refers to the random effects from replicate injections from the same sample preparation. Standard statistical data pre-processing techniques, including quantile normalization and randomization of measurement orders, were applied first to eliminate technical bias due to random variations from biological samples and their replicates. The model fitting was performed in the SAS software (version 9)

**Table 1 The experimental design for proteomics analysis of bone loading in rat**

|      | Samples | Replicates | Injection runs (Subtotals) |
|------|---------|------------|----------------------------|
| CTRL | 3       | 2          | 6                          |
| L    | 3       | 2          | 6                          |
| FL   | 3       | 2          | 6                          |

The LC-MS/MS experiment consists of 3 groups × 3 samples × 2 replicates = 18 LC/MS injections run in random order. The three groups are: Controls (CTRL), Loaded (L), and Fully-Loaded (FL).

using PROC MIXED. The REML method was used as a fit mechanism and degrees of freedom were computed using the Satterthwaite method. The RANDOM statement was used to model the covariance with the NOBOUND parameter option in the PROC statement. The p-value estimates the proportion of times a change at least as big as evaluated will be observed if in fact there is no real change. All the p-values were then transformed into q-values that estimate the False Discovery Rate (FDR) [25].

### Homologous gene mapping of rat and human proteins

Due to the lack of protein-protein interaction data coverage in Rat, we map all Rat protein-encoding genes to their human gene homolog to take advantage of large sets of protein interaction data available in human. The homologous gene mapping involved four steps. First, we extracted all the Rat protein identifiers (IPI number and protein GI accessions) from the sequence annotation field of the proteomics search results. Second, we downloaded Rat IPI reference database version 1.2, which contains 38,873 sequence identifier mapping relationships among Rat Swissprot IDs, sequence accession numbers, and gene names. Third, we downloaded NCBI Homologene release 49.1. We filtered out genes from other organisms to include proteins only from Rat and human. After applying the filter, 14,558 remained in the homologene groups, which contain homology mapping relationships between 15,125 Rat genes and 14,753 human genes. We defined a “homolog gene match” between a Rat gene and a human gene as each pair found within the same homologene group. In the fourth step, we map the matched human genes back to human proteins, using Uniprot sequence annotation files. Note that the mapping between Rat protein to human protein based on gene homology relationships has the limitation of aggregating all alternative spliced protein isoforms together. However, this will not be a major concern, since the majority protein-protein interaction data are collected based on gene-level experimentation data and therefore do not offer isoform-level resolution anyway.

### Method for selecting candidate significantly differentially-expressed proteins

For candidate proteins, we refer to the list of proteins that satisfies statistical protein-selecting filters but still needs further scrutiny before a subset of them can be confirmed as biologically relevant. It is tempting to control false positives using high FC threshold and q-value (false discovery Rate adjusted p-value) when we try to select candidate proteins that are differentially expressed with statistically rigor. For example, the following threshold filter (the F1 filter) was suggested by the proteomics analysis software by default to control possible

false positives that may arise due to potential sources of variability (estimated to be up to 15%) from different sample and experimental errors:

$$F1 : FC(x|i) \geq 1.5 \& q - \text{value}(x|i) < 0.05$$

While a stringent filter is generally necessary for proteomics experiments, protein expression level changes in proteomics experiments are generally expected to be smaller than those often observed in expression microarrays, because changes in signaling proteins or regulatory proteins are expected to be subtle in general. In addition, the problem with applying default filters directly is that these filters fail to take into account of data that may be highly correlated from controlled comparative experiments with more than two conditions. In our case, we have three conditions FL for fatigue loading, L for normal loading, and CTRL for normal controls. If we can observe high degree of correlation of results that occur in FL vs. CTRL and in F vs. CTRL, the FC requirement and q-value requirement may be both relaxed to allow more interesting proteins that change barely in the “twilight zone” of >10%, as long as these proteins can be further validated using additional computational or experimental techniques.

Therefore, in complementary to fold change filter in F1, we developed a second experimental filter (the F2 filter) to select candidate proteins that changed significantly above 10% ( $FC \geq 1.1$ ) to show up, when we try to compare two *similar* conditions, FL\_vs\_L (Fatigue Loading against Normal Loading), in which data for L\_vs\_CTRL (Fatigue Loading against Controls) and FL\_vs\_CTRL (Normal Loading against Controls) are also available:

$$F2: FC(x|FL\_vs\_L) \geq 1.1 \text{ and} \\ q\text{-value}(x|FL\_vs\_CTRL) * q\text{-value}(x|L\_vs\_CTRL) < 0.0025 \text{ and} \\ p\text{-value}(x|FL\_vs\_CTRL) < 0.05 \ \& \ p\text{-value}(x|L\_vs\_CTRL) < 0.05$$

Here in this F2 filter, in addition to relaxing the FC threshold, we also modified how we should apply statistical q-value. Here, we introduce a concept that we'll refer to as the *triangulation property* of comparable analysis. Briefly, this property is met if and only if pairwise comparison results from three conditions, for example, CTRL, L, and FL, are consistent among themselves. In other words, we say a triangulation property exists among CTRL-L-FL if and only if proteins passing FL\_vs\_CTRL and L\_vs\_CTRL q-value filters with FC changes of  $f1$  and  $f2$  respectively are the same set of proteins that pass FL\_vs\_L with and same q-value filter and a FC threshold of  $f1/f2$  independently. In fact, no

proteomics search software that we know today guarantee such triangulation property due to inherent errors in the model that estimates statistical significance of peptides and proteins. In fact, we understand that the  $q$ -value was derived from a more stringent statistical model in early years of proteomics licensed from Eli Lilly (private communication with Dr. Mu Wang, who provided the proteomics service for this experiment). Therefore, we developed an easy-to-understand meta-analysis method, *q-value triangulation method*, in the F2 filter, so that we can rely primarily on better-understood  $p$ -value statistics. In this method, we assume the  $p$ -value calculations of two independent experiments, FL\_vs\_CTRL and L\_vs\_CTRL, are generally reliable and therefore can be controlled at 0.05. The *q-value* triangulation calculation for FL\_vs\_L is done by multiplying the respective *q-values* for FL\_vs\_CTRL and L\_vs\_CTRL comparisons controlled at the  $0.05^2 = 0.0025$  level. The reason why the  $p$ -values are chosen comparing to the control samples rather than comparing FL vs L is that comparing to the control samples with our statistic method can reduce baseline noise in proteomics data and detect weak patterns.

#### Normality probability plot calculation

To determine normality of the residual distribution, we use the normal probability plot to calculate the *normal quantiles* of all values in *Residue (i)*, or *Res\_FL\_L*. The values and the normal quantiles are then plotted against each other. Normal quantiles are computed using the  $f$ -value,  $f_i$ , which is calculated as:

$$f_i = \frac{i - 0.5}{n}$$

where  $i$  is the index of the value and  $n$  is the number of values. The *normal quantile*,  $q(f)$ , for a given  $f$ -value is the value for which  $P[X \leq q] = f$ , where  $X$  is a standard normally distributed variable [26].

#### Creation of bone tissue stimulated protein sub-networks

Differentially expressed candidate Rat proteins, which we successfully mapped to human proteins through homologous gene matching, are used as *seed proteins* to build a protein-protein interaction subnetwork. We derive this protein interaction sub-network using a nearest-neighbor expansion method initially described in [27]. In summary, we searched the seed proteins against a human protein-protein interaction database. We include additional proteins in this subnetwork if and only if these additional proteins are found to directly interact with at least one seed protein. The protein-protein interactions involved are also collected into the subnetwork. If the subnetwork does not form a large connected graph, the biological

functional distance among such seed proteins would be regarded as high. On the other hand, if the subnetwork does form a large connected graph, the biological functional distance among these seed proteins would be very close. The sub-network offers a good model to integrate proteomics results, from which drug target may be developed [20,27]. Since the seed proteins used are all proteins that are quantitatively changed under the FL\_vs\_L condition, this subnetwork is essentially an activated protein signaling network specific to bone cells' response to mechanical stress.

We use the Human Annotated and Predicted Protein Interaction (HAPPI) database [28] (<http://bio.informatics.iupui.edu/HAPPI/>) to retrieve high-quality protein interacting. We choose a human protein interaction database due to limited protein-protein interaction data available for Rat and the fact that Rat and human share the majority of biological processes in common. The HAPPI database is an open-access web-based relational database that contains a comprehensive collection of computer-annotated human protein-protein interactions involving 10,592 human proteins (identified by UniProt ID). Data in the HAPPI database are derived from both experimental data sources and computational predictions publicly available. Different from most protein-protein interaction databases, reliability of protein-protein interaction information is provided in the HAPPI database as *H scores*, which range between 0 to 1 or a quality star rank grade of 1, 2, 3, 4 and 5. Increased protein interaction grade from 1 to 5 have been shown to be associated with improved quality of physical interacting proteins and decreased amount of non-physical interactions found primarily in text mining or gene co-expression studies [29]. For this study, we only use protein interactions in the HAPPI database with star grade of 3 and higher (consisting of more than 280,000 human protein interactions of primarily physical interactions), which are comparable to the overall quality of HPRD, a much smaller reference human protein interaction database commonly used in bioinformatics.

#### Visualization of differentially expressed protein sub-network

To perform interaction network visualization, we used an internally developed software platform, ProteoLens [30], which can be freely downloaded from <http://bio.informatics.iupui.edu/teolens/>. ProteoLens is a biological network data mining and annotation platform that supports both standard GML files and relational data in Oracle or PostgreSQL Database Management System. It is a scalable data-driven biological network visualization software that enables expert bioinformatics users to browse database schemas and tables, filter and join relational data using SQL queries, and customize data fields to be visualized as network graphs.



## Results

### Cellular changes in bone tissues after mechanical stimulations

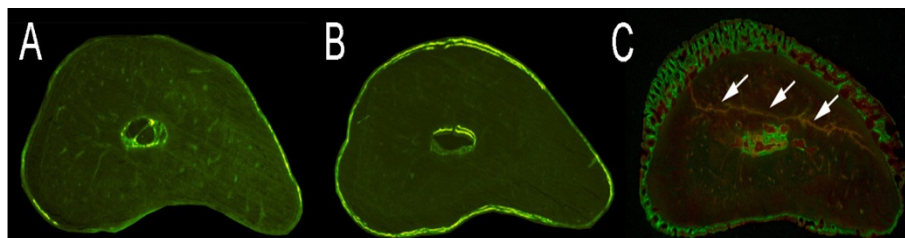
In Figure 2, we show a comparison of histological changes for bone tissues under control, normal loading, and fatigue loading conditions. In Figure 2A, we show a control without any mechanical stimulations. In Figure 2B, we show that bone formation in female SD Rat is significantly increased compared with the control, when one bout of axial loading of the ulna with a peak force of 20 N at 2 Hz for 360 cycles periosteal is applied. In Figure 2C, we show that substantial periosteal bone formation and microdamage in the cortex are generated, when fatigue loading with a peak force of 20 N at 2 Hz until 15% stiffness loss is applied.

### Proteomics changes between normal loading and fatigue loading conditions

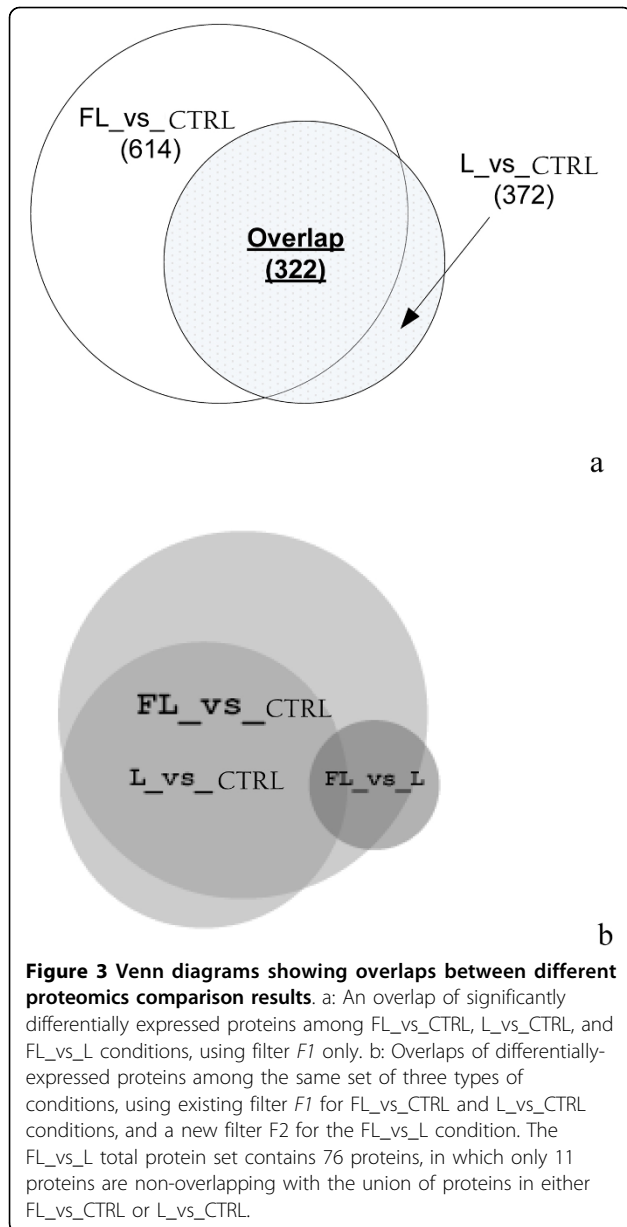
The Proteomics software mentioned in the method section reported a comprehensive list of 1,058 proteins that are differentially expressed among normal loading, fatigue loading, and controls. This list was derived from 5,361 IPI-identified Rat proteins observed in the LC-MS/MS experiment of all Rat samples. Among the 5,361 IPI-identified proteins, 578 have Xcorr = 'H' (i.e., "high confidence") and 4,783 have Xcorr = "L" (i.e., "low confidence"). The 1,058 differentially expressed Rat proteins can be mapped to 1,171 human proteins using homologous gene mapping methods (see Experimental Procedures for details). Note that only a fraction of these 1,058 proteins may have undergone through real quantitative changes, due to inherent variations of the proteomics platform and the high-variability nature of biological samples.

In Figure 3, we used Venn Diagrams to show overlaps among three proteomics comparative analysis results, i.e., FL\_vs\_CTRL (Fatigue Loading against Control), L\_vs\_CTRL (Normal Loading against Control), and FL\_vs\_L (Fatigue Loading against Normal Loading), by applying two different types of candidate protein selection filters, F1 and F2 (see Experimental Procedures for

details), for results derived from LC-MS/MS proteomics analysis of Rat samples. In Figure 3A, only F1 default filter was applied. It showed that there are 322 proteins overlapping between FL\_vs\_CTRL and L\_vs\_CTRL proteomics results. Combined together, the two data sets represented  $614 + 372 - 322 = 664$  total proteins that are quantitatively changed from either loading condition to controls. Note that FL\_vs\_L produced no "significant" protein list using the standard filter criteria, F1 (see Experimental Procedures for details). A plausible explanation is that FL and L are biologically "equivalent" conditions, which make their proteomics level expression indistinguishable. This is very unlikely, since the FL\_vs\_CTRL and L\_vs\_CTRL results overlap in significant portions but differently (for FL\_vs\_CTRL, overlap is  $322/614 = 52\%$ , for L\_vs\_CTRL, overlap is  $322/372 = 87\%$ ). A second and alternative explanation is that the filter F1 may be too stringent (requiring 1.5 fold change differences between loading conditions and controls) to allow detection of quantitative protein expression level changes, which may be quite subtle for FL\_vs\_L comparisons. Therefore, we applied the second filter, F2 (also see Experimental Procedures for an explanation), which provides relaxed (requiring  $FC \geq 1.1$ ) yet still statistically significant candidate protein selecting threshold for FL\_vs\_L differentially expressed proteins. By substituting filter F2 for F1 in the FL\_vs\_L condition, we show the new overlapping relationship among FL\_vs\_CTRL (using the original filter F1), L\_vs\_CTRL (using the original filter F1), and the new FL\_vs\_L (using the new filter F2) in Figure 3B. The new Venn Diagram has an added FL\_vs\_L protein set of 76 candidate proteins. Interestingly, 65 out of the 76 protein ( $65/76 = 86\%$ ) are overlapped with the existing 664 proteins differentially expressed and detected using the stringent filter F1. The high degree of overlap resulted in only a slight increase in the final combined data set of 679 candidate rat proteins associated with loading conditions. This observation is consistent with the assumption that applying the F2 filter to the FL\_vs\_L condition can still control false positives well. However,



**Figure 2 Cellular changes of bone tissues under control, normal loading, and fatigue loading conditions.** A: Control condition (no loading); B: Normal loading condition. The thick staining at the perimeter of bone tissues indicates bone formation; C: Fatigue loading condition. The microdamage (indicated by arrows) and bone formation at the peripherals of bone tissues are clearly visible.



since filter *F2* uses a fold change threshold of 1.1—much smaller than the 1.5 threshold used in filter *F1*, we believe that only a subset of the 76 candidate proteins that changed at the subtle amount may have true biological significance.

#### Statistical validation of candidate proteins based on correlated loading conditions

To examine how well the quantitative changes measured between FL\_vs\_CTRL and L\_vs\_CTRL conditions—a sign that should indicate how consistent and accurate fold changes reported in the proteomics results are, we performed a linear regression on two variables, FC\_CTRL\_FL as  $x$  variable and FC\_CTRL\_L as  $y$

response variable. All the 679 proteins were used but only the data points with both fold change reported were reported. In Table 2, we show the linear regression results, which has an  $R^2 = 0.98$ . This surprisingly high degree of correlation is perhaps attributable to the commercial operations (use of standard protocols and well-tested proteomics analysis platform that also supports high-volume commercial operations at Monarch Life Sciences). It also supports the use of filter *F2* that sets FC threshold at 1.1—a level normally too low to be trustworthy when CV (covariance) of proteomics results are at approximately 15% yet still acceptable for this particular experimental setup, due to high degree of correlations found for fold changes between FL\_vs\_CTRL and L\_vs\_CTRL condition.

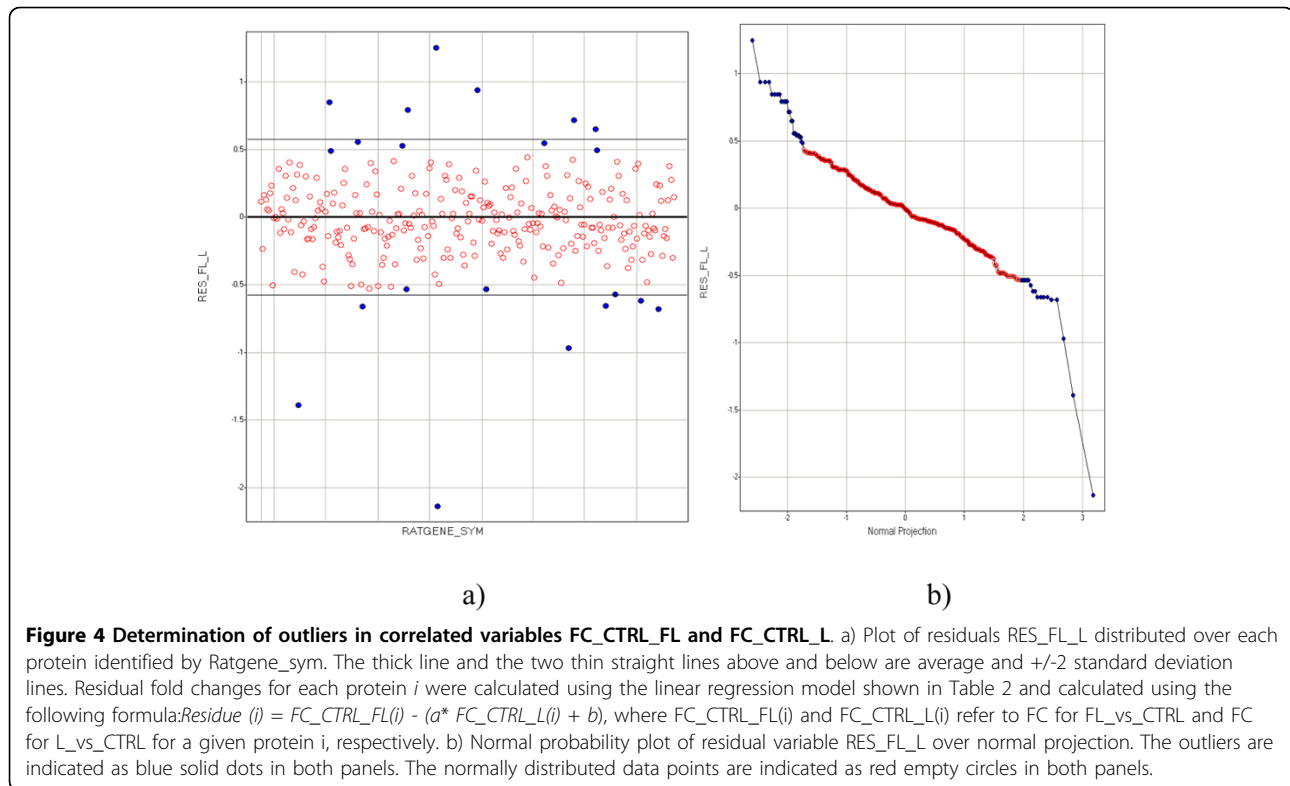
We further analyzed the residual plot for the above linear regression model and determined the normalcy data range (Figure 4). In Figure 4A, we observed that most residuals are evenly distributed within the  $\pm 2.0$  standard deviation range (between thin lines), with the exception of several residual extreme values that seemed not normally distributed around the mean (shown as a thick line in the center). To test if the residuals are normally distributed around the mean, we studied the residual normal probability plot (shown in Figure 4B). In regions showing normality, the plot follows a diagonal line. This suggests that residual values in the range vary as expected due to random errors predicted by the linear regression model. Otherwise, we could suspect that the residuals differ from one another by following a different model. In Figure 4B, we observed that the normal probability plot of *Res\_FL\_L* (Residuals of the FL\_vs\_CTRL against L\_vs\_CTRL after fitting the model described earlier) has good normality (linear) in the range of normal projection between  $-1.85$  and  $+1.85$  standard deviations of the mean. Outside this range, the *Res\_FL\_L* has a different slope, suggesting non-normality for the outliers from the bulk of data.

#### Validated proteomics results – proteins that quantitatively changed in fatigue loading conditions

Based on the residual distribution and normality probability test results, we reset the data outlier threshold to be within  $\pm 1.85$  standard deviation range in the residual plot, with which we narrow down to 42 proteins. Interestingly, the collection of these 42 proteins is a

**Table 2** Linear regression results of FC\_CTRL\_FL and FC\_CTRL\_L variables on differentially expressed proteins in all 3 conditions of the study

| Regression parameter | Slope (a) | Intercept (b) | Data point count | $R^2$ |
|----------------------|-----------|---------------|------------------|-------|
| value                | 1.09      | 0.03          | 679              | 0.98  |



subset of the 76 candidate proteins from the FL\_vs\_L condition that passed filter *F2*. These 42 proteins correspond to 21 genes, which we showed in Table 3.

In this table, we can further make several observations. First, protein ranks (indicator of confidence of detection during search) derived from MS search software result as a default is not a reliable predictor for the proteins' biological significance. All significantly differentially expressed proteins in Table 3 have quite low protein ranks, varying between 1500 and 2100. Second, the patterns for differential expression changes are varied from one gene to another. For example, Capon, Ddx21a, Rab40b (predicted), pdcd8, Serbinb13 (predicted) are all induced multiple folds from the resting stage; Fbf1 (predicted), Pik4cb (predicted), Fcho2 (predicted), Slc1a3 (predicted) are all suppressed significantly from the resting stage; and Ddx18, Mrpl53 (predicted), and Mrpl45 (predicted) are all significantly changed for FL\_vs\_CTRL conditions from L\_vs\_CTRL conditions. Third, we have shown that at least in some cases, a protein may be significantly differentially expressed in the FL\_vs\_L condition for many reasons, not necessarily due to a high FC\_FL\_L, e.g., Capon and Rab40 (predicted)—both due to high FC\_CTRL\_L and FC\_CTRL\_FL. Additional details of the protein quantification results for the proteins corresponding to the 21 genes are shown in Supplementary Table 1.

#### Activated protein signaling sub-network of molecular response to fatigue loading

We mapped all significant Rat proteins to human proteins using gene homolog matching method describe in the Experimental Procedures. 1,058 significantly changed Rat IPI-identified proteins (using the *F2* filter on all comparative studies) out of 5,361 IPI-identified Rat proteins from the LC-MS/MS experiment were involved in the mapping. These IPI-identified Rat proteins can be mapped to 513 unique known Rat gene names (the decrease was primarily due to aggregation of proteins isoforms mapped to the same gene). 482 out of the original 513 Rat genes were successfully mapped to 484 human genes using the NCBI Homologene database. The 484 human genes were mapped to 1,171 human proteins identified with UniProt IDs. The slight increase in total protein count from initial 1,058 Rat proteins to 1,171 human proteins suggest that there were a small percentage of one-to-many homologous mapping relationships between Rat and human proteins.

Then, using the 42 Rat proteins representing 21 Rat genes (as shown in Table 3) as seed proteins, we built a protein interaction subnetwork. This network represented a coarse biological model that integrated prior knowledge of the functional interaction relationships among proteins and the latest acquired proteomics knowledge on proteins quantitatively changed under



**Table 3 A list of 21 Rat genes whose proteins are found to be differentially expressed with statistical significance between FL\_vs\_CTRL and L\_vs\_CTRL conditions**

| Rat Gene             | Human Gene | FC (CTRL_L) | FC (CTRL_FL) | FC (FL_L) | Max Confidence | Peptide Evidence |
|----------------------|------------|-------------|--------------|-----------|----------------|------------------|
| Capon                | NOS1AP     | 6.72884     | 6.00145      | 1.1212    | 0.98           | ≥6               |
| Ddx18                | DDX18      | 1.14716     | 2.13095      | -1.85759  | 0.98           | ≥6               |
| Ddx21a               | DDX21      | 3.28614     | 4.10949      | -1.25055  | 0.96           | ≥6               |
| Fbf1_predicted       | FBF1       | -3.10292    | -2.81444     | -1.1025   | 0.98           | ≥6               |
| Fcho2_predicted      | FCHO2      | -1.97277    | -2.79227     | 1.41541   | 0.98           | ≥6               |
| Klk14_predicted      | KLK14      | 1.2212      | 1.88874      | -1.54662  | 0.98           | ≥6               |
| LOC301506            | FSD1       | -2.77612    | -3.54757     | 1.27789   | 0.99           | ≥6               |
| LOC306805            | ASPN       | 1.83348     | 2.8254       | -1.54101  | 0.99           | ≥12              |
| Mrpl45_predicted     | MRPL45     | 2.47117     | 3.98149      | -1.61118  | 0.99           | ≥6               |
| Mrpl53_predicted     | MRPL53     | 3.70412     | 1.94325      | 1.90615   | 0.96           | ≥6               |
| Pdcd8                | PDCC8      | 2.91378     | 4.15437      | -1.42577  | 0.96           | ≥6               |
| Pik4cb               | PIK4CB     | -2.77612    | -3.54757     | 1.27789   | 0.99           | ≥6               |
| RGD1562139_predicted | RPL29      | 2.47771     | 3.28214      | -1.32467  | 0.98           | ≥6               |
| Rab40b_predicted     | RAB40B     | 5.42109     | 4.99103      | 1.08617   | 0.98           | ≥6               |
| Raf1                 | RAF1       | -2.1328     | -1.59117     | -1.3404   | 0.97           | ≥6               |
| Sema5b_predicted     | SEMA5B     | 1.75998     | 2.60246      | -1.47869  | 0.99           | ≥6               |
| Serpinb13_predicted  | SERPINB13  | 3.01946     | 3.82539      | -1.26691  | 0.97           | ≥6               |
| Slc1a3               | SLC1A3     | -1.97126    | -2.78988     | 1.41528   | 0.98           | ≥6               |
| Slc4a3               | SLC4A3     | 2.15184     | 1.80834      | 1.18995   | 0.96           | ≥6               |
| Tex101               | TEX101     | 2.007       | 1.60395      | 1.25128   | 0.97           | ≥6               |
| Upf2_predicted       | UPF2       | -1.72341    | -2.54157     | 1.47474   | 0.98           | ≥6               |

\* "Max Confidence" was calculated as 1- smallest q-value among all the comparison conditions (FL\_L, CTRL\_FL, and CTRL\_L). "Peptide evidence" refers to total number of peptides per group used to calculate Fold Change (FC) and q-value in groupwise comparisons for protein quantifications.

fatigue loading conditions compared with normal loading conditions. After the protein interaction network expansion, the initial 42 seed proteins became expanded into a set of 394 human protein interacting pairs covered by 297 human proteins. In Figure 5, we show a visualization of the FL\_vs\_L expanded human protein interaction sub-network (network with only one pair of interactions are not shown). The largest connected component of this network consists of 9 genes (to be discussed in the next section), which can be used to reason about molecular mechanisms why these proteins changed during mechanical stress conditions that ultimately lead to microdamage in bones.

#### Pathway-protein association analysis

The 42 Rat proteins representing 21 Rat genes (as shown in Table 3) were also used to perform pathway-protein association analysis using the Kyoto Encyclopedia of Genes and Genomes (<http://www.genome.ad.jp/kegg/>) [33]. Significance level for pathway comparisons was set by represented number >3 due to results of small counts. This allows avoiding any assumptions about the shape of sampling distribution of population.

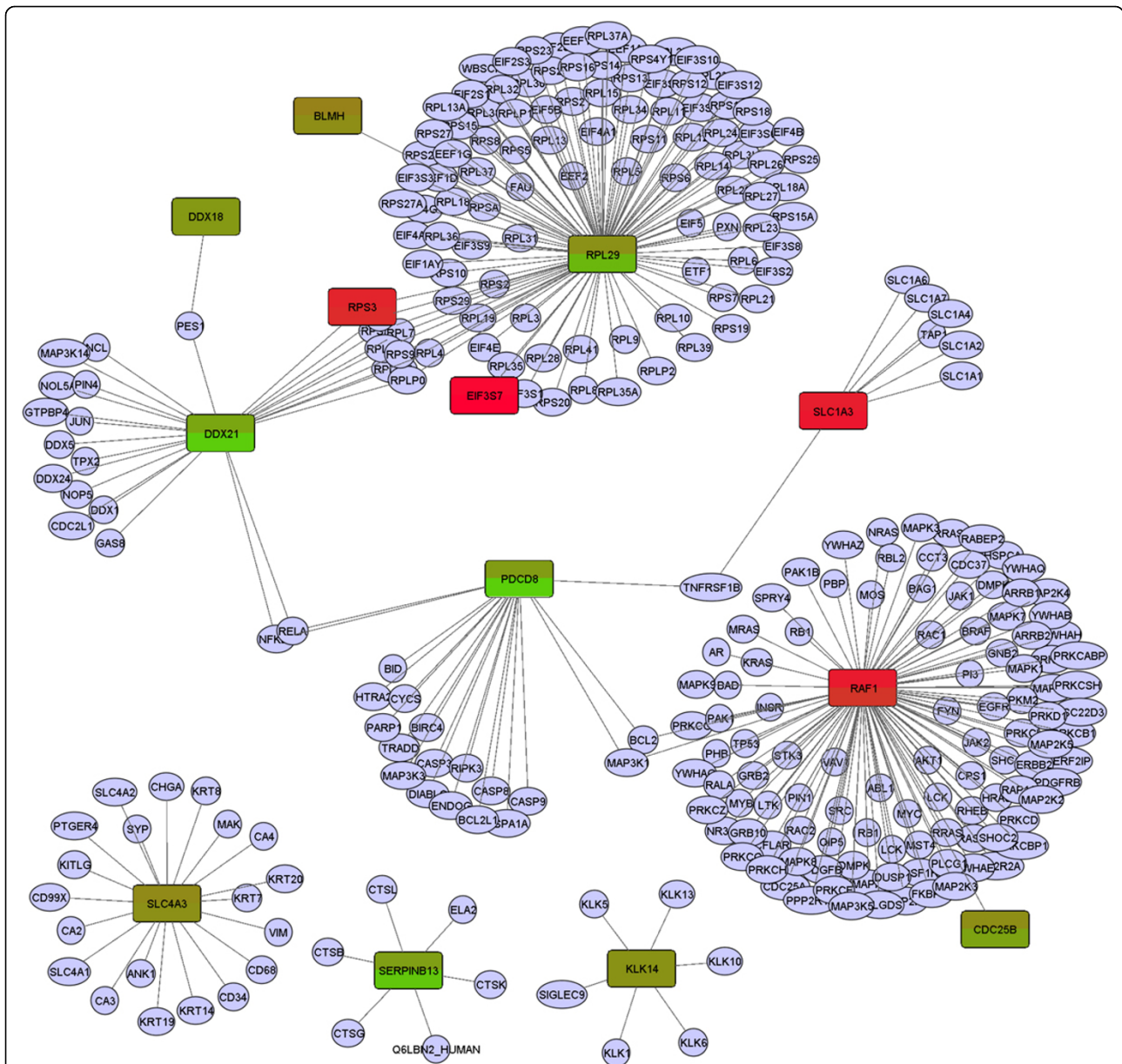
This pathway protein association matrix maps all the biological pathways with pathway proteins. It enriches the top frequent pathways in a given list of pathways, which helps in discovering pathway markers. In Figure

6, 36 pathways and 21 proteins are associated with each other for three comparisons (red for CTRL\_L; green for CTRL\_FL; and blue for FL\_L).

#### Discussions

Mechanical stimulation may cause bone cells to express mechano-sensitive genes and proteins through membrane receptors and ion channels and downstream intracellular signaling cascades [34-36]. These would lead to differentiation of osteoblast progenitor cells and osteoblast proliferation [5]. Besides increase in bone formation, fatigue loading produce microdamage [9] in the cortex which also leads to osteocyte apoptosis and further activate bone remodeling through which the damaged cortical bone is repaired [6,37].

In our study, we have found the enhanced expression of proteins involved in receptor binding, RNA processing, cell division and etc. Cell division cycle 25 homolog B (CDC25B), DEAD (Asp-Glu-Ala-Asp) box polypeptide 21 (DDX21), ribosomal protein L29 (RPL29) (seed proteins) and the expanded proteins as shown in Figure 5 were up-regulated. CDC25B that plays a role in cell division seems to allow cell to go into cell division during fatigue loading [38]. DDX21 and RPL29 all are elevated in exercise conditions, and further elevated in fatigue exercise conditions. DDX21 is putative RNA helicase involved in RNA secondary structure alteration,



**Figure 5 A protein interaction sub-network of FL\_vs\_L expanded differentially expressed proteins.** Nodes colored in red or green are FL\_vs\_L differentially expressed proteins (*seeds*) and nodes in light purple are non-seed expanded proteins recruited through human protein interactions. Edges represent protein interactions recorded in the HAPPI database. Only HAPPI database protein interactions with quality ratings at or above 3 are used. Proteins that are significantly differentially expressed in FL\_vs\_CTRL or L\_vs\_CTRL conditions are also shown using the same color legend for FL\_vs\_L seed proteins, with the rectangle split into two half panels: the upper panel shows the gradient red (FC\_CTRL\_L > 0) or green (FC\_CTRL\_L < 0) colors for the FC\_CTRL\_L value, while the lower panel shows the gradient red or green color using the same color profile for the FC\_CTRL\_FL value. Standalone networks with only one pair of interactions are not shown.

and Ribosome reassembly [39]. RPL29 is ribosomal protein L29 involved in cell surface hairpin protein binding [40].

NOS (Nitric Oxide Synthase) is increased under the loading condition and further elevated by fatigue loading in this study. NOS is the enzyme to produce Nitric Oxide (NO) in cells [41]. NO has been shown to

increase in response to mechanical stimulation in osteoblastic cells [42]. It is also involved in mechanically induced bone formation in vivo [43]. Our study further verifies that NOS may mediate load induced bone formation at the periosteal surface in loading and fatigue loading groups. In addition, the further elevated NOS level under fatigue loading condition suggests NO may

also play a key role in mediating the repair of bone damage, such as recruitment of osteoclast precursor, because its actions include changes of the vascular permeability of the damaged area and stimulation of angiogenic activity [41].

Several apoptosis related proteins have been found to change significantly in the current study. Raf1 human (RAF proto-oncogene serine/threonine-protein kinase) was down regulated in the present study. It has a role in the transduction of mitogenic signals from the cell membrane to the nucleus [44]. Raf1 may promote cell survival by antagonizing apoptosis signals-regulating kinase [45]. Our study indicates that loss of Raf1 coincides with increased number of apoptotic osteocytes resulting from fatigue loading, suggesting that Raf1 has a role in

protection of osteocytes apoptosis. On the other hand, PDCD8 (Programmed cell death 8) is up-regulated under fatigue loading condition. Because PDCD8 is an apoptosis-inducing factor [46], it may induce osteocytes apoptosis following fatigue loading. Taken together, our study shows that the combination of decreased anti-apoptotic factor, Raf1, and increased pro-apoptotic factor, PDCD8, results in significant increase in the number of apoptotic osteocytes following fatigue loading. Several downstream proteins of Raf1 and PDCD8 pathways, such as Bcl2 and caspase proteins have previously been shown to be involved in osteocyte apoptosis induced by fatigue loading [37,47]. Therefore, this study suggests that drugs targeting on Raf1 and PDCD8 may regulate bone metabolism via prevention of osteocyte apoptosis.

| PATHWAY ID | PATHWAY NAME                              | # Represented | # of Pathway Molecules | A4QP84_HUMAN | A6PVV8_HUMAN  | A6PVV9_HUMAN | A1FMI_HUMAN | B1ALN1_HUMAN | B5ME80_HUMAN | B7ZBU1_HUMAN | C9IY14_HUMAN | C9J2U6_HUMAN | C9J3L4_HUMAN | C9J9Z7_HUMAN | C9JKR3_HUMAN | C9JTV9_HUMAN | C9JTX2_HUMAN | C9JVG6_HUMAN | C9JX17_HUMAN | P14KB_HUMAN | Q6IP11_HUMAN | RAF1_HUMAN | RL29_HUMAN | SEM5B_HUMAN |
|------------|---|---------------|------------------------|--------------|---------------|--------------|-------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------------|--------------|------------|------------|-------------|
|            |   |               |                        | hsa04360     | Axon guidance | 6            | 129         |              |              |              |              |              | ↑↑↓          |              |              |              |              | ↑↑↓          | ↑↑↓          | ↑↑↓         | ↑↑↓          |            |            |             |
| hsa00562   | Inositol phosphate metabolism             | 5             | 93                     | ↓↑           | ↓↑            | ↓↑           |             |              |              | ↓↑           |              |              |              |              |              |              |              |              |              | ↓↑          | ↓↑           |            |            |             |
| hsa04070   | Phosphatidylinositol signaling system     | 5             | 103                    | ↓↑           | ↓↑            | ↓↑           |             |              |              | ↓↑           |              |              |              |              |              |              |              |              |              | ↓↑          | ↓↑           |            |            |             |
| hsa03010   | Ribosome                                  | 4             | 86                     |              |               |              |             |              |              | ↑↑↓          |              |              |              |              |              |              |              | ↑↑↓          |              |             |              |            |            | ↑↑↓         |
| hsa04010   | MAPK signaling pathway                    | 3             | 275                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              | ↑↑↓          |              |             |              |            |            | ↑↑↓         |
| hsa04012   | ErbB signaling pathway                    | 3             | 91                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04062   | Chemokine signaling pathway               | 3             | 196                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04210   | Apoptosis                                 | 3             | 89                     | ↑↑↓          |               |              | ↑↑↓         | ↑↑↓          |              |              |              |              |              |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04270   | Vascular smooth muscle contraction        | 3             | 139                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04370   | VEGF signaling pathway                    | 3             | 81                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04510   | Focal adhesion                            | 3             | 204                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04540   | Gap junction                              | 3             | 102                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04650   | Natural killer cell mediated cytotoxicity | 3             | 141                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04660   | T cell receptor signaling pathway         | 3             | 112                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04662   | B cell receptor signaling pathway         | 3             | 79                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04664   | Fc epsilon R1 signaling pathway           | 3             | 89                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04666   | Fc gamma R-mediated phagocytosis          | 3             | 105                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04720   | Long-term potentiation                    | 3             | 79                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04722   | Neurotrophin signaling pathway            | 3             | 131                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04730   | Long-term depression                      | 3             | 80                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04810   | Regulation of actin cytoskeleton          | 3             | 221                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04910   | Insulin signaling pathway                 | 3             | 142                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04912   | GnRH signaling pathway                    | 3             | 107                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa04916   | Melanogenesis                             | 3             | 109                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa05200   | Pathways in cancer                        | 3             | 351                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa05210   | Colorectal cancer                         | 3             | 84                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa05211   | Renal cell carcinoma                      | 3             | 73                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa05212   | Pancreatic cancer                         | 3             | 74                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa05213   | Endometrial cancer                        | 3             | 53                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa05214   | Glioma                                    | 3             | 69                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa05215   | Prostate cancer                           | 3             | 100                    |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa05218   | Melanoma                                  | 3             | 72                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa05219   | Bladder cancer                            | 3             | 42                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa05220   | Chronic myeloid leukemia                  | 3             | 76                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa05221   | Acute myeloid leukemia                    | 3             | 61                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |
| hsa05223   | Non-small cell lung cancer                | 3             | 60                     |              |               |              |             |              |              |              | ↓↓           | ↓↓           | ↓↓           |              |              |              |              |              |              |             |              |            |            | ↑↑↓         |

**Figure 6 A pathway-protein association matrix of differentially expressed proteins.** The proteins in the first row from the fifth column to 25<sup>th</sup> column are differentially expressed with statistical significance between FL\_vs\_CTRL and L\_vs\_CTRL. The first column is KEGG pathway ID, the second column is KEGG pathway name, the third column is number of represented proteins in a pathway, and the fourth column is the total number of molecular in a pathway. The up-arrow represents up-regulated expression, and the down-arrow represents down-regulated expression. Three comparisons are shown (red for CTRL\_L, green for CTRL\_FL, and blue for FL\_L).

In the pathway-protein association analysis, a list of 42 rat proteins differentially expressed with statistical significance between FL\_vs\_CTRL and L\_vs\_CTRL is used to identify topmost frequent pathways. Of the 36 pathways in Figure 6, 13 are related to cancers; 18 to cellular processes (6 immune system, 3 nervous system, 3 endocrine system, 2 cell communication, 1 cell growth and death, 1 cell motility, 1 circulatory system, 1 development); 4 to signal transduction; and 1 to carbohydrate metabolism. The top eight pathways are Axon Guidance, Inositol Phosphate Metabolism, Phosphatidylinositol Signaling System, Ribosome, MAPK Signaling Pathway, Erbb Signaling Pathway, Chemokine Signaling Pathway, and Apoptosis. Some of those pathways have been reported to be related to bone metabolism. For example, neural regulation of bone metabolism mediated in osteoblastic and osteoclastic cells via Axon Guidance pathway has been demonstrated in histochemical and pharmacological studies [48] and Togari et al., in their paper, suggested the extension of axons of peripheral sensory and sympathetic neurons to osteoblastic and osteoclastic cells and the possible neural regulation of bone metabolism in these osteogenic cells. Inositol phosphate metabolism and signal transduction pathways was reported to regulate cytoplasmic Ca<sup>2+</sup> concentrations in osteoblastic bone cells [49]. In addition, Kennea et al. suggested that there would be robust and functional intrinsic and extrinsic apoptotic pathways in human fetal mesenchymal stem cells or bone marrow-derived stromal cells which could participate in the repair of mesodermal tissues, such as bone in osteogenesis imperfecta and heart muscle in cardiac ischaemia [50].

Of the 21 proteins, PDCD8 (A4QPB4\_HUMAN; AIFM1\_HUMAN; B1ALN1\_HUMAN) which is up regulated with statistical significance between FL\_vs\_CTRL and L\_vs\_CTRL are involved in Apoptosis pathway, RAF1 (C9J2U6\_HUMAN; C9J3L4\_HUMAN; RAF1\_HUMAN) which is down regulated with statistical significance between FL\_vs\_CTRL and L\_vs\_CTRL is involved in Cancers, Cellular Processes, and Signal Transduction Pathways. This further indicates the effect of decreased anti-apoptotic factor, Raf1, and increased pro-apoptotic factor, PDCD8, on the increase in the number of apoptotic osteocytes following fatigue loading.

We also found other pathway-protein associations such as PI4KB in Inositol phosphate metabolism and Phosphatidylinositol signaling system pathways, SEMA5B in Axon guidance, and RPL29 in Ribosome. Some of them are linked to bone metabolism or bone formation by previous reports. For example, Miller et al. reported that the presence of HIP/RPL29 during early chondrogenesis is essential for normal skeletal growth and patterning. They designed a ribozyme-mediated

knock-down approach to partially down-regulate HIP/RPL29 expression in the multipotent mouse embryonic skin fibroblast cell line C3H/10T. To investigate the role of HIP/RPL29 normal expression during cartilage formation [51]. And Mary showed that SEMA5B is a nerve guidance factor which is involved in invasive growth, vascular patterning, axon guidance, and bone development [52].

In addition, Rab40b is a member of Ras oncogene family [53]. Ras oncogenes are small GTP-binding proteins [53]. Besides their role in cell proliferation, Ras paradoxically induce both pro- and anti-apoptotic signaling [54]. It remains to be investigated whether Ras plays any role in osteocyte apoptosis following fatigue loading.

There is a possibility that other proteins, such as MRPL45, SLC1A3, UPF2 and ASPN identified in this study are involved in bone response to mechanical loading. ASPN has been found to be related to osteoarthritis [32]. It is remained to be investigated if MRPL45, SLC1A3 and UPF2 as intracellular transports could be stimulated by mechanical stimulation.

In conclusion, using an integrated LC-MS/MS proteomics analysis for the first time in bone mechanical stimulation studies, we have identified several essential proteins related to cell division, which can be linked to osteoblast differentiation and proliferation and bone formation eventually in response to loading. More importantly, our study identified several new proteins associated with osteocyte apoptosis induced by fatigue loading. Our results suggest new insights for future investigation of these proteins as candidate drug targets to regulate bone metabolism and repair bone damage.

## Additional material

**Additional file 1: Protein quantification data for the 21 Rat genes whose proteins levels are significantly changed in Loaded (L) or Fully Loaded (FL) conditions compared with controls (CON).** "CON\_L" refers to comparing L to CON. "CON\_FL" refer to comparing FL to CON. "FL\_L" refers to comparing L to FL. q-value refers to adjusted p-values. While p-value is an estimate of false positive rate, q-value is an estimate of false discovery rate (FDR). FC refers to Fold Change. "Mean CON/L/FL" refers to mean protein intensities. "%CV Injection" refers to % Coefficient of Variation for injection variation, "%CV Inj + Sample %" refers to the Coefficient of Variation for injection plus sample variation. "# of peptides/group" refers to the number of distinct identified peptides for this protein in any of the three groups: CON, L, or FL. "Mean Xcorr" refers to the mean Xcorr of the peptides identified for this protein.

## Acknowledgements

This work was partially supported by a Clinical Proteomic Technology Assessment for Cancer (CPTAC) grant to Jake Chen (co-investigator) from the National Cancer Institute (U24CA126480-01), seed funding to Indiana Center for Systems Biology and Personalized Medicine from the Indiana University - Purdue University Indianapolis, and NASA grant to Jiliang Li (NNA04CD04G). We thank Dr. Mu Wang for consultation with us to develop

a proteomics service plan for this project and performing the LC-MS/MS proteomics service. We also thank Ragini Pandey for helping us organize materials during the manuscript writing process.

This article has been published as part of *BMC Systems Biology* Volume 5 Supplement 3, 2011: BIOCAMP 2010 - The 2010 International Conference on Bioinformatics & Computational Biology: Systems Biology. The full contents of the supplement are available online at <http://www.biomedcentral.com/1752-0509/5?issue=53>.

#### Author details

<sup>1</sup>Department of Biology, Purdue School of Science, Indiana University Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202, USA. <sup>2</sup>Indiana University School of Informatics, Indianapolis, IN 46202, USA. <sup>3</sup>Indiana Center for Systems Biology and Personalized Medicine, Indianapolis, IN 46202, USA. <sup>4</sup>Department of Computer and Information Science, Purdue University School of Science, Indianapolis, IN 46202, USA.

#### Authors' contributions

JYC conceived the initial work, designed the method for the data construction. JL implemented the design of bone loading experiments and generated the proteomics data using rat models. FZ collected and analyzed the MS data, performed the statistical analyses. All authors are involved in the drafting and revisions of the manuscript.

#### Competing interests

The authors declare that they have no competing interests.

Published: 23 December 2011

#### References

- Robling AG, Castillo AB, Turner CH: **Biomechanical and molecular regulation of bone remodeling.** *Annu Rev Biomed Eng* 2006.
- Ziegler R, Scheidt-Nave C, Scharla S: **Pathophysiology of osteoporosis: unresolved problems and new insights.** *J Nutr* 1995, **125**(7 Suppl):2033S-2037S.
- Stromsoe K: **Fracture fixation problems in osteoporosis.** *Injury* 2004, **35**(2):107-113.
- Villemure I, Chung MA, Seck CS, Kimm MH, Matyas JR, Duncan NA: **Static compressive loading reduces the mRNA expression of type II and X collagen in rat growth-plate chondrocytes during postnatal growth.** *Connect Tissue Res* 2005, **46**(4-5):211-219.
- Turner CH, Owan I, Alvey T, Hulman J, Hock JM: **Recruitment and proliferative responses of osteoblasts after mechanical loading in vivo determined using sustained-release bromodeoxyuridine.** *Bone* 1998, **22**(5):463-469.
- Bentolila V, Boyce TM, Fyhrie DP, Drumb R, Skerry TM, Schaffler MB: **Intracortical remodeling in adult rat long bones after fatigue loading.** *Bone* 1998, **23**(3):275-281.
- Burr DB, Martin RB, Schaffler MB, Radin EL: **Bone remodeling in response to in vivo fatigue microdamage.** *J Biomech* 1985, **18**(3):189-200.
- Li J, Miller MA, Hutchins GD, Burr DB: **Imaging bone microdamage in vivo with positron emission tomography.** *Bone* 2005, **37**(6):819-824.
- Li J, Waugh LJ, Hui SL, Burr DB, Warden SJ: **Low-intensity pulsed ultrasound and nonsteroidal anti-inflammatory drugs have opposing effects during stress fracture repair.** *J Orthop Res* 2007.
- Tami AE, Nasser P, Schaffler MB, Knothe Tate ML: **Noninvasive fatigue fracture model of the rat ulna.** *J Orthop Res* 2003, **21**(6):1018-1024.
- Verborgt O, Gibson GJ, Schaffler MB: **Loss of osteocyte integrity in association with microdamage and bone remodeling after fatigue in vivo.** *J Bone Miner Res* 2000, **15**(1):60-67.
- Pavalko FM, Norvell SM, Burr DB, Turner CH, Duncan RL, Bidwell JP: **A model for mechanotransduction in bone cells: the load-bearing mechanosomes.** *J Cell Biochem* 2003, **88**(1):104-112.
- Rubin J, Rubin C, Jacobs CR: **Molecular pathways mediating mechanical signaling in bone.** *Gene* 2006, **367**:1-16.
- Armstrong VJ, Muzylak M, Sinters A, Zaman G, Saxon LK, Price JS, Lanyon LE: **Wnt/beta-catenin signaling is a component of osteoblastic bone cell early responses to load-bearing and requires estrogen receptor alpha.** *J Biol Chem* 2007, **282**(28):20715-20727.
- Lau KH, Kapur S, Kesavan C, Baylink DJ: **Up-regulation of the Wnt, estrogen receptor, insulin-like growth factor-I, and bone morphogenetic protein pathways in C57BL/6J osteoblasts as opposed to C3H/HeJ osteoblasts in part contributes to the differential anabolic response to fluid shear.** *J Biol Chem* 2006, **281**(14):9576-9588.
- Ott LW, Resing KA, Sizemore AW, Heyen JW, Cocklin RR, Pedrick NM, Woods HC, Chen JY, Goebel MG, Witzmann FA, et al: **Tumor Necrosis Factor-alpha- and interleukin-1-induced cellular responses: coupling proteomic and genomic information.** *J Proteome Res* 2007, **6**(6):2176-2185.
- Xie L, Pandey R, Xu B, Tsapralis G, Chen QM: **Genomic and proteomic profiling of oxidative stress response in human diploid fibroblasts.** *Biogerontology* 2009, **10**(2):125-151.
- Wang M, You J, Bemis KG, Tegeler TJ, Brown DP: **Label-free mass spectrometry-based protein quantification technologies in proteomic analysis.** *Brief Funct Genomic Proteomic* 2008, **7**(5):329-339.
- Harezlak J, Wu MC, Wang M, Schwartzman A, Christiani DC, Lin X: **Biomarker discovery for arsenic exposure using functional data. Analysis and feature learning of mass spectrometry proteomic data.** *J Proteome Res* 2008, **7**(1):217-224.
- Chen JY, Pinkerton SL, Shen C, Wang M: **An integrated computational proteomics method to extract protein targets for Fanconi anemia studies.** In *21st Annual ACM Symposium on Applied Computing. Volume 1.* Dijon, France; 2006:173-179.
- McBride WJ, Schultz JA, Kimpel MW, McClintick JN, Wang M, You J, Rodd ZA: **Differential effects of ethanol in the nucleus accumbens shell of alcohol-preferring (P), alcohol-non-preferring (NP) and Wistar rats: a proteomics study.** *Pharmacol Biochem Behav* 2009, **92**(2):304-313.
- Higgs RE, Knierman MD, Gelfanova V, Butler JP, Hale JE: **Comprehensive label-free method for the relative quantification of proteins from biological samples.** *J Proteome Res* 2005, **4**(4):1442-1450.
- Kersey PJ, Duarte J, Williams A, Karavidopoulou Y, Birney E, Apweiler R: **The International Protein Index: an integrated database for proteomics experiments.** *Proteomics* 2004, **4**(7):1985-1988.
- Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR: **Direct analysis of protein complexes using mass spectrometry.** *Nat Biotechnol* 1999, **17**(7):676-682.
- Benjamini Y, Yekutieli D: **The control of the false discovery rate in multiple testing under dependency.** *Ann Statist* 2001, **29**(4):1165-1188.
- Rice JA: Belmont, CA: Duxbury Press; 2 1995.
- Chen JY, Shen C, Sivachenko A: **Mining alzheimer disease relevant proteins from integrated protein interactome data.** *Pac Symp Biocomput* 2006, 367-378.
- Chen JY, Mamidipalli S, George B: **HAPPI: A Database of Human Annotated Protein-Protein Interactions.** 2007, 2006.
- Chen JY, Mamidipalli S, Huang T: **HAPPI: an online database of comprehensive human annotated and predicted protein interactions.** *BMC Genomics* 2009, **10**(Suppl 1):S16.
- Huan T, Sivachenko A, Harrison S, Chen JY: **ProteoLens: a visual analytic tool for multi-scale database-driven biological network data mining.** *BMC Bioinformatics* 2008, **9**(Suppl 9):S5.
- Kizawa H, Kou I, Iida A, Sudo A, Miyamoto Y, Fukuda A, Mabuchi A, Kotani A, Kawakami A, Yamamoto S, et al: **An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis.** *Nat Genet* 2005, **37**(2):138-144.
- Jiang Q, Shi D, Yi L, Ikegawa S, Wang Y, Nakamura T, Qiao D, Liu C, Dai J: **Replication of the association of the aspartic acid repeat polymorphism in the asporin gene with knee-osteoarthritis susceptibility in Han Chinese.** *J Hum Genet* 2006, **51**(12):1068-1072.
- Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, et al: **KEGG for linking genomes to life and the environment.** *Nucleic Acids Res* 2008, **36**: Database: D480-484.
- Duncan RL, Turner CH: **Mechanotransduction and the functional response of bone to mechanical strain.** *Calcif Tissue Int* 1995, **57**(5):344-358.
- Li J, Duncan RL, Burr DB, Gattone VH, Turner CH: **Parathyroid hormone enhances mechanically induced bone formation, possibly involving L-type voltage-sensitive calcium channels.** *Endocrinology* 2003, **144**(4):1226-1233.
- Li J, Duncan RL, Burr DB, Turner CH: **L-type calcium channels mediate mechanically induced bone formation in vivo.** *J Bone Miner Res* 2002, **17**(10):1795-1800.
- Verborgt O, Tatton NA, Majeska RJ, Schaffler MB: **Spatial distribution of Bax and Bcl-2 in osteocytes after bone fatigue: complementary roles in bone remodeling regulation?** *J Bone Miner Res* 2002, **17**(5):907-914.



38. Lindqvist A, Kallstrom H, Karlsson Rosenthal C: **Characterisation of Cdc25B localisation and nuclear export during the cell cycle and in response to stress.** *J Cell Sci* 2004, **117**(Pt 21):4979-4990.
39. Henning D, So RB, Jin R, Lau LF, Valdez BC: **Silencing of RNA helicase II/ Gualpha inhibits mammalian ribosomal RNA production.** *J Biol Chem* 2003, **278**(52):52307-52314.
40. Liu JJ, Huang BH, Zhang J, Carson DD, Hooi SC: **Repression of HIP/RPL29 expression induces differentiation in colon cancer cells.** *J Cell Physiol* 2006, **207**(2):287-292.
41. Garthwaite J, Charles SL, Chess-Williams R: **Endothelium-derived relaxing factor release on activation of NMDA receptors suggests role as intercellular messenger in the brain.** *Nature* 1988, **336**(6197):385-388.
42. Pitsillides AA, Rawlinson SC, Suswillo RF, Bourrin S, Zaman G, Lanyon LE: **Mechanical strain-induced NO production by bone cells: a possible role in adaptive bone (re)modeling?** *Faseb J* 1995, **9**(15):1614-1622.
43. Turner CH, Takano Y, Owan I, Murrell GA: **Nitric oxide inhibitor L-NAME suppresses mechanically induced bone formation in rats.** *Am J Physiol* 1996, **270**(4 Pt 1):E634-639.
44. Morrison DK: **The Raf-1 kinase as a transducer of mitogenic signals.** *Cancer Cells* 1990, **2**(12):377-382.
45. Chen J, Fujii K, Zhang L, Roberts T, Fu H: **Raf-1 promotes cell survival by antagonizing apoptosis signal-regulating kinase 1 through a MEK-ERK independent mechanism.** *Proc Natl Acad Sci USA* 2001, **98**(14):7783-7788.
46. Susin SA, Lorenzo HK, Zamzami N, Marzo I, Snow BE, Brothers GM, Mangion J, Jacotot E, Costantini P, Loeffler M, et al: **Molecular characterization of mitochondrial apoptosis-inducing factor.** *Nature* 1999, **397**(6718):441-446.
47. Follet H, Li J, Phipps RJ, Hui S, Condon K, Burr DB: **Risedronate and alendronate suppress osteocyte apoptosis following cyclic fatigue loading.** *Bone* 2007, **40**(4):1172-1177.
48. Togari A, Mogi M, Arai M, Yamamoto S, Koshihara Y: **Expression of mRNA for axon guidance molecules, such as semaphorin-III, netrins and neurotrophins, in human osteoblasts and osteoclasts.** *Brain Res* 2000, **878**(1-2):204-209.
49. Hughes AR, Horstman DA, Takemura H, Putney JW Jr: **Inositol phosphate metabolism and signal transduction.** *Am Rev Respir Dis* 1990, **141**(3 Pt 2):S115-118.
50. Kennea NL, Stratou C, Naparus A, Fisk NM, Mehmet H: **Functional intrinsic and extrinsic apoptotic pathways in human fetal mesenchymal stem cells.** *Cell Death Differ* 2005, **12**(11):1439-1441.
51. Miller SA, Brown AJ, Farach-Carson MC, Kirn-Safran CB: **HIP/RPL29 down-regulation accompanies terminal chondrocyte differentiation.** *Differentiation* 2003, **71**(6):322-336.
52. Halloran MC, Severance SM, Yee CS, Gemza DL, Raper JA, Kuwada JY: **Analysis of a zebrafish semaphorin reveals potential functions in vivo.** *Dev Dyn* 1999, **214**(1):13-25.
53. Pereira-Leal JB, Seabra MC: **Evolution of the Rab family of small GTP-binding proteins.** *J Mol Biol* 2001, **313**(4):889-901.
54. Cox AD, Der CJ: **The dark side of Ras: regulation of apoptosis.** *Oncogene* 2003, **22**(56):8999-9006.

doi:10.1186/1752-0509-5-S3-S7

**Cite this article as:** Li et al.: An integrated proteomics analysis of bone tissues in response to mechanical stimulation. *BMC Systems Biology* 2011 **5**(Suppl 3):S7.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

