[New hosts and vectors for genome cloning] DE-FG03-90ER Calif Inst of

T3

DOE/ER/60997--T3

DE92 007029

Progress Report

The main goal of our project remains the development of new bacterial hosts and vectors for the stable propogation of human DNA clones in E. coli. During the past six months of our current budget period, we have (1) continued to develop new hosts that permit the stable maintenance of unstable features of human DNA, and (2) developed a series of vectors for (a) cloning large DNA inserts, (b) assessing the frequency of human sequences that are lethal to the growth of E. coli, and (c) assessing the stability of human sequences cloned in M13 for large-scale sequencing projects.

Our work on this project over the past 13 months (10 mo budget period 01 + 3 mo budget period 02) of funding has resulted in three publications:

1) A. Greener, P. Youderian, and J. A. Sorge, (1992) Inverted repeats in Escherichia coli are DNA lesions to be repaired. J. Bacteriol., in press.
2) M. Lorincz, A. A. Kumamoto, and P. Youderian, (1992) Boil and fry: direct electroporetic transfer of plasmid DNA in the absence of a donor counterselection. Trends in Genetics, submitted.
3) A. A. Kumamoto, E. Rockenstein, M. Lorincz, and P. Youderian, (1992) The stability of human DNA inserts in a cosmid cloning vector. Genomics, in preparation.

Our current work involves:

A) Greener subcontract (host development):

1. Construction of a set of otherwise isogenic E. coli tester strains for assessing plasmid and cosmid stability.

We have constructed a set of otherwise isogenic derivatives of E. coli strain JM101 carrying mutations in each of the known ts recombination and repair pathways. These strains, for the most part, were constructed in a two-step process: a Tn10 insertion was isolated in or near the recombination/repair gene, and tetracycline-sensitive segregants were isolated and screened for deletions that result in loss of the Tn10, and for loss of the recombination/repair function. These strains carry single mutations in the recA, B, D, F, J, N, O, R, gyrA, himA, uvrC, uvrB, umuC, mcrA, mcrB, mrr, xse, mutS, mutD, and ada genes.

2. Stabilization of human DNA clones with (CA)n and (GA)n repeats.

We have constructed derivatives of the plasmid, pBluescript, carrying segments of the repeating dinucleotides $(CA)_{13}$ and $(GA)_{13}$, repeat features found frequently in human DNA sequences, and often components of highly informative RFLP's. These plasmids, unlike the parental pBluescript plasmid, are very unstable in all bacterial hosts among a constellation of otherwise isogenic E. coli strains with mutations in each of the known genes involved in recombination and repair. Nonselective growth for 20 generations results in >99% loss of the plasmid; in addition, these repeated sequences somehow cause the high-frequency loss of F replicons in the same hosts.

Because these repeats are inserted in the multicloning site of pBluescript, they result in a shift in the frame of the N-terminal α-complementing fragment of lacZ coding sequence. When these plasmids are

transformed into our set of hosts encoding the C-terminal α-complementing fragment of lacZ, the transformants form white colonies on lactose indicator plates. Loss of two bp from the repeat restores reading frame, and results in a blue colony on LB Xgal plates. In a mutS genetic background, transformants with the repeat plasmids segregate blue colonies at very high frequencies. After UV-mutagenesis of the mutS, plasmid-carrying hosts, we have obtained several independent mutant strains of E. coli that (1) reduce the frequence of repeat rearrangements, and (2) allow for the stable maintenance of the plasmid and F. We are currently mapping the mutations in 10 independent mutants that permit the stable inheritance of these plasmids. These mutants are healthy prototrophs, and show no obvious growth defects. We plan to introduce the strongest of these mutations into the SURE genetic background, and reassess the frequency of unstable human cosmid inserts in this background. We expect that these mutations will facilitate the cloning of many currently unstable human sequences in cosmids, and thereby the closure of cosmid contig maps of human chromosomes.

3. Assessing the stability of human inserts with hyphenated inverted repeats.

We have constructed tester plasmids carrying long hyphenated inverted repeats, a feature that occurs frequently in human genomic DNA. Surprisingly, we find that these plasmids are relatively stable and rearrange very slowly in the SURE genetic background, indicating that we need not be concerned with the instability of these frequent features of human DNA in the SURE host.

A) Youderian contract (vector development):

1. Vectors to assess the frequency of lethal human sequences.

We are constructing plaque-forming derivatives of Salmonella phage P22 carrying an F origin of replication and antibiotic-resistance determinant, for the cloning of human sequences potentiall lethal to the growth of E. coli. These phage vectors form turbid plaques. When a lethal sequence of DNA is cloned into one of these vectors, the recombinant phage forms clear plaques, because the lethal gene prevents the stable growth of lysogens that turn a plaque turbid. Such sequences do not impair the rapid, lytic growth of phage P22, because P22 lytic development only relies on pre-existing bacterial gene products, and occurs with an eclipse time of 20 min. We will test these "lethality" vectors in two ways. First, in collaboration with Betty Kutter at Evergreen State College in Olympia, WA, we will clone the 5% of the phage T4 genome that encodes lethal functions, and thereby close the map of phage T4. Second, we will clone random human sequences into this vector, and simply score the frequency of clones that form clear plaques as a measure of the frequency of lethal features in the human genome. Because our vector carries opposing T3 and T7 promoters flanking the site of insertion, we can generate RNA probes for lethal sequences, to assess their distribution in the human genome, as well as use PCR primers complementary to these promoter sequences to obtain STS tags for such lethal sequences.

2. Vectors for cloning large human DNA fragments.

As we continue to develop improved bacterial hosts for the stable maintenance of unstable features of human genomic DNA, we will be able to clone larger and larger fragments of human DNA without having these fragments undergo rapid rearrangements. Currently, we have two steps remaining for the construction of

vector STEALTH S117, a vector for general use in closing genomic maps. This vector has: (a) a single-copy F origin of replication, (b) two selectable antibiotic resistance genes (cat and spcA) flanking the insertion region, to selectively prevent rearrangements of an insert that delete the ends of a cloned fragment, (c) an insertion region consisting of opposing T7 and T3 promoters flanking unique NotI, SfiI, MluI, M.Cla/DpnI, and SalI (insertion) sites, to facilitate restriction mapping of inserts, and (d) the Salmonella phage P22 early region, to allow for the terminal amplification of a cloned insert.

To use this vector, large Sau3aI or MboI partial fragments of human DNA (50-300 kbp) are filled in with dGTP and dATP, the vector is filled in with dCTP and dTTP, and insert is ligated with an excess of vector. After electroporation of a library, single colonies are resuspended in rich media, and divided. Half of each colony is frozen, and the other half is grown for two more generations then amplified by temperature-induction of the P22 repressor gene. Terminally amplified DNA from these clones will represent the structure of a human insert before rearrangement, giving us a way to look at the structure of inserts prior to extensive rearrangement. End probes will permit standard walking strategies, and inserts also can be ordered with respect to a chromosome map by hybridization with STS or unique probes. The presence of rare restriction sites adjacent to the site of insertion will permit the rapid Smith-Bernstiel type mapping of inserts.

3. Vectors to assess the large-scale sequencing of human sequences.

Given that serveal pilot projects are underway to sequence megabase regions of individual human chromosomes, it is important to assess the frequency of features of human DNA that rearrange rapidly in phage M13 vectors, the vectors of choice for automated, large-scale sequencing. Using site-directed mutagenesis, we have constructed a derivative of M13mp19 that carries a unique BglII site immediately downstream of the N-terminal α-complementing fragment of lacZ in this vector. We will cut this vector with BglII, phosphatase or partially fill-in the ends, and construct libraries of small human inserts generated by Sau3a, SalI, or XhoI cleavage of human DNA. Because insertion into the BglII site does not disrupt the lacZ' gene, all clones should form blue plaques on a complementing host. Individual plaques will be picked, resuspended, grown as though for DNA sequencing, and replated on indicator plates, to measure the frequency of inserts that rearrange rapidly, and cause the deletion of adjacent lacZ' sequences (resulting in the formation of white plaques). This experiment should allow us to assess the frequency of unstable features of human DNA in M13, both for subclones from cosmids representing gene-rich regions and for clones of total genomic DNA.

Research proposed for the next budget period

The research goals for the next budget period will be essentially as proposed, and described in the Progress Report (above).

Budget justification for year 03

The budget is requested as initially proposed, with the exception that an additional month of support is requested for the PI (change from 2 mo to 3 mo support), with a commensurate decrease in the Materials and Supplies category. The PI is currently supported by one other grant from the NIH for a total of 25%

of his time, and contributes to this project 3 mo/yr by doing laboratory work on vector development at the bench.

DISCLAIMER

END

DATE
FILMED
4/24/92