A sequence tagged connector (STC) approach with a non selective AFLP fingerprinting

A protocol for tomato BAC walking

Assembly

Using the PREGAP4 interface of the Staden package 2005, raw trace data from tomato chromosome 6 BACs was processed into assembly ready sequences. Sequences were base-called by the PHRED basecaller. Clipping was performed to remove sequencing vector, cloning vector, and bad quality sequences. Processed sequences were subsequently assembled with Gap4, with a sequence percentage mismatch threshold of 8%, and parsed into the GAP4 assembly database. The gap4 contig editor interface was used for editing and finishing. Consensus calculations with a quality cut-off score of 40 were performed from within GAP4 using a probabilistic consensus algorithm based on the expected error rates output by PHRED.

Identification of minimal overlapping and maximal extending BACs for walking

To identify overlapping BAC clones for walking, we use tomato BAC end sequences from the SOL Genomics Network available at ftp://ftp.sgn. cornell.edu/tomato_genome, and perform a BlastN or MegaBlast analysis against assembled tomato contigs. Position and direction of overlap were verified, and candidate BAC clones are preselected setting a threshold expect value to 0.0 or an identity score treshold of 0.99. When meeting constraints, corresponding ABI traces are subsequently assembled onto BAC contig sequences to which the Blast hit is found, and verified at nucleotide level for integrity. Assembled BAC end sequences showing high quality basecall differences compared to contig consensus sequences, or showing its assembly start more than 50 basepairs downstream from a candidate *Hin*dIII, *Mbo*I, or *EcoRI* cloning site are rejected. Remaining candidate BAC clones are further analyzed by fingerprint analysis.

AFLP fingerprinting and BAC insert sizes

BAC DNA was isolated by standard alkaline lysis method and *Eco*RI/*Mse*I, *Hin*dIII/*Mse*I and *Pst*I/*Mse*I AFLP templates were prepared as described by Vos *et al*., 1995. 5 µl of the restriction ligation mix was diluted 10 fold in 10 mM Tris-HCl pH 7.5, 0.1 mM EDTA buffer. A non selective amplification with [γ-33]ATP labeled *Eco*RI+0 and a *Mse*I+0 primers was performed in a total volume of 20 µl. Typically a 30 sec DNA denaturing step at 94˚C, a 1 min annealing step at 56 ˚C, and a 1 min extension step at 72 ˚C for 35 cycles was performed. For the *Hin*dIII/*Mse*I and *Pst*I/*Mse*I templates respectively, the *Hin*dIII+0 and *Pst*I+0 [γ-33]ATP-labeled primers were used in combination with the *Mse*I+0-primer. All amplification reactions were performed in a PE-9700 thermocycler (Perkin Elmer). After the amplification step

electrophoretic gel analysis of the reaction mix was carried out, and the fingerprint patterns were visualized using a Fuji BAS-2000 phosphoimaging analysis system (Fuji Photo Film Company Ltd, Japan). Band sizes were calculated relatively to a 10 basepairs size ladder with AFLP-Quantar® fingerprint analysis software, and co-migrating bands were scored by visual inspection. AFLP-Quantar® fingerprint analysis software (http://www.keygene.com/technologies/technologies_keymaps.htm) is distributed by KeyGene and is not part of TOPAAS. For insert size determination BAC DNA was prepared by a standard alkaline lysis method from a 3-ml overnight culture. BAC DNA was digested with *Not*I (New England Biolabs) to completion and separated by field inversion gel electrophoresis (BioRad FIGE MAPPER™) on a 1% agarose gel in 0.5x TBE, with a linear run time, forward (3-30s) reverse (1-10s), 14 hrs and 160 Volts, along with a mid-range PFGE marker I (New England Biolabs).

This protocol is to be published in: *Peters et al*., 2006. TOPAAS, a Tomato and Potato Assembly Assistance System for selection and finishing of BACs. Accpeted for publication in *Plant Physiology.*