

## Survey of transcripts in the adult *Drosophila* brain

### ABSTRACT

By examining this unique brain library in *Drosophila* we believe that the number of genes in this organism may be underestimated, and that this information complements the *Drosophila* genome project by allowing for more complete annotation of the genomic sequence. This library is expected to provide valuable information about how basic brain functions work at the molecular level.

### INTRODUCTION

A chronic debilitating autoimmune disease called Rheumatoid arthritis causes synovitis (synovial swelling) of the joints, bursae and tendon sheaths, but is not often manifested as a synovial disease; instead, immunologic processes appear to mediate these systemic and articular manifestations. The presence of synovial abnormalities in rheumatoid arthritis is marked by synovial hyperplasia, angiogenesis, and inflammatory cell infiltration involving the myeloid, macrophage, or lymphoid lineages. There has been considerable debate over the relevance of the specific cell types (and their products) to inflammation in rheumatoid arthritis, but it seems likely that all these cell kinds play some role in disease pathogenesis. Rheumatoid arthritis is supported by the identification of certain types of T cells in blood and synovial tissue that express specific surface membrane proteins or express a limited range of antigen receptors. To exemplify, clonal amplifications of CD8<sup>+</sup> CD57<sup>+</sup> T cells are frequently found in the repertoire of T-cells from rheumatoid arthritis patients. Similarly, CD4<sup>+</sup>CD28<sup>-</sup> T cell replications occur in both the blood and synovial compartments or tumor sites of these patients and appear to be auto-reactive. Ultimately, the T-cell receptors for antigen expressed by these and other T-cell subsets often display bias in favor of receptor that expresses some V gene expression (in contrast to the detailed studies of T cell clonality in rheumatoid arthritis), but much remains unclear about the extent to which B-cell diversity varies in this disease. The interpretation that the B-cell repertoire is also restricted is supported by previous studies. Circulating B cells were found to have oligoclonality according to flow cytometry, and cell culture experiments revealed the spontaneous secretion of immunoglobulins of restricted heterogeneity by synovial tissue explants as defined by immunological subclass, isoelectric focusing, or idiotype expression. Molecular studies of immunoglobulin genes found in the synovial tissue of rheumatoid arthritis patients provide new evidence for these concepts. These findings are significant because they reveal an ongoing immune response that targets specific (auto)antigens at low and high levels of B- and T-cells. By using the length of the third complementarity determining region (CDR3) of this rearranged copy of an enzyme called HCDR3, we were able to study the level of genetic alterations in B cells affected by rheumatoid arthritis through our current investigation. The immunoglobulin VH gene fingerprinting method, which has been used to identify the diversity of B cells and T cells in multiple clinical trials, was modified to address this problem. Recent findings indicate that B-cell clonal expansion is a typical and widespread feature of rheumatoid arthritis, that it includes both non-responsive and activated cells, and that this can persist for several months. These results provide support for the notion that autoimmune responses are involved in underlying chronic (auto)immune responses.

## CONCLUSION

In order to avoid basic sequence contingency loci of limited epidemiological value, we confined our study of tandem repeats to minisatellites, which are repeat units longer than 9 base-pairs, and to make typing alleles easier with agarose gel electrophoresis. Simple sequence contingency loci are also included in the database, which is particularly useful for molecular pathogenicity studies. The tandem repeats database was used to test the use of *Y. pestis* and *B. anthracis* on two of the most genetically homogeneous human pathogens. It is possible that a common database format will be established to identify and conduct epidemiological analyses of small-scale pathogen types that can be characterized by minisatellite typing, with additional data on tandem repeat polymorphism added to the current database to prevent duplication of work and nomenclature. The abundance and use of tandem repeats vary greatly among different bacterial species, with some having an excess of more than three tandem repetitions in their genome, such as *M. tuberculosis* and *P. aeruginosa*. The occurrence of polymorphism in tandem repeats is likely to affect protein structure rather than gene activity. In *M. tuberculosis*, all tandem repeats with total length (L) above 100 bp and 9 or 15 base-pairs long units are located with ORFs. This means that a significant proportion of these tandem repetitions correspond to the PE and PPE multigene families. Tandem repeat polymorphism is strongly associated with one or more sequenced allele characteristics in the two species examined, as shown in Figure 7. In *Yersinia pestis*, a strong correlation exists between the number of alleles observed and the homogeneity of the tandem array. In *Bacillus anthracis* (pictured) the strongest correlations are found with total array length and GC content. The two species do not exhibit comparable correlations, implying that the tandem repeat's polymorphism cannot be deduced from its primary sequence. Minisatellites, like those in the human genome are known for their high polymorphism, despite having low internal homogeneity. In order to dispel any suggestion that polymorphism is linked to subclasses of alleles displaying a higher level of internal homogeneity, more comprehensive allele sequencing will be necessary. Likewise, allele sequencing would be required to officially prove that variations in allele size are likely due to differences in the number of repeats. Of particular importance, five among the *B. anthracis* markers listed (Ceb-Bams1, 3, 7, 13, and 30) are highly polymorphic with PIC values (or Nei's index) greater than 0.7. It is worth noting, however, that the allele length observed for Ceb-Bams1 in the Ames strain is not typical according to the sequence data (Table 2). This may be due to a high mutation rate at Ceb-Bams1/3 and subsequently to some sequencing error. The Ames strain is unlikely to have allele 4 (Table 3), as the expected allele size does not correspond to any existing allele in the product set. The locus is moderately polymorphic, with a PIC value of 0.26 and only three alleles present in Table 2, which suggests that interpreting it may result in ambiguity among sequencing researchers. This problem could be remedied by using Ceb-Bams1 and Ceb-Bean28, the same strain used for the sequencing project. The phylogenetic tree illustrated in Figure 6 tends to cluster strains with alleles of similar size at the most variable loci, despite the lack of consideration for the magnitude of allele size difference in building the distance matrix. This is comparable to observations made in *H. influenzae*. It is suggested by *influenzae* and others that mutation events are primarily small-scale changes. To better understand the process of event generation, it is necessary to conduct more detailed studies using full allele sequencing.