Worlds Collide

SIGHTINGS

Virologists come in two flavors: those who study DNA viruses and those who study RNA viruses. The RNA types can be further subdivided into those who study retroviruses and those who study everything else. Basic virologists occasionally cross paths when different viral life cycles have mechanistic parallels, but the camps seem to have led rather separate existences. Nowhere can this segregation can be seen more clearly than in the rapidly evolving field of gene therapy.

Gene therapy investigators usually focus on achieving the most efficient delivery of a DNA virus or an RNA retrovirus. Adenoviruses are commonly used as DNA vectors because they can be easily engineered and grown to high titer in the laboratory. They also infect a wide variety of human cells quite easily. However, because adenoviruses do not integrate into the host genome, they will be lost after several generations, along with any passenger DNA they were designed to carry. A retrovirus, on the other hand, will integrate into the host genome and remain a permanent resident throughout the life of the target cell, expressing the passenger gene of interest. Unfortunately, it is often difficult to produce high titers of recombinant retroviruses in culture systems. Could the best of the DNA viral delivery vectors be fused in some way with the best of the retroviruses?

A recent report (*1*) answers the question with a resounding "yes." The group, led by D. Curiel in Alabama, has succeeded in designing a system that produces a virus that grows to high titers in vitro, efficiently infects dividing and quiescent cells, and integrates into the host genome.

First, Feng *et al.* took all the genes required for the packaging of a functional retrovirus (*gag*, *pol*, and *env*) and inserted them into an infectious adenovirus vector, thus making a "packaging virus." Next, they modified a second adenovirus by inserting a packageable retroviral genome containing a passenger gene, thus making a "passenger virus." For testing, they used green fluorescent protein (GFP) as their passenger gene. They then infected cells in culture with a mixture of the two engineered adenoviruses. The combination of all the various protein components made by both viruses yielded individual cells that could now express GFP from both an episomal adenovirus vector and an integrated retrovirus. Infected cells became essentially retroviral factories to produce additional infectious retroviruses in situ. These viruses could then efficiently infect neighboring cells.

Feng *et al.* also performed in vivo experiments to test the ability to target tumor cells. They injected nude mice with both a human ovarian tumor line and, intraperitoneally, the packaging and passenger viruses. The combined injection of viruses was able to produce a high proportion of GFP-positive cells in the subcutaneous tumor nodule. Thus, high infectivity coupled with extra- and intrachromosomal replication creates a new gene delivery system that works in vivo.

Applications of this new technology to humans will require a bit of redesigning to make the vector systems safer, but the concept is clearly proven. In the mouse system, the use of defective adenovirus strains will allow limited infection of injected tissues. The use of HIV-derived retroviral packaging systems may widen the overall utility as well, because HIV-based systems do not require cell division for integration.

–Robert Sikorski and Richard Peters

References

1. M. Feng *et al.*, *Nat. Biotechnol.* **15**, 866 (1997).

Digital Mailbox:
www.sciencemag.org/dmail.cgi?53372a

BLAST Off!

SIGHTINGS

Most biologists are by now aware of the Basic Local Alignment Search Tool (BLAST) program available through the Web site of the National Center for Biotechnology Information (NCBI) (*1*). In today's genomic era, BLAST is essential, providing fast and accurate search algorithms for analyzing protein and nucleic acid sequences. However, the current windfall of genomic data has created an amazing situation in which the size of protein and nucleic acid databases is growing almost faster than the processing power of computers. And, as the processing rate of the program is increased, the accuracy of the results drops. The original BLAST program, while quite accurate, can occasionally miss important similarities between sequences.

These issues have been addressed by the team at NCBI in their just-released version 2.0 of BLAST. There are two major new features in the software: gapped BLAST and PSI-BLAST. The details behind these improvements can be found in a recent article (*2*).

The BLAST alignment algorithm is based on the fact that a statistically significant alignment is likely to contain a high-scoring pair of aligned words. Such an aligned word pair is referred to as a "hit." Each hit is checked to see if it lies within an alignment by extending the initial hit in both directions. It turns out that this last step accounts for more than 90% of BLAST's execution time.

To increase the speed of BLAST, the NCBI team created gapped BLAST (www.ncbi.nlm.nih.gov/BLAST/). It uses a clever, two-hit method. Hits are extended only when two nonoverlapping hits are found occurring within a predefined distance from one another. Also, any ungapped alignment that achieves a threshold score triggers a gapped alignment extension. This new algorithm leads to a threefold increase in speed over the original BLAST. As the name implies, gapped BLAST also allows for the introduction of gaps (deletions and insertions) into alignments. With a gapped alignment tool, homologous domains are not broken into several segments when they are compared.

PSI-BLAST (which stands for Position-Specific Iterated BLAST) is a search algorithm that uses position-specific scoring matrices as the query rather than a simple sequence (www.ncbi.nlm.nih.gov/BLAST/). PSI-BLAST performs an iterative search in which sequences found in one round of searching are used to build a score model for the next round of searching. Three operations are performed: (i) constructing a multiple alignment from BLAST output data; (ii) processing this alignment into a position-specific score matrix; and (iii) using this matrix to search the database. The NCBI team modified BLAST to operate with such a matrix rather than with a simple query. PSI-BLAST runs at about the same speed as the original BLAST but is more sensitive and can detect even weak relationships among candidate sequences.

As an example of the increase in processing speed brought by these new tools, the team searched SWISS-PROT with the length-567 influenza A virus hemagglutinin precursor sequence as a query. The original BLAST program took 45.8 seconds, whereas gapped BLAST took only 15.8 seconds.

In another example, they tested the approach with the COOH-terminal domain of human BRCA1, which includes two BRCT domains. The initial BLAST search yielded significant alignments to other BRCA1 sequences and BARD, another BRCT protein. Subsequent PSI-BLAST iterations retrieved more than 30 distinct homologs. Seven of these were recent additions to the databases and had not previously been recognized as BRCT proteins.

In summary, BLAST 2.0 is a more robust, versatile, all-purpose alignment tool that can be applied to general sequence alignment problems. Gapped BLAST is faster than the original BLAST and allows for gaps in the sequence being aligned. Finally, PSI-BLAST

(continued on page 502)