

SUPPLEMENTAL METHODS

Directly or indirectly, retinoic acid regulates at least 580 genes (*ibid.*, updated). The literature contains references to 220 genomic elements thought to be associated with this regulation—fully tested receptor (dimer) binding sites, unparsed “active” regions, untested motifs, and the like. Once inappropriate sites were eliminated for a variety of reasons—and once orthologous redundancies and sites in transcribed regions were removed—a list of 77 unique binding sites in noncoding, nontranscribed DNA remained. Each was from one of three target species: mouse (Mm), rat (Rn), or man (Hs). Supplemental Table 1 lists all elements of the starting set, together with literature references, GenBank identifiers, and the reasons for occasional eliminations. Note, however, that site conservation was never a consideration.

We located each binding site in a native context of about 1000 bases and used BLASTn to search for homologous regions in the other two species. We multiply aligned the resulting homologs using MACAW (<ftp://ftp.ncbi.nih.gov/pub/schuler/macaw/>), extending the sequences with BLASTn as necessary. There are no standard criteria of significance in short, noncoding blocks. However, TF binding sites are generally between 6 and 20 bases in length; and clustering, if present, would impose some limit on runs of intra-modular mismatches (or indels). We therefore worked up- and downstream from each binding string (conserved or not) until we reached a non-aligning zone of at least 60 nucleotides that did not overlap an aligning substring of at least nine nucleotides containing at least six matches. This method is outside the domain of traditional “significance”, but it was applied within larger homology blocks that *do* rank as significant in the larger search spaces of our original BLASTn hits (see Karlin & Altschul* for MACAW and BLAST significance).

* Karlin, S. & Altschul, S. F. 1990. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc.Natl.Acad.Sci.U.S.A.*, **87**, 2264-2268.