EÖTVÖS LORÁND UNIVERSITY

FACULTY OF SCIENSE

DEPARTMENT OF GENETICS

Identifying novel genes involved in both deer physiological and human pathological osteoporosis

theses of PhD dissertation

Adrienn Éva Borsy

DOCTORATE SCHOOL IN BIOLOGY

Head of the school: Dr. Anna Erdei, MHAS

CLASSICAL AND MOLECULAR GENETICS DhD PROGRAM

Head of the program and Supervisor: Dr. László Orosz, MHAS

BUDAPEST

2009

Introduction

Osteoporosis

Osteoporosis is one of the most common metabolic bone diseases, attacking mainly women in the industrialized countries. It is characterized by reduced bone mass, decreased amount of normally mineralized bone and microarchitectural deterioration of bone tissues, resulting in an increased risk of fracture.

It is widely established that genetic factors are responsible for the acceleration of bone loss up to 60-80%, leaving the remaining 20-40% to environmental factors. Therefore we aim to reveal genetic determinants underlying bone metabolism. Various approaches have been used to identify genes whose malfunction leads to compromised bone metabolism, including linkage analyses as well as studies of allele polymorphisms, multivariate statistical analyses of expression data and model animals.

The most preferable and best established animal model organisms of osteoporosis are mainly mice, though occasional reports on cats, dogs, ewes, pigs, hens and non-human primates. However, none of these model animals suffering from "the hidden epidemic disease" is able to regenerate their porous bone in a natural way as efficiently as deer stags.

The model system is the cyclic physiological osteoporosis of red deer

Antler development in the *Cervinae* subfamily represents the most robust bone deposition in the animal kingdom. Since the growth rate may reach over 100 g per day between May and July, enormous bone mass – generally 7-9 kg, occasionally 13-15 kg, as reported from the Danube-Drava-Gemenc-Bilje National Park in Hungary and Croatia – develops within 100-120 days. The antler's bone mass can reach up to 25-30% of the skeletal mass, which is estimated to be 10% of the live weight, about 30 kg in our case. In the period of intense antler growth, the demand for mineral precursors exceeds the dietary intake by browsing. As a consequence, the gap is filled by mobilizing minerals from the skeleton, thus leading to a temporary bone defect termed cyclic physiological osteoporosis. Before the rutting season, during the fitness recovery period in July and August, the process reverses and bone mineral density (BMD) is restored. In mule deer, mineral resorption is highest in the ribs, reaching 23% during the middle period of antler growth, and falling to less than 3% by

the time antler growth is completed. Full regeneration of skeletal elements around the completion of antler development (velvet shedding) is the unique feature of this phenomenon that never occurs in human osteoporotic bone loss.

Goal

Our hypothesis is that the genetic networks underlying human age-related osteoporosis and physiological osteoporosis of deer stag substantially overlap. Therefore, we performed comparative genetic analyses of deer and human to identify novel genes and genetic pathways which are involved in the development of osteoporosis in the skeleton during the antler cycle of deer and in human bone tissues.

Materials and methods

Tissue Sampling

Rib Bone Samples of Red Deer Stags: Approximately 2-3 g flying rib bone pieces in the entire cross section of bony rib were surgically removed from 3 anaesthetized 6, 7 and 8 year old *C. elaphus* stags. Sampling was made 3 times during one antler cycle at the Deer Farm of the Pannonian Equestrian Academy, Bőszénfa, Hungary. The time of tissue collections was (i) within the period of the active mineralization of antler, at the beginning of June when skeletal osteoporosis takes place, (ii) in the fitness improvement period with velvet shedding in late July (the „regenerating time") and (iii) in the period of late autumn dwell at the end of November when in the skeleton the mineral mobilization and deposition are dynamically equilibrated (BMD is in steady state).

Human Bone Tissue Samples: Gene expression profiles in bone samples were determined in 7 postmenopausal, unrelated, consecutive, Hungarian, Caucasian women suffering from age-related osteoporosis (PP group). The control group included 10 bone tissue samples from postmenopausal non-osteoporotic, healthy women (PNP group).

Direct polyA-RNA Isolation and Amplification

mRNA was isolated from human and red deer bone samples in order to apply in microarray hybridization experiment and Relative Quantitative Real-Time PCR. Before microarray hybridization, the red deer mRNA samples were linearly amplified in a single-round *in vitro* transcription method.

Microarray Construction

Linearly amplified mRNA of the osteoporotic, regenerating or late autumn dwell deer rib samples were reverse transcribed and hybridized onto the Human A 20K and Human B 20K standard cDNA microarrays. Hybridization procedure was carried out under the very strict, standard human-human hybridization conditions. As a result, we obtained genes by replicates matrix of size 40 000 by 3 for the comparison of osteoporotic status *versus* late

autumn dwell, for the regenerating status *versus* late autumn dwell and for the osteoporotic *versus* regenerating status.

Relative Quantitative Real-Time RT-PCR

Deer microarray data were not evaluated here, because they only served the purpose of selecting candidate genes from humans. So human mRNA was reverse transcribed and cDNA for 25 human orthologs of identified and selected deer genes was amplified by Realtive Quantitative Real-Time PCR.

Univariate and Multivariate Data Analyses

Gene expression data were evaluated by univariate and multivariate methods in order to see genetic differences between the two groups of patients and to reveal similarity structure among genes.

Mann-Whitney test: The two groups of patients were compared for each gene using the non-parametric Mann-Whitney test. $p \leq 0.1$ were considered statistically significant.

Principal components analysis (PCA): Mann-Whitney tests cannot fully recover the information hidden in the data, and more exhaustive multivariate procedures are called for. PCA is a widely used technique to summarize multidimensional data structure in terms of a few important and uncorrelated dimensions called the components. This simultaneous representation allows for the evaluation of the grouping of patients and the assessment of the relative importance and correlations of genes in influencing this configuration.

Canonical Variates Analysis (CVA): Whereas PCA explores the total variance in the data, CVA or discriminant analysis provides axes in order to maximize separation of a priori defined groups of observations examining whether the two groups overlap on the canonical axis or not provides equally useful information.

Results and Discussion

Osteoporosis as a phenotype appeared along the evolutionary lineages of both species, deer and human. An extremely intensive form termed the cyclic physiological osteoporosis developed in red deer stag associated to antlerogenesis. In contrast, humans are affected by a slowly manifesting, pathological, age-related osteoporosis that often develops after the reproductive period. At the level of orthologous genes, human and red deer are highly similar. We assume that the genetic networks underlying these two types of osteoporosis should share many common genes.

High sequence homology between orthologous genes of deer and human provide reliable option to detect expression differences in interspecific microarray setups. We probed pair wise deer rib mRNA pools against the template of Human A 20K and Human B 20K standard cDNA microarrays. Three comparisons were made:

(i) Osteoporotic status *versus* late autumn steady state,

(ii) osteoporotic *versus* regenerating status and

(iii) regenerating status *versus* late autumn steady state.

A total of 167 differential gene expression changes were recorded at or above the twofold threshold. These represented 138 different genes, 29 of them appearing in more than one comparison. The fold changes distributed as 22 genes were upregulated and 27 downregulated in comparison ii, 57 genes were upregulated and 2 downregulated in comparison i, and 3 were upregulated and 56 downregulated in comparison iii.

Based on the differentially expressed deer genes we selected 25 human orthologous genes for further investigation in order to compare the groups of age-related osteoporotic and non-osteoporotic patients. Before the statistical analyses we predefined several sortiments of the selected genes based on their functional relationships. For human osteoporosis-reference, 10 genes – *ALPL, BGLAP, BGN, COL1A1, ESR1, FN1, MGP, SPARC, SPP1* and *VDR* – were also selected.

The expression intensity of human reference and selected „deer" genes was compared by Relative Quantitative Real-Time RT-PCR analysis in bone tissue of postmenopausal, age-

related osteoporotic and postmenopausal, non-osteoporotic patients then subjected to a series of univariate and multivariate statistical evaluations.

The Mann-Whitney test detected 6 human orthologous genes, *COL1A1* (0.06), *IGSF4* (0.03), *FABP3* (0.07), *FABP4* (0.06), *TIMP2* (0.07) and *TRIB2* (0.07), which were coupled in deer to physiological osteoporosis ($p \leq 0.1$). Among the human reference genes, *ALPL* (0.02), *BGN* (0.09) and *COL1A1* (0.06) showed significantly different expression in the two groups of patients.

To understand more fully the relationships between the activity and the phenotypic manifestation of genes in osteoporosis, we applied standardized principal components analysis (PCA) to various gene sortiments. The group of age-related osteoporotic women was found to be fairly homogeneous and characterized by decreased expression activities of almost all genes investigated here.

Among the Wingless genes, the low density lipoprotein receptor-related protein 4 (LRP4) deserves particular attention. LRP4 is a potential negative regulator of the Wnt pathway. Our PCA results support this hypothesis, showing remarkably strong correlation between the expression of *LRP4* and Wnt genes, especially *CTNNB1*.

We analyzed the data set by canonical variates analysis (CVA, alias Discriminant Analysis). The analysis of the expression pattern of gene sortiments (human, deer genes and combined) detected unambiguously that deer genes have remarkably high diagnostic value for differentiating between age-related osteoporotic and non-osteoporotic states. We demonstrated the diagnostic power of *CTNNB1, LRP4, NLK, TCF7L2,* and *WIF1* expression. Inclusion of Wnt genes improved significantly the discriminating power in combination with either "deer" or human genes, although if used alone these discriminated poorly between the two human groups. Especially clear cut separation resulted for the combination of Wnt and 10 human reference genes.

Discriminant analysis was also performed to evaluate each of the possible 126 combinatorial gene sortiments from 10 human reference and 5 deer genes. Segregation was especially sharp between osteoporotic and non-osteoporotic patients in 6 out of 126 combinatorial gene sortiments. Extremely high discriminating power was observed for a sortiment in which human reference genes were combined with *CKB*, *EIF3S4*, *FKBP2*, *OSTF1*, *SFRS7* deer genes.

It deserves particular attention that *FKBP2* was involved in all of the 6 best combinatorial gene sortiments. *TMSB4X* was also included in 5 combinations. Immunophilin FK506-

binding protein 2 (*FKBP2*) and X chromosomal thymosin beta-4 (*TMSB4X*) have not yet been known to be involved in bone metabolism and development. However, it is worth mentioning that a few members of immunophilins are able to mediate the inhibitory effects of immunosuppressant cyclosporin A and FK506 on calcium-dependent signaling pathways, causing bone loss.

In biomedical research, comparative genomics serves as a compass to search for genetic markers, to develop diagnostics, and to set pharmaceutical targets.

Our results demonstrate that studying a human pathological phenomenon, disease by using a healthy animal model in comparative genomics provides a powerful approach to broaden our knowledge about the genetic background of a disease.

The sensitive bio-statistical methods – previously used to solve ecological questions – applied here may open an innovative approach in molecular biological, especially gene expression data analysis.

Multivariate statistical analyses highlighted genes and defined gene sets that are potential pharmaceutical targets for osteoporosis research and are suggested to be included in an osteoporosis diagnostical marker set. The exact role of these genes in bone biology needs further clarification.

Even the possibility emerged, that osteoporosis might develop in two independent ways, due to a defect in either the BMP-Hedgehog or Wingless signaling.

Summary

Cyclic physiological osteoporosis is a consequence of the annual antler cycle. In the period of intense antler growth, the demand for calcium and phosphate is partially filled by mobilizing minerals from the skeleton. This phenomenon raises the possibility to identify genes involved in the regulation of bone mineral density on the basis of comparative genomics between deer and human.

We compare gene expression activity of osteoporotic and regenerating rib bone samples *versus* autumn dwell control in red deer by microarray hybridization. Identified genes were tested on human femoral bone tissue from non-osteoporotic controls and patients affected with age-related osteoporosis. Expression data were evaluated by Principal Components Analysis and Canonical Variates Analysis.

Separation of patients into a normal and an affected group based on 10 formerly known osteoporosis reference genes was significantly improved by expanding the data with newly identified genes. These genes include *IGSF4*, *FABP3*, *FABP4*, *FKBP2*, *TIMP2*, *TMSB4X*, *TRIB2* and elements of the Wnt signaling. Moreover, our data strengthened the hypothesis that *LRP4* is a member of the Wnt signaling pathway; even raised the possibility that osteoporosis might develop in two independent ways, namely, due to a defect in either the BMP/Hedgehog or Wingless signaling.

This study supports that extensive comparative genomic analyses, in our case healthy deer and diseased human, provide a novel approach to identify genetic markers. The highly sensitive biostatistical methods applied here may open an innovative way in gene expression data analysis. Genes identified by the combination of comparative genomics and statistics in this study are potential diagnostical and pharmaceutical targets for osteoporosis prevention and treatment. The exact function of these factors in bone biology needs further clarification.