Alu alert

SIR — The presence of Alu elements in protein coding regions has been a controversial issue for some time[1] and is periodically reactivated[2–4].

The polymorphic short interspersed elements (approximately 300 base pairs long) of the Alu family are ubiquitous in the genome of primates. They constitute about 5 per cent of the human genome, but can also be found in dense clusters of up to two Alu elements per kilobase[5]. In relation to regular protein transcription units, Alu elements are often found in introns, or in 3′ untranslated regions. A single base deletion or insertion while sequencing a genomic clone or a complementary DNA can easily cause Alu regions to be mistakenly incorporated into open reading frames[1]. Also, cDNA libraries can contain partially spliced mRNA, be contaminated with genomic material, or contain rearranged clones from artefactual ligation and recombination.

In such rearranged cDNAs, Alu elements will probably be inserted within bona fide exon sequences. A few Alu elements are still transcriptionally active, and are thought to be spread within the genome through retrotransposition. This suggests a mechanism by which Alu elements might have played an important role in the recent evolution of certain primate proteins[6]; but the sequence data supporting such a hypothesis are difficult to distinguish from the many possible artefacts cited above[4]. A recent example of interpretational difficulties comes from BRCA1 (breast cancer) mRNA. One of the rare forms of this mRNA contains exon 4, composed entirely of an Alu cassette. As this was not mentioned by Miki *et al.* [7] and exon 4 was found only in one of 15 variants of BRCA1 mRNA, this particular cDNA is probably a splicing error that results in a non-functional product.

While the issue is still in debate, an increasing number of Alu-containing protein sequences, conceptually translated from cDNA or genomic open reading frames, are finding their way into data bases (we detected at least 37 entries in NR, the NCBI non-redundant database). Despite the introduction of Alu warning entries in protein databases[1], the presence of Alu-like segments often escape detection by the original authors, thus generating entries with no mention of this feature. These potentially anomalous protein sequences, then, constitute a growing source of misleading homology matches.

To help researchers to detect Alu-like segments in putative coding regions early on, a select subset of 325 Alu elements is now available for a quick search via the NCBI BLAST server (blast@ncbi.nlm.nih.gov). This subset, its conceptual translation, and the various programs of the BLAST suite allow protein–protein, nucleotide–protein and nucleotide–nucleotide comparisons. This small selection of Alu elements is highly representative and can detect 99 per cent of all known Alu elements in GenBank. In conjunction with the xblast[8] program, search against this database can be used for reliably masking all Alu-like segments from large (for example, genomic contig) or multiple (partial cDNA) queries before scanning the protein databases or GenBank. This step greatly improves the biological relevance of the output by drastically reducing reported hits; and limiting them to functionally meaningful regions of the sequences.

The 325-Alu subset, its conceptual translation and a help file are available by anonymous FTP (ncbi/pub/jmc/alu). Also 8 Alu consensus sequences, representing each of the currently defined subfamilies, have been incorporated as warning entries into GenBank (accession numbers U14567–U14574).

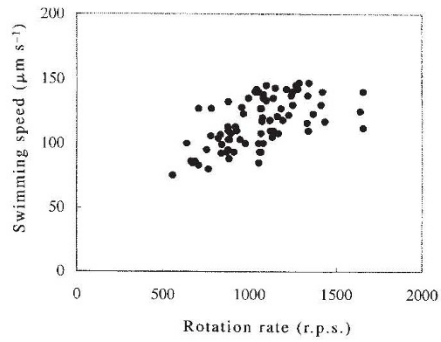
Jean-Michel Claverie*
Wojciech Makalowski†
National Center for Biotechnology Information,
National Library of Medicine, NIH,
Bethesda, Maryland 20894, USA
*(jmc@ncbi.nlm.nih.gov); †(makalow@ncbi.nlm.nih.gov)

1. Claverie, J.-M. *Genomics* **12**, 838–841 (1992).
2. Margalit, H. *et al. Cell* **78**, 173–174 (1994).
3. Tugendreich, S. *et al. Nature* **370**, 106 (1994).
4. Zietkiewicz, E. *et al. Science* **265**, 1110–1111 (1994).
5. Iris, F. J. M. *et al. Nature Genet.* **3**, 137–145 (1993).
6. Makalowski, W. *et al. Trends Genet.* **10**, 188 (1994).
7. Miki, Y. *et al. Science* **266**, 66-71 (1994).
8. Claverie, J.-M. in *Automated DNA Sequencing and Analysis* (eds Adams, M. *et al.*) 267(Academic, New York, 1994).

Very fast flagellar rotation

SIR — Bacteria swim by rotating their helical flagellar filaments. Each filament is driven with a flagellar motor embedded in the cell surface. The rotation rates of the motor had been reported to be high, for example, 270 r.p.s. for *Escherichia coli*[1] and 170 r.p.s. for *Salmonella typhimurium*[2]. We have observed remarkably faster rotation, up to 1,700 r.p.s. for polar flagellum of *Vibrio alginolyticus*.

We measured the flagellar rotation rates of individual free-swimming cells of *V. alginolyticus* as well as their swimming speeds by using laser dark-field microscopy[2]. The figure shows the result obtained at 35 °C. The average value of rotation rate was 1,100 r.p.s. and the highest value was 1,700 r.p.s. The average ratio of swimming speed to rotation rate was 0.113 μm rev^{-1}, which means that a cell of *V. alginolyticus* progresses by 7% of the pitch of its flagellar helix (1.58 μm). In addition, we observed a tendency for swimming speed to saturate at high rotation rate, which we suggest reflects both the torque versus rotation rate characteristics of flagellar motor, and the distribution in lengths of flagella (in preparation).



Swimming speed as a function of flagellar rotation rate obtained for *V. alginolyticus* mutant YM42 (smooth swimming, no lateral flagella). Each point shows data for an individual cell in the presence of 300 mM NaCl at pH 7.0.

The sliding velocity between rotor and stator of the flagellar motor rotating at 1,700 r.p.s. is estimated to be about 160 μm s^{-1}, because the diameter of the rotor is about 30 nm. Sliding velocities reported for other molecular motors are smaller than this value. For example, myosin moves along actin at a speed of about 10 μm s^{-1} in skeletal muscle and around 100 μm s^{-1} in cytoplasmic streaming of plant cells such as *Chara* and *Nitella*[3]. So the polar flagella of *V. alginolyticus* should be one of the fastest molecular motors in biological systems.

Our observation of fast flagellar rotation raises some questions about the underlying mechanisms of flagellar motors. The motor of *V. alginolyticus* is driven by Na^+ flow from the outside of the cell to the inside. If we assume that the motor requires 1,000 ions per revolution, as was reported for the H^+-driven motor of *Streptococcus*[4], the number of Na^+ ions that move into the cell through the motor per second amounts to about 5% of cytoplasmic Na^+ (concentration of 50 mM; ref. 5). Does fast rotation require such a large influx of ions? Are there any possibilities that the required number of ions per revolution is smaller than 1,000 or reduces at higher rotation rates?

Y. Magariyama, S. Sugiyama
Tsukuba Research Laboratory,
Yaskawa Electric Corporation,
5-9-10 Tokodai, Tsukaba 300-26, Japan
K. Muramoto, Y. Maekawa
I. Kawagishi, Y. Imae*
Department of Molecular Biology,
Faculty of Science, Nagoya University,
Chikusa-Ku, Nagoya 464-01, Japan
S. Kudo
PRESTO, JRDC†
*Deceased.
† C/O Yaskawa Electric Corporation.

1. Lowe, G. *et al. Nature* **324**, 637–640 (1987).
2. Kudo, S. *et al. Nature* **346**, 677–680 (1990).
3. Shimmen, T. *Bot. Mag.,Tokyo* **101**, 533–644 (1988).
4. Meister, M., Lowe, G. & Berg, H. C. *Cell* **49**, 643 (1987).
5. Yoshida, S. *et al. J. biol. Chem.* **265**, 20346 (1990).