DNA Skyline: fonts to facilitate visual inspection of nucleic acid sequences

Jonas Jarvius and Ulf Landegren
Uppsala University, Uppsala, Sweden

In a time of explosive increase of genomic information, there is a need for improved DNA literacy. Despite the availability of computer programs to evaluate DNA and RNA sequences, the human eye can still play a vital role in sequence analysis for patterns and motifs. We describe a set of fonts that represent DNA sequences in an intuitive manner, suitable for teaching and for visual inspection of strand complementarity, conserved motifs,

and oligonucleotide probes. Similar notations to aid visual inspection of nucleic acid sequences have been described earlier (1–3), but they typically require dedicated software in order to operate, thereby reducing their simplicity. The fonts we have developed represent nucleotides in a manner similar to musical notes, with each nucleotide extending to one of four levels (see Figure 1). Gs are placed on top, followed by As and Ts, with Cs at the lowest level. The two nucleotides that form base pairs via three hydrogen bonds, G and C, occupy the extreme positions, while A and T, held together by two hydrogen bonds, take up the innermost two. Consecutive nucleotides are connected by an arrow at the top level in all positions, displaying the 5′ to 3′ orientation. This GATC font is designed so that complementary sequences exhibit a precise fit, and a variant CTAG font is also provided to represent the complement of a selected sequence. The fonts have proven useful in our own work on DNA probe assays [e.g., involving padlock (4), selector (5), and proximity (6) probes]. Once installed on a computer by dragging the font files into the local font directory, DNA sequence information represented in any text/image software can be easily viewed

in DNA Skyline by selecting one of the fonts. DNA Skyline is a TrueType font set, making this notation platform independent, simple to use, and compatible with most computer programs that use fonts. The fonts and a user guide can be downloaded at www.genpat.uu.se/Skyline.html.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

1. **Cowin, J.E., C.H. Jellis, and D. Rickwood.** 1986. A new method of representing DNA sequences which combines ease of visual analysis with machine readability. Nucleic Acids Res. *14*:509-515.
2. **Schneider, T.D. and R.M. Stephens.** 1990. Sequence logos: a new way to display consensus sequences. Nucleic Acids Res. *18*:6097-6100.
3. **Zimmerman, P.A., M.L. Spell, J. Rawls, and T.R. Unnasch.** 1991. Transformation of DNA sequence data into geometric symbols. BioTechniques *11*:50-52.
4. **Nilsson, M., H. Malmgren, M. Samiotaki, M. Kwiatkowski, B.P. Chowdhary, and U. Landegren.** 1994. Padlock probes: circularizing oligonucleotides for localized DNA detection. Science *265*:2085-2088.
5. **Dahl, F., M. Gullberg, J. Stenberg, U. Landegren, and M. Nilsson.** 2005. Multiplex amplification enabled by selective circularization of large sets of genomic DNA fragments. Nucleic Acids Res. *33*:e71.
6. **Fredriksson, S., M. Gullberg, J. Jarvius, C. Olsson, K. Pietras, S.M. Gustafsdottir, A. Ostman, and U. Landegren.** 2002. Protein detection using proximity-dependent DNA ligation assays. Nat. Biotechnol. *20*:473-477.

Address correspondence to Jonas Jarvius, Uppsala University, Department of Genetics and Pathology, Rudbeck Laboratory, SE-751 85 Uppsala, Sweden. e-mail: jonas.jarvius@genpat.uu.se
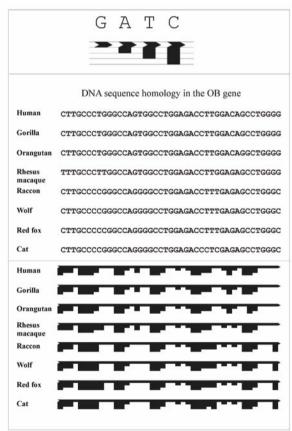
Figure 1. Sequence homology in the OB gene displayed using the Skyline "GATC" notation. The top panel illustrates the design of the Skyline GATC font. In the middle panel, sequences from the obesity gene in different mammalian species are displayed using the Courier font, and in the lower panel, the same nucleotides are displayed using the GATC version of the Skyline font.