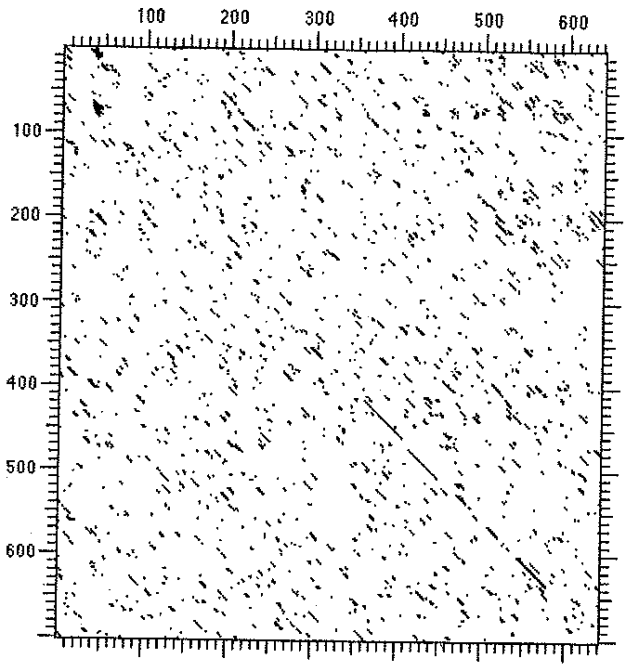

Figure 3.4. Dot matrix analysis of DNA sequences encoding phage λ *cI* (*vertical sequence*) and phage P22 *c2* (*horizontal sequence*) repressors. This analysis was performed using the dot matrix display of the Macintosh DNA sequence analysis program DNA Strider, vers. 1.3. The window size was 11 and the stringency 7, meaning that a dot is printed at a matrix position only if 7 out of the next 11 positions in the sequences are identical.

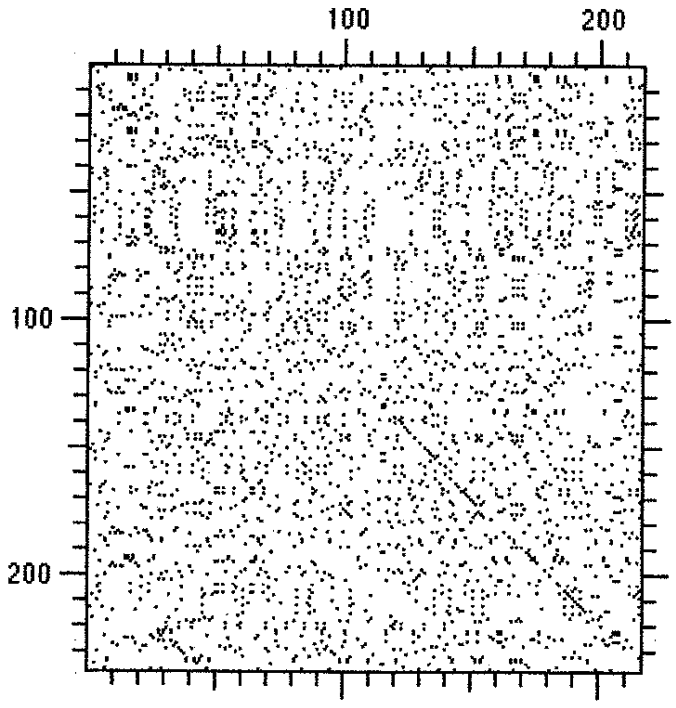

Figure 3.5. Dot matrix analysis of the amino acid sequences of the phage λ *cI* (*horizontal sequence*) and phage P22 *c2* (*vertical sequence*) repressors performed as described in Fig. 3.4. The window size and stringency were both 1.

from: David M. Mount, Bioinformatics (2004), pp. 60-61

David M. Mount:
Bioinformatics, 2004

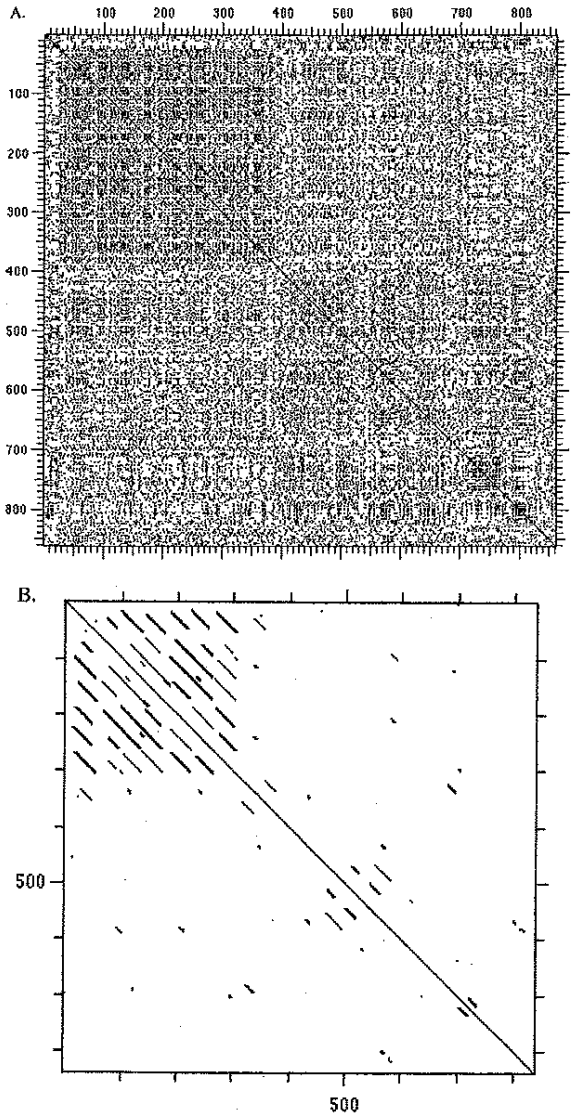

Figure 3.6. Dot matrix analysis of the human LDL receptor against itself using DNA Strider, vers. 1.3, on a Macintosh computer. (*A*) Window 1, Stringency 1. There is a diagonal line from upper left to lower right due to the fact that the same sequence is being compared to itself. The rest of the graph is symmetrical about this line. Other (quite hard to see) lines on either side of this diagonal are also present. These lines indicate repeated sequences perhaps 50 or so long. Patches of high-density dots, e.g., at the position corresponding to position 800 in both sequences representing short repeats of the same amino acid, are also seen. (*B*) Window 23, Stringency 7. The occurrence of longer repeats may be found by using this sliding window. In this example, a dot is placed on the graph at a given position only if 7/23 of the residues are the same. These choices are arbitrary and several combinations may need to be tried. Many repeats are seen in the first 300 positions. A pattern of approximate length 20 and at position 30 is repeated at least six times at positions 70, 100, 140, 180, 230, and 270. Two longer, overlapping repeats of length 70 are also found in this same region starting at positions 70 and 100, and repeated at position 200. Since few of these diagonals remain in new analyses at 11/23 (stringency/window) and all disappear at 15/23, they are not repeats of exactly the same sequence but they do represent an average of about 7/23 matches with no deletions or insertions. The information from the above dot matrix may be used as a basis for listing the actual amino acid repeats themselves by one of the other methods for sequence alignment described below.