INTRODUCTION: Osteoarthritis (OA) is a debilitating joint disease that may occur in up to 9.4% of the adult population. Currently, there are no effective disease-modifying therapies available to treat this condition. In addition, although there has been significant effort applied to identification of biochemical markers to help diagnose and monitor disease progression, a clearly validated marker waits to be identified. Analysis of differential gene expression between non-diseased and OA cartilage is one approach to identify genes that may be used as candidate markers for disease diagnosis and/or robust diagnostic and prognostic tools. Over the past twenty years, significant effort has been applied to identification of articular cartilage matrix components. The majority of these studies have focused on the biochemical identity of these components upon extraction from articular cartilage. The difficulty with this approach is that proteins which are less abundant, or which have relatively short half-lives within the tissue, may be missed through biochemical characterization. An alternative approach is to identify expressed genes through high throughput sequencing of cDNA libraries prepared from non-diseased and diseased human cartilage. This approach has been taken to identify 1400 genes expressed in mouse growth plate cartilage (1). Due to the relatively low cellularity of adult human articular cartilage, it has been rather difficult to obtain and generate enough high quality mRNA for preparation of cDNA libraries. Consequently, while a tremendous amount of sequence data exists for other tissues in public gene databases, sequence data related to bone and especially cartilage tissue are very limited. In this report we have prepared and analyzed 5000 ESTs expressed in normal and osteoarthritic cartilage cDNA libraries.

METHODS: Human OA cartilage was obtained with informed consent and with the approval of the institutional review board from patients undergoing knee replacement surgery for OA over a period of three years. Control tissue was obtained at the time of autopsy through the Anatomical Gift Foundation. Cartilage was characterized histologically as control tissue (Table 2) and confirm the authenticity of the tissue used for the preparation and sequencing of cDNA libraries from normal and osteoarthritic human cartilage cDNA libraries.

RESULTS AND DISCUSSION: Out of 5226 ESTs sequenced from the normal cartilage library, 4415 (85%) were grouped into 2819 assemblies. A total of 1110 (40%) assemblies were unique to normal cartilage (Table 1). Similarly, 5023 ESTs were sequenced from the osteoarthritic library and 4085 were grouped into 2377 (81%) assemblies with 666 (28%) assemblies being unique to OA cartilage (Table 1). Overall 56-66% of the assemblies code for known genes. Collectively, our data indicate that about 5% of the known assemblies code for the genes reported to be highly expressed in cartilage tissue (Table 2) and confirm the authenticity of the tissue used for the preparation of the cDNA libraries. In addition to housekeeping genes, the known genes include aggrecan, COMP, and stromelysin. A small number of assemblies code for genes that are well characterized but whose expression has not yet been localized to cartilage (Table 2). Expression of several of these genes have been confirmed using Northern blotting and/or in situ hybridization. Approximately 28-40% of assemblies from both libraries are unknown. We have isolated full-length clones for some of the unknown ESTs and have partially characterized them (Table 2).

Table 1. Analysis of ESTs generated from high throughput sequencing of normal and osteoarthritic human cartilage cDNA libraries

Table 2. Partial list of genes identified by sequence analysis of ESTs from normal and osteoarthritic human cartilage cDNA libraries

REFERENCES:
3) Kumar et al., 1999, J. Biol. Chem. 274:17123

E-# - All authors are employees of SmithKline Beecham Pharmaceuticals.