Copyright © 1998 by the Genetics Society of America

Mapping the Genome of a Model Protochordate. I. A Low Resolution Genetic Map Encompassing the Fusion/Histocompatibility (Fu/HC) Locus of *Botryllus schlosseri*

Anthony W. De Tomaso,* Yasunori Saito,† Katharine J. Ishizuka,* Karla J. Palmeri* and Irving L. Weissman*

**Hopkins Marine Station, Pacific Grove, California 93950 and Department of Pathology, Stanford University School of Medicine, Stanford, California 94305 and †Shimoda Marine Research Station, University of Tsukuba, Shimoda, Shizuoka, Japan*

Manuscript received October 14, 1997
Accepted for publication February 10, 1998

ABSTRACT

The colonial protochordate, *Botryllus schlosseri*, undergoes a genetically defined, natural transplantation reaction when the edges of two growing colonies interact. Peripheral blood vessels of each colony touch and will either fuse together to form a common vasculature between the colonies, or reject each other in an active blood-based inflammatory process in which the interacting vessels are cut off and the two colonies no longer interact. Previous studies have demonstrated that allorecognition in Botryllus is principally controlled by a single Mendelian locus named the fusion/histocompatibility (Fu/HC) locus, with multiple codominantly expressed alleles. However, identification and cloning of this locus has been difficult. We are taking a genomic approach in isolating this locus by creating a detailed genetic linkage map of the 725 Mbp Botryllus genome using DNA polymorphisms (primarily identified as AFLPs) as molecular genetic markers. DNA polymorphisms are identified in inbred laboratory strains of Fu/HC defined Botryllus, and their segregation and linkage is analyzed in a series of defined crosses. Using bulk segregant analysis, we have focused our mapping efforts on the Fu/HC region of the genome, and have generated an initial map which delineates the Fu/HC locus to a 5.5 cM region.

BOTRYLLUS schlosseri is a colonial ascidian which inhabits shallow subtidal marine habitats around the world. As protochordates, these organisms occupy a key phylogenetic position in the evolution of vertebrates: protochordates are likely the direct predecessors of the chordate line (reviewed in Satoh and Jeffrey 1995). The product of ascidian sexual reproduction is a free swimming tadpole with a chordate body plan, including a notochord, dorsal hollow nerve tube and segmented musculature. Following hatching from the mother colony, the tadpole swims to a nearby surface, attaches to the substratum and metamorphoses to an invertebrate form. In the colonial ascidians such as *B. schlosseri*, metamorphosis is followed by a budding process which gives rise to a large colony of asexually derived, genetically identical offspring united by a common vascular network.

B. schlosseri is an excellent model for genetic manipulation (Milkman 1967; Sabbadin 1971; Boyd *et al.* 1986). As described above, each colony is made up of many (2 to >1000) asexually derived, genetically identical individuals, and individual sections of a colony (called subclones) can be isolated, transferred to another substrate and will continue to survive. The generation time is very short. Each asexually derived bud has a lifespan of one week, and during this time it reproduces both asexually, by budding off a new individual, as well as sexually. The gestation time is 5–7 days, and thus each week a new clutch of tadpole larvae is released. These new individuals can be induced to become sexually mature within 3–4 weeks (Boyd *et al.* 1986). Additionally, the animals are hermaphrodites and can be self-crossed in the laboratory.

The primary focus of our laboratory has been on elucidating the molecular mechanisms which underlie a genetically defined, natural transplantation reaction in Botryllus. As a colony asexually expands outwards, it often interacts with other colonies which settled nearby. When two colonies of *B. schlosseri* come into close contact, terminal projections of the colony vasculature, called ampullae, reach out from each individual and contact each other. Two outcomes can result from this interaction: either the ampullae will fuse together and form a single chimeric colony with a common vasculature, or the two colonies will reject each other in an inflammatory reaction during which the interacting ampullae are destroyed and the two colonies no longer interact. These allorecognition phenomena are controlled by a single Mendelian locus with multiple codominantly expressed alleles, called the fusion/histocompatibility (Fu/HC) locus (Bancroft 1903; Oka and Watanabe 1957, 1960; Sabbadin 1962; Scofield *et al.*

Corresponding author: Anthony W. De Tomaso, Stanford University, Hopkins Marine Station, Pacific Grove, CA 93950.
E-mail: tdet@leland.stanford.edu

1982). In order to fuse, two colonies must have at least one Fu/HC allele in common; those sharing none will undergo rejection.

The Fu/HC-based allorecognition system has been extensively characterized at a genetic level for almost a century. In contrast to other invertebrate and vertebrate self/nonself recognition systems, Fu/HC-based allorecognition is species specific. Colonies of *B. schlosseri* have overlapping distributions, but will completely ignore, and can even be overgrown by other, closely related Botryllid ascidians (Oka and Watanabe 1960), even though these other species have completely analogous single locus allorecognition systems (reviewed in Saito *et al.* 1994).

The Fu/HC locus is also extremely polymorphic, and several studies have demonstrated hundreds of alleles in small populations. Rinkevich *et al.* (1995) estimated that there were as many as 300 different alleles in a 5000 m^2 area in the Mediterranean, and Grosberg and Quinn (1986) found an estimated 50–100 Fu/HC alleles along a 20-meter transect in Eel Pond (Woods Hole, MA). This level of polymorphism effectively ensures that all wild-type colonies are heterozygotic at the Fu/HC locus, and that only closely related individuals are able to fuse. Furthermore, Fu/HC-based allorecognition events may control other processes besides interacting colonies as described above. It has been suggested that detection of Fu/HC polymorphisms by the tadpole larvae causes them to settle near histocompatible individuals (and hence kin; Grosberg and Quinn 1986). Fu/HC-based allorecognition may also be a barrier to self-fertilization in the hermaphroditic colonies, as some (but not all) experiments suggested that sperm which did not share Fu/HC alleles were more competitive at fertilization than sperm which shared alleles with the mother colony (Watanabe 1962; Scofield *et al.* 1982; Grosberg 1987).

Allorecognition in *B. schlosseri* is analogous to that seen in the adaptive immune system of the higher vertebrates. Recognition events in the latter are centered around the major histocompatibility complex (MHC), and the parallels between the Fu/HC and MHC are clear. Both mediate allorecognition events as single Mendelian loci, both are extremely polymorphic (ensuring that most wild-type individuals are heterozygotic), and both may be involved in other functions, such as controlling the outcome of certain mating events (see above; Boyse *et al.* 1987).

The phylogenetic relationship of the protochordates to the vertebrates, and the analogies between Fu/HC- and MHC-based recognition events make it tempting to speculate on the relationship of the Fu/HC to the MHC, the possible role of the Fu/HC in the origins of adaptive immunity, or whether a Fu/HC-based allorecognition system is still present in the higher vertebrates. However, these questions cannot be answered until the Fu/HC locus has been characterized at the molecular level.

Over the last several years, our laboratory has developed partially inbred laboratory-reared strains of *B. schlosseri* with defined Fu/HC alleles, which allows us to analyze the segregation of the Fu/HC locus in defined crosses. Using this system we have decided to take a genomic approach in isolating the Fu/HC locus via making a genetic map, isolating tightly linked flanking markers and using these as starting points for a genomic walk. Using defined crosses and bulk segregant analysis (Michelmore *et al.* 1991), we have identified eleven markers tightly linked to the Fu/HC locus, two of which delineate the Fu/HC to a 5.5-cM region.

MATERIALS AND METHODS

Mariculture and partially inbred lines: Conditions for raising and crossing *B. schlosseri* in the laboratory and an in depth discussion of the life history have been described (Boyd *et al.* 1986; Sabbadin 1971; Milkman 1967). Derivation of the partially inbred laboratory strains has also been described briefly (Rinkevich *et al.* 1993), and is discussed in the results section.

Mapping population: The map was generated from the analysis of an F_2 population (presently at 75 individuals) from the parents BBYd72 (Fu/HC^{BX}) and Yw1023 (Fu/HC^{AA}). The pedigree of the two individuals used to create the mapping population and the F_1 population used in this study are shown in Figure 1. The BBYd72 animal was derived from an outcross which increased the genetic diversity between the two colonies. However, the Fu/HC^{X} allele is defined as we have several animals which are Fu/HC^{X-}. Botryllus colonies reproduce on a weekly basis with a gestation time of 5–7 days and are able to be continually mated, and this cross is still in progress (discussed below; Boyd *et al.* 1986). Progeny take 5–10 wk to obtain the size where they can be Fu/HC genotyped and material can be collected for DNA extraction.

DNA extraction and AFLP analysis: All enzymes were from New England Biolabs (Beverly, MA), except Taq Polymerase (Boehringer Mannheim, Indianapolis). Chemical reagents were from Sigma (St. Louis). Oligonucleotides were synthesized at the PAN facility at the Stanford University Medical School (Stanford, CA). Frozen tissue samples were individually ground to a fine powder using a mortar and pestle. DNA was extracted on silica columns (Nucleobond C+T Kit; Macherey Nagel, Duren, Germany) using proprietary buffers provided and according to the manufacturer's instructions. Amplified Fragment Length Polymorphisms (AFLPs) were performed essentially as described previously (Vos *et al.* 1995). DNA (200 ng) was cut with restriction enzymes *Eco*RI and *Mse*I for 2 hr at 37° in a 30 μl reaction volume containing 10 mm Tris, 10 mm $MgCl_2$, 50 mm NaCl, 1 mm DTT, pH 7.9. Oligonucleotide adaptors (*Eco*RI = 5 pmol; *Mse*I = 50 pmol) for each sticky end, T4 DNA ligase (400 units) and ATP (1 mm final) were then added for a further 3 hr in a total volume of 40 μl and incubated for at least three hours to overnight. Adaptors are an equimolar amount of two oligos (for *Eco*RI: 5′-CTCGTA GACTGCGTACC-3′ and 5′-AATTGGTACGCAGTC-3′; the *Mse*I adaptor was 5′-GACGATGAGTCCTGAG-3′ and 5′-TACT CAGGACTCAT-3′). After the ligation reaction, 60 μl of $TE_{0.1}$ (10 mm Tris, pH 8, 0.1 mm EDTA) was added. Four microliters of the diluted DNA mix was preamplified with one selective nucleotide on each primer (*Eco*RI = A; *Mse*I = T; preamplification primers are the "base" primers shown below) in a buffer containing 10 mm Tris, pH 8.3, 1.5 mm $MgCl_2$, 50 mm KCl, 0.1% gelatin, 200 μmol dNTPs, 0.5 units of Taq polymerase

and 30 ng of each primer. The PCR reaction was carried out in a Genius Thermocycler (Techne, Cambridge, UK) in the following manner: 1 cycle of 72° for 2 min, followed by 20 cycles of 94° (30 sec), 56° (30 sec), and 72° (1 min). The preamplification mix was diluted 1:20 in $TE_{0.1}$ and 3 μl was used for AFLP fingerprinting with each primer containing 3 selective nucleotides. Primer combinations in the text are designated as a letter (the *Eco*RI primer), followed by a number (the *Mse*I primer). The base primers were 5′-AGACTGCGTAC CAATTCAxx-3′ for the *Eco*RI side, and 5′-GATGAGTCCT GAGTAATxx-3′ for the *Mse*I side. The 3-bp extensions were designated as follows: A = AGG; B = ACT; C = ACA; D = AGC; E = ATG; F = AGT; G = AGA; H = ACG; I = AAG; J = ATC for the *Eco*RI primers, and 1 = TAC; 2 = TCA; 3 = TGA; 4 = TTG; 5 = TCT; 6 = TCG; 7 = TGG; 8 = TGC; 9 = TAG; 10 = TTC for the *Mse*I primers. The *Eco*RI primer was end-labeled with ^{33}P-ATP (New England Nuclear, Boston) using polynucleotide kinase according to the manufacturer's instructions. PCR for AFLPs were done using 5 ng labeled RI primer, and 30 ng *Mse*I primer, in the same buffer described above. There was an initial 2 min denaturation at 95°, followed by the first cycle of 94° (30 sec); 65° (30 sec); 72° (1 min). For the next 11 cycles the annealing temperature was decreased by 0.7 degrees in each cycle. This was followed by 24 cycles of 94° (30 sec); 56° (30 sec); 72° (1 min). PCR reactions were diluted 1:1 in stop solution (98% formamide, 10 mm EDTA pH 8, 0.1% bromophenol blue, 0.1% xylene cyanol), denatured for 5 min at 95°, and resolved on a standard TBE sequencing gel at 75 W for 2.5 hr. Gels were dried and autoradiograms were exposed for 36 hr. Preliminary experiments (not shown) revealed that banding patterns were >99% reproducible on an individual sample basis; however, most data points on the present map were repeated twice. No variation in the scored AFLP markers was found.

Genotyping and segregation analysis: F_2 progeny were scored for their fusion alleles using a colony allorecognition assay with Fu/HC-defined colonies as described previously (Oka and Watanabe 1957). Briefly, small sections (subclones) of the colony were removed and placed in contact with a Fu/HC-defined colony. Fusion or rejection was visibly assayed after 24–48 hr. AFLP polymorphic alleles were scored as present or absent by hand. Linkage analysis of the segregated markers was done using maximum likelihood estimates in the computer program MAPMAKER (Lander *et al.* 1987). Initially, the LOD score for linkage between each pair of loci was determined. Loci were considered linked if the LOD score exceeded 3.0, which represents a relative likelihood of linkage to nonlinkage of 1000:1. Gene order was determined using multipoint analysis with a log-likelihood threshold of 2. Recombination distances were converted to cM using the Kosambi mapping function.

In order to focus our mapping efforts on the region of the genome containing the Fu/HC, we used bulk segregant analysis (BSA; Michelmore *et al.* 1991). Bulk pools consisted of DNA from animals which either contained the Fu/HC^{A} allele, or did not. Each bulk contained DNA from at least 10 individuals, although the same results could be obtained with as little as five in these experiments. Different AFLP primer combinations were screened on the parents and bulks. Putative linked markers that were revealed using this strategy were then scored on the entire mapping population.

Genome size analysis: Individual Botryllus cells were isolated by repeatedly chopping a colony on a glass slide with a razor blade. This chopped colony was then transferred into a 70-μm Nylon Cell Strainer (Falcon, Lincoln Park, NJ) which fit onto the top of a 50-ml conical tube. The colony was ground into the filter with the rubber end of a 25-ml syringe plunger. Every 30 sec, 2 ml of isolation buffer [0.2 μm filtered seawater; 0.5% bovine serum albumin (BSA) and 10 mm DTT] was poured through the filter to flush the individual cells through the filter and into the tube. Isolated cells were gently spun down (1000 × *g*; 5 min) and resuspended in 10 ml of isolation buffer. This treatment appeared to keep the majority of the cells intact. While some cell lysis occurred as detected by hemocyte pigments released into the buffer, >98% of the cells were viable after the first wash as assayed by trypan blue exclusion. Cells were counted on a hemocytometer and 3×10^6 were pelleted and resuspended in 1 ml of filtered seawater containing 1% Tween-20, 10 μg/ml RNAse A, and 10 μg/ml propridium iodide, and stored overnight at 4°. Chicken nucleated peripheral blood and spleen cells isolated from a C57Bl/Ka-Thy 1.1 mouse were used as standards, isolated and prepared as described previously (Tiersch *et al.* 1989; Galbraith *et al.* 1983), and resuspended for flow cytometry in Hepes Buffered Saline (25 mm HEPES, pH 7.4; 150 mm NaCl), Tween-20 and propridium iodide as described above. Released nuclei were then analyzed for DNA content by FACS. 20,000 nuclei were analyzed per sample, done on both a log scale using all three samples, as well as a linear scale with chicken and Botryllus. Forward and side scatter were gated so analysis was restricted to single nuclei. This experiment was repeated five times using 12 different animals [both inbred laboratory strains as well as wild-type animals collected from the Monterey Marina (Monterey, CA)] with no significant variation (<5%). In all experiments both standards were used, and each correctly estimated the size of the other. One experiment is shown in Figure 2.

RESULTS

Derivation of partially inbred lines and genetic crosses: Fu/HC-defined strains were initially created in two ways. First, we isolated a wild-type animal from the Monterey Marina, heterozygous at the Fu/HC locus, and the hermaphroditic colony was self-crossed. Progeny were genotyped for their Fu/HC alleles by colony allorecognition assays, where subclones of the colony are placed in contact at their growing surfaces and fusion or rejection visibly assayed. Any two colonies which rejected each other were necessarily homozygous for different Fu/HC alleles, as only one Fu/HC allele needs to be shared for fusion. Unfortunately, severe inbreeding depression prevented systematic inbreeding of these two strains, such that after two rounds of self-crossing few sexually mature colonies could be obtained, as had been described previously (Sabbadin 1971), and these lines died out.

To increase the genetic diversity in the first cross and resulting Fu/HC homozygous individuals, two adjacent wild-type colonies were isolated, which fused in the laboratory, demonstrating they shared at least one Fu/HC allele in common (*i.e.*, Fu/HC^{AB} and Fu/HC^{AD}). These two colonies were crossed and the progeny were analyzed and grouped according to similar Fu/HC phenotypes by testing a subclone of each F_1 progeny against all the other progeny. This resulted in animals being categorized into three groups: one group where all the members fused with every other F_1 progeny, and two groups in which all the members fused with the first group, but which rejected each other. The number of

progeny in each group was in the expected 1:2:1 ratio, with the larger group containing the animals which always fused.

One of the two smaller groups was necessarily homozygous at the Fu/HC locus. To determine which group it was, members of each group were intercrossed, and these F_2 progeny were again tested for Fu/HC phenotype by testing a subclone of each animal among themselves, as well as with all of the original F_1 animals. We expected to see the progeny of one of the F_1 rejecting groups (the Fu/HC homozygotes) showing a single Fu/HC phenotype, and the other F_2 group (heterozygous for the other two Fu/HC alleles in the cross) to break down into another 1:2:1 Fu/HC phenotype grouping. Surprisingly, individuals in each group instead had a single Fu/HC phenotype, suggesting that both F_1 groups were homozygous. As described above, the tadpole larvae of *B. schlosseri* have a tendency to aggregate near kin, and the Fu/HC is highly polymorphic (Grosberg and Quinn 1986), thus fusible adjacent colonies are probably half-sibs (Fu/HC^{AB} and Fu/HC^{AD}). However, we had serendipitously isolated two colonies which shared both Fu/HC alleles and this cross resulted in creating two homozygous lines.

These original Fu/HC alleles were designated as Fu/HC^{A} or Fu/HC^{B}, and these animals were the founders of our present day strains. However, severe inbreeding depression again prevented systematic breeding of these two strains, with lethality occurring at defined stages of embryonic and fetal development. From these and other studies it appears that there are at least 14 independent recessive lethal genes which act during embryogenesis and/or fetal development (Sabbadin 1971; our unpublished data) and we have not yet been able to purge these high penetrance genes from our laboratory populations.

We then selected the highest viability colonies which could be self- or outcrossed to other colonies sharing the Fu/HC allele and produced the highest number of healthy offspring. In general, we have attempted to maintain homozygotes of the Fu/HC^{A} and Fu/HC^{B} alleles by breeding tested individuals derived from the original cross, and we often bred siblings, cousins, *etc.*, and simply selected for the highest production offspring for the next generation. Thus offspring may represent crosses between individuals derived from the original cross at several different generations (F_1, F_2, BC, *etc.*; see Figure 1). These lines are tending toward homozygosity, but not to the extent that could be obtained by a rigorous program of self-crossing or by brother/sister mating. In addition, we have several other defined alleles as well as laboratory-reared Fu/HC wild-type strains available for experiments.

The pedigree of the two parents used in this cross (Yw1023 and BBYd72) and their relationship to the founder colonies is shown in Figure 1. Colony Yw1023 has been bred completely in the laboratory and its lineage traced back to the original founder. Colony BYd72 was the offspring from an outcross of a Fu/HC^{BB} homozygote to a wild-type colony, and this was done for several reasons. Primarily, it increased the genetic diversity between the two strains prior to mapping so that general inbreeding depression did not cause an undetectable segregation distortion in the cross. However, we have also observed that as our defined lines become more inbred, we see a specific segregation distortion at the Fu/HC locus: a lower number of Fu/HC homozygotes appears to be able to survive to adulthood and grow to the size were they can be Fu/HC genotyped and sampled, which requires at least 10 systems. However, the larvae appear in correct Mendelian ratios (unpublished data). Because we wanted to use a bulk segregant strategy to focus our initial mapping efforts on the Fu/HC region of the genome, it was imperative that at least one allele be absent in many of the F_2 progeny; having the third Fu/HC allele in the cross allowed this. The Fu/HC^{X} allele is defined and can be genotyped.

Our initial mapping population is an F_2 intercross and presently consists of 75 Fu/HC genotyped adults as well as approximately 200 unscored juveniles which are currently too small for genotyping, and is still in progress. This type of cross was done as we originally envisioned using codominant RFLPs as genetic markers, an F_2 intercross being the most informative. A backcross would be more efficient at detecting recombinants using the dominant AFLP genetic markers (Allard 1956); however, the Fu/HC itself is codominant, and we initially focused our mapping efforts on this area of the genome. We are also presently at the F_1 stage of several backcrosses between the Fu/HC^{AA} homozygote Yw1023 and wild-type colonies (discussed below).

Genome size of *B. schlosseri*: Very little is known about the genomes of the Botryllid ascidians. Colombera (1974) karyotyped *B. schlosseri* from a number of sites around the Mediterranean and found a haploid chromosome content of 16. Preliminary studies in our lab show that this appears to be true of *B. schlosseri* found in Monterey Bay (our unpublished data).

The haploid genome size (C-value) of some of the solitary tunicates has also been determined (Lambert and Laird 1971; Taylor 1967) and found to range from approximately 150 Mbp for the primitive species *Ciona intestinalis*, to near 300 Mbp in *Styela plicata*. We estimated the genome size of *B. schlosseri* from Monterey Bay by comparing the propridium iodide staining of isolated Botryllus nuclei to those from chicken (*Gallus domesticus*; 1.2×10^9 bp; 2.4 pg/nucleus) and mouse (*Mus musculus*; 3×10^9 bp; 6.2 pg/nucleus) by FACS analysis as described in materials and methods, and the results are shown in Figure 2. Isolated Botryllus nuclei contain approximately 1.48 ± 0.04 ($n = 12$) pg DNA/nucleus, corresponding to a haploid genome size of approximately 725 Mbp. There was no significant

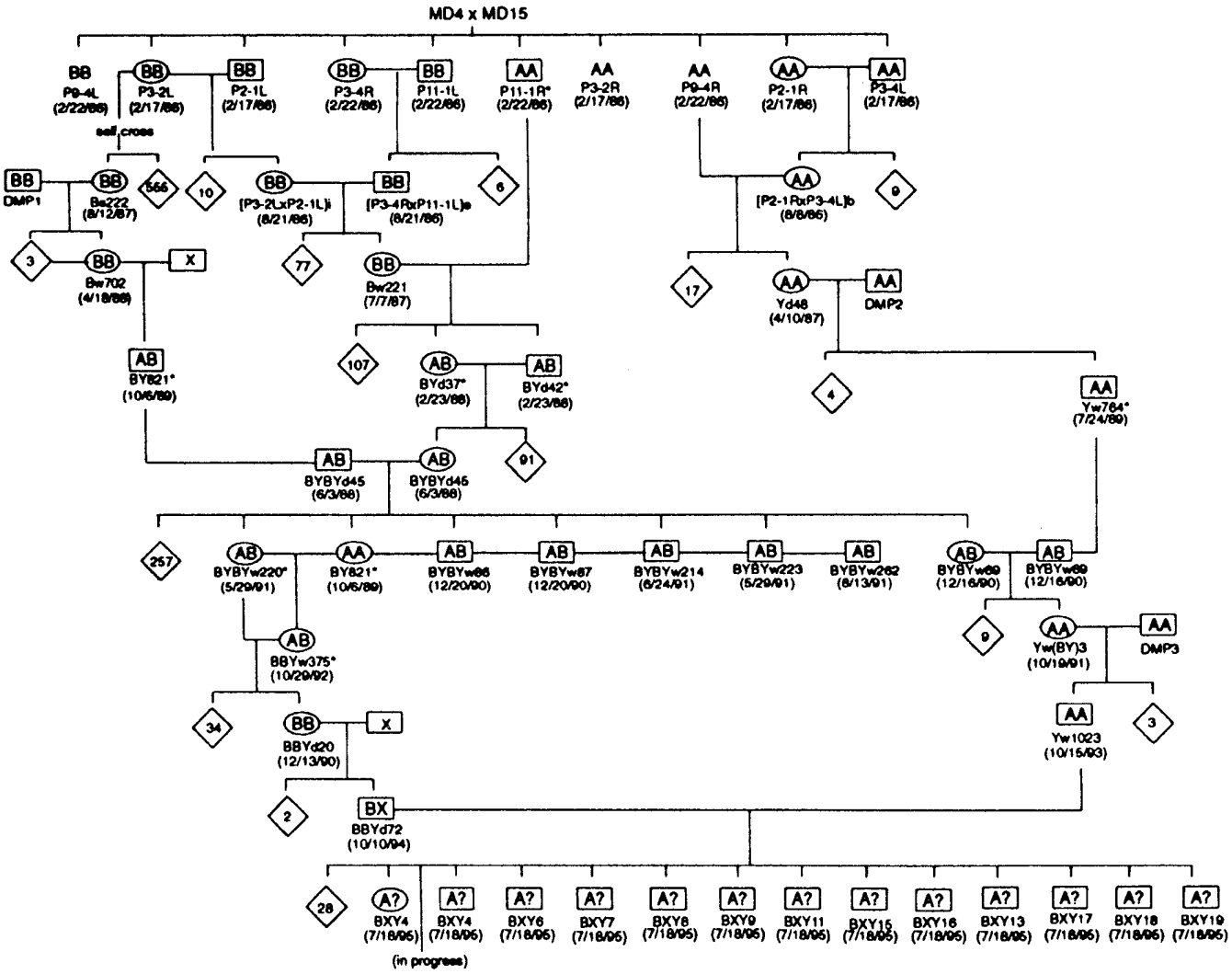

Figure 1.—Pedigree of inbred strains Yw1023 and BBYD72. The relationship of the partially inbred lines used in this study to the founder colonies is shown. MD4 and MD15 were adjacent wild-type colonies isolated from the Monterey marina which shared at least one Fu/HC allele. The crosses used to derive these individuals are shown, as well as the first clutch of F_1 individuals in this mapping population. Symbols: ◇ = number of individuals produced by maternal colony; ○ = maternal colony; □ = paternal colony.

variation between inbred lab colonies and colonies collected from the Monterey marina.

Genetic mapping: To map the genome of *B. schlosseri*, we are using DNA polymorphisms as molecular genetic markers. These polymorphisms are being identified using AFLPs (Vos *et al.* 1995). We began our mapping efforts with 100 AFLP primer combinations and 75 F_2 progeny. F_2 individuals were genotyped for their Fu/HC alleles and DNA was isolated and prepared for AFLP analysis. F_2 DNA pools were set up based on the segregation of the Fu/HC^A allele, and were created from both Fu/HC^A+ and Fu/HC^A- F_2 progeny for bulk segregant analysis.

Each AFLP primer combination (+3/+3 selective nucleotides; see materials and methods) on average amplified 55 loci, of which an average of 12 were polymorphic between the parental animals, for a total of 5500 loci scanned, revealing 1200 polymorphic markers. Since our strains are not completely inbred, we first analyzed segregation of the positive markers in both F_1 and F_2 generations. In over 90% of the cases, marker loci segregated in their expected ratio in the F_2 progeny (1:1 or 3:1), depending if they were homo- or heterozygous in the parental strains, with a χ^2 threshold of $P > 0.10$ (Figure 3, Table 1). When two unlinked homozygous markers were analyzed for their segregation in the F_2 intercross mapping population (*i.e.*, heterozygous $F_1 \times$ heterozygous F_1), they showed the dihybrid 9:3:3:1 segregation ratios expected for dominant markers (Table 2).

After establishing Mendelian segregation of the AFLP loci, we next did bulk segregant analysis on the pooled F_2 DNA samples (Fu/HC^A+ and Fu/HC^A-). These pools were initially screened with 64 primer sets which revealed 768 polymorphic markers, and resulted in 10

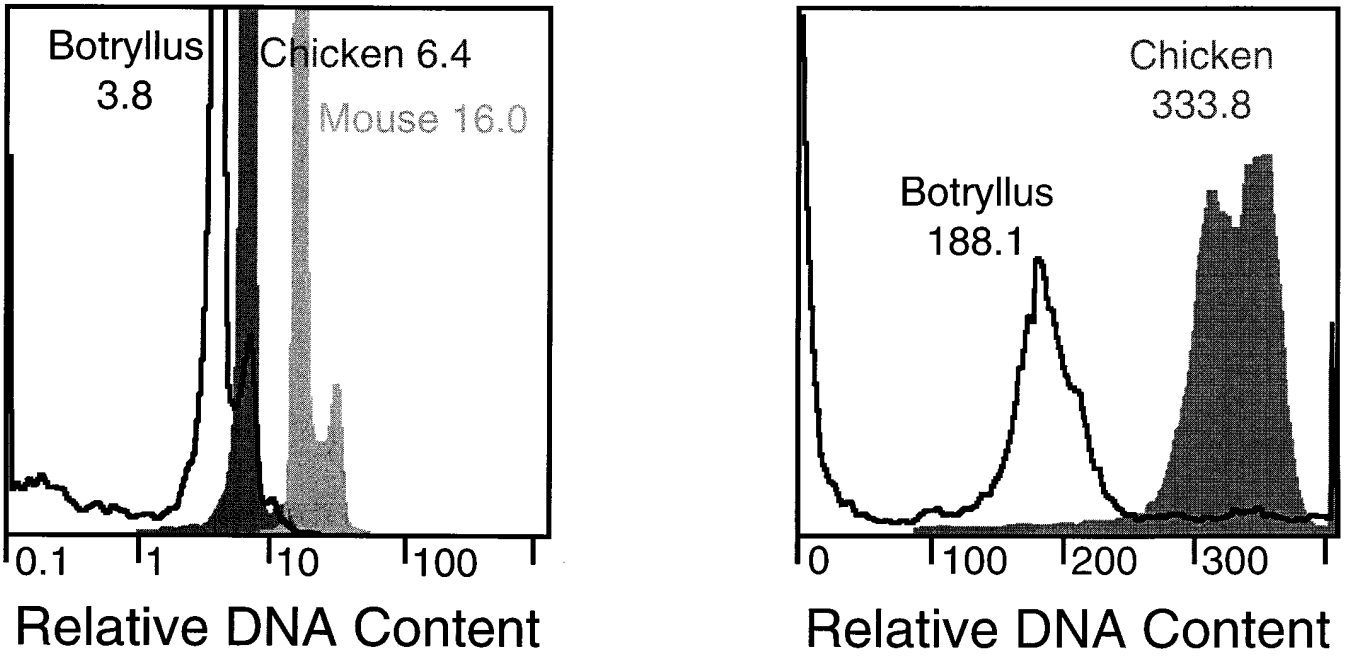

Figure 2.—FACS analysis of the Botryllus genome. The size of the Botryllus genome was determined by FACS analysis using mouse spleen cells, and nucleated chicken red blood cells as standards. The DNA content of the Botryllus cells was determined by comparing the propridium iodide signal from the flow cytometer of all three samples to the published values of the mouse (6.2 pg/nucleus) and chicken (2.5 pg/nucleus) genomes. All three samples were analyzed on a log scale (left), and chicken and Botryllus on a linear scale (right). The nucleus of Botryllus contains 1.48 ± 0.04 (n = 12) picograms of DNA, corresponding to a haploid genome size of approximately 725 Mbp.

putative AFLP loci linked to the Fu/HC (Figure 4). Each individual in the mapping population was individually typed for these markers; eight of the ten showed tight linkage (LOD > 3.5) and were ordered using multipoint analysis in the computer program MAPMAKER (Lander *et al.* 1987). Two of these markers (E4 and F4) flanked the Fu/HC and delineated it to a 9.2-cM region of the genome.

Next, new DNA pools from animals which showed crossovers at the E4 or F4 AFLP loci were made, and another 36 AFLP primer sets were tested on both the original *Fu/HC*^{A}+ and *Fu/HC*^{A}− bulk pools, as well as the E4 and F4 crossover (CX) pools (Figure 5). This reiteration of the pooling strategy allows us to narrow our search for linked markers to within the original 9.2-cM region. Any new marker which shows tight linkage in the *Fu/HC*^{A}+ and *Fu/HC*^{A}− bulk pools, but is absent from the one of the crossover pools is likely to be distal to the original flanking marker. Conversely, any marker which shows linkage on all the pools is likely proximal to the Fu/HC locus. Three new markers (I7, D5 and C10) were revealed using this strategy and all three were genotyped on the entire mapping population. These eleven markers were then reanalyzed in MAPMAKER, and as shown in Figure 6, nine of these AFLP markers were ordered with a log-likelihood of 2.0, and three of these markers (D5, E4 and I7) have narrowed the region in which the Fu/HC must reside to less than 5.5 cM. We also have one marker (C10) which showed no re-

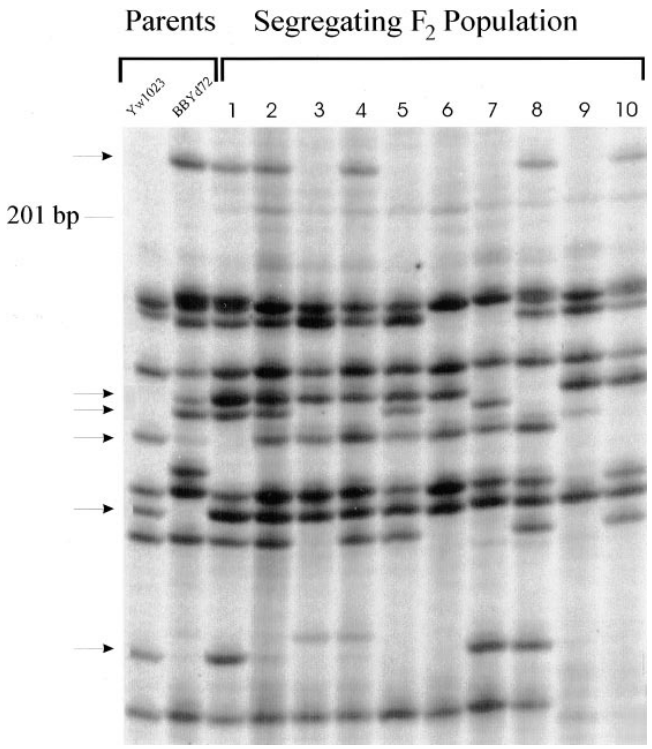

Figure 3.—Inheritance of AFLP loci. AFLP fingerprints from a single primer set performed on the two parents, and 10 F_2 individuals of the Yw1023 × BBYd72 F_2 mapping population. The segregation of six polymorphic loci (arrows) are shown. A small section of the complete gel is shown for clarity.

TABLE 1

Mendelian inheritance of 10 randomly picked AFLP loci

Locus (zygosity)	No. indiv. scored	Band present (expected)	Band present (observed)	Band absent (expected)	Band absent (observed)	$\Sigma\chi^2$	$P >$
1 (homo)	58	43.5	42	14.5	16	0.21	0.90
2 (homo)	58	43.5	49	14.5	9	2.70	0.10
3 (homo)	58	43.5	45	14.5	13	0.20	0.90
4 (homo)	58	29.0	25	29.0	33	1.10	0.25
5 (homo)	59	44.3	46	14.8	14	0.05	0.75
6 (hetero)	59	29.5	26	29.5	33	0.84	0.25
7 (homo)	58	43.5	40	14.5	18	1.12	0.25
8 (homo)	58	43.5	43	14.5	15	0.02	0.75
9 (homo)	52	39.0	42	13.0	10	0.92	0.25
10 (homo)	52	39.0	38	13.0	14	0.09	0.75

Ten AFLP loci were randomly picked and their segregation analyzed in both the F_1 and F_2 mapping populations. Results from the F_1 population revealed zygosity. F_2 individuals were scored for these loci and marker segregation analyzed by a χ^2 test for goodness of fit.

combination with the Fu/HC locus, but due to the size of the mapping population it could still be several cM away. The other linked AFLP markers revealed by bulk segregant analysis could not be placed with this log-likelihood threshold, *i.e.*, the best fit order was only slightly better than the second alternative, and are not shown on this figure. Thus we have narrowed down the location of the Fu/HC locus to less than 6 cM using only 2.5% of the available AFLP primer sets with this particular restriction enzyme combination. Furthermore, all three markers that were revealed using the crossover pools were inside the original flanking markers. As the mapping population grows, we will be able to continue this process and saturate genetic markers very close to the Fu/HC locus.

Although the AFLPs were very reliable when analyzed for Mendelian inheritance, we were curious to see how they performed when actually mapping the Fu/HC locus. We reanalyzed the data set of all the markers scored on the mapping population, looking for individuals who at any particular locus had a genotype which was different from both flanking loci. This could indicate a double crossover, but if the loci are tightly linked it probably represents a scoring error or an artifact. This was done both by hand, and by using the "genotype" command in MAPMAKER. Out of the 825 data points, we found two individuals with this double crossover genotype. Both were genotyped again and the results were the same. In one of these individuals, the flanking markers were 8-cM and 6-cM apart, making it probable that a double crossover could have been observed in this size mapping population. In the other individual, the genotype did look like an artifact, as it was very tightly flanked by two other markers. Considering that this technique is a blind PCR-based assay, one potential artifact out of 825 data points did not seem too disturbing, and AFLP-identified DNA polymorphisms appear to be reliable genetic markers.

TABLE 2

Segregation analysis of two unlinked loci show dihybrid ratios

Number scored	Radio (expected)[a]	Ratio (observed)	$\Sigma\chi^2$	$P >$
57	32:11:11:4	36:6:11:4	2.61	0.5
55	31:10:10:3	36:8:8:3	1.92	0.5
59	33:11:11:4	34:8:12:5	1.05	0.5
58	33:11:11:4	35:11:8:4	0.97	0.7
56	32:10:10:4	36:8:10:2	1.91	0.5

Five pairs of homozygous, unlinked AFLP loci from Table 1 were picked and scored in the progeny of an F_2 intercross. Expected and observed ratios were compared by a χ^2 test for goodness of fit. In all cases examined ($n = 20$), pairs of unlinked markers segregate in expected dihybrid ratios for dominant markers.

[a] Expected ratio was rounded to whole numbers for presentation.

Using the dominant AFLP markers and an F_2 mapping population is not the most effective way of detecting recombinants; a backcross would be more efficient as there is no ambiguity in scoring the markers (Allard 1956). For independent verification of linkage groups, as well as for a future genome-wide linkage map, our strategy is to analyze marker segregation in the present mapping population, as well as in at least two independent backcross populations. These backcrosses, which consist of each of the two parents used in this cross (Yw1023 and BBYd72), crossed to *Fu/HC* wild-type laboratory strains, are presently at the F_1 stage. In these crosses we are focusing our efforts on analyzing any unresolved areas from the F_2 intercross, as well as rechecking established linkage groups if possible. We have found that 60% of the AFLP markers that are polymorphic in the F_2 cross are also polymorphic with our *Fu/HC* wild-type

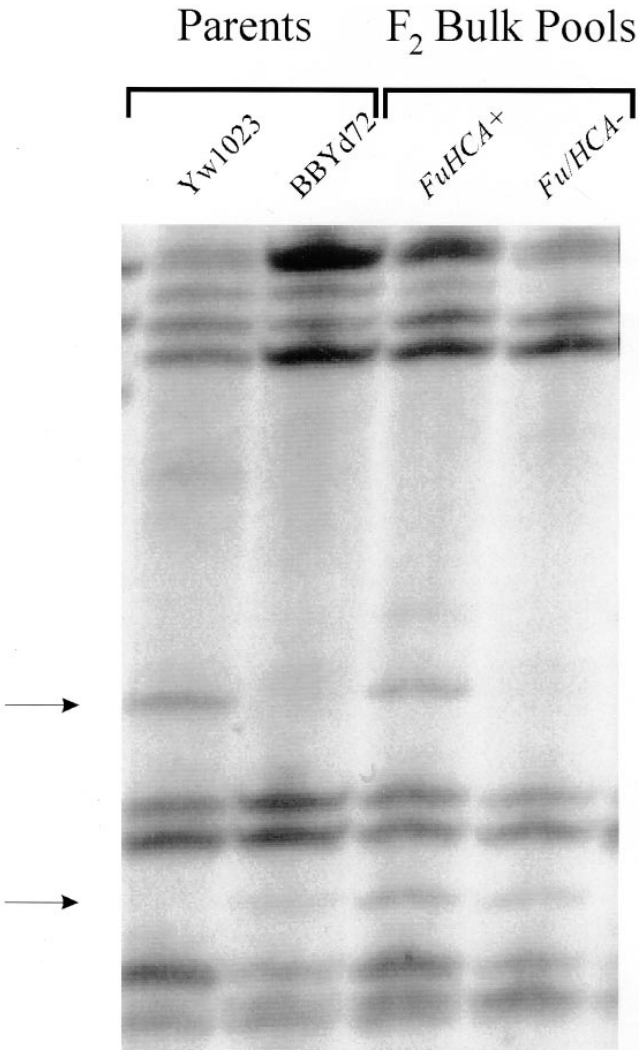

Figure 4.—Bulk segregant analysis using AFLP markers. AFLP fingerprints of the two parents and two bulk pools of F_2 individuals (n = 10) as described in the text. The top arrow shows an AFLP allele which is polymorphic between the parents and is only present in the Fu/HC^{AA} parent. This marker is also only present in the Fu/HC^{A}+ F_2 pool, suggesting that it is linked to the Fu/HC^{A} allele. The bottom arrow shows a marker present in the Fu/HC^{BX} parent, but present in both bulk pools, suggesting that it is unlinked to the Fu/HC locus. Putative linked alleles are then scored on the entire mapping population to confirm linkage.

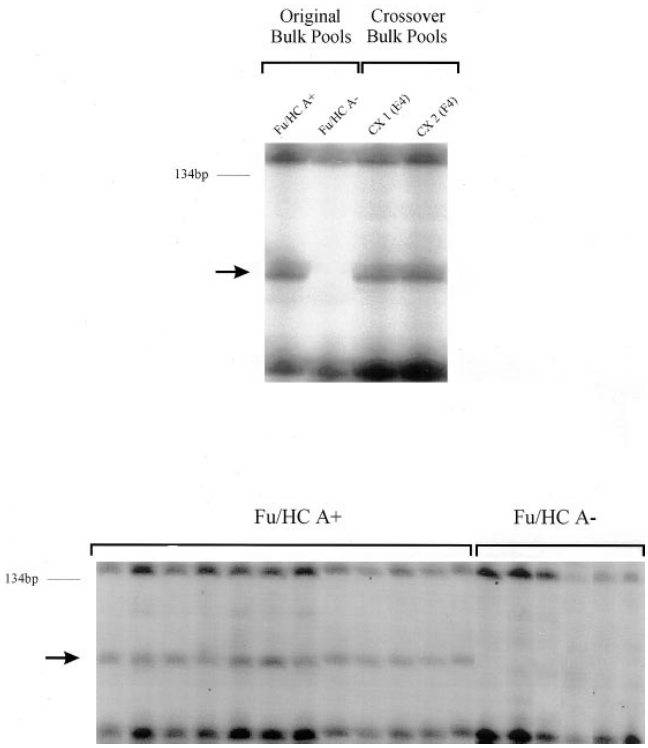

Figure 5.—Reiteration of the bulk segregant strategy using recombinant individuals. DNA from Fu/HC A+ individuals recombinant for AFLP markers E4 and F4 were combined into crossover pools CX 1 (E4 recombinants) or CX 2 (F4 recombinants), and screened with 36 AFLP primer sets as described in the text. (A) A portion of the gel using primer set C10. The arrow points to an AFLP linked to the Fu/HC A allele in the original bulk pools, and also present in both crossover pools, suggesting that it is closer to the Fu/HC locus than the original flanking markers. (B) The same AFLP allele scored on 12 Fu/HC A+, and 6 Fu/HC A− F_2 individuals, which includes those recombinant individuals used to make the crossover pools.

laboratory strains, which allows independent verification of their segregation in these crosses (Figure 7).

DISCUSSION

In this report we present our initial molecular genetic studies on the colonial protochordate, *B. schlosseri*. As part of our studies on the genetics of allorecognition, we are developing a detailed genetic map of the Botryllus, using partially inbred lines with defined Fu/HC alleles developed in our laboratory over the last 13 years. Here we used bulk segregant analysis (Michelmore *et al.* 1991) to focus our initial mapping efforts on the Fu/HC region of the genome and present a low resolution genetic map which delineates the Fu/HC locus to at least a 5.5-cM region of the genome. We have also determined the genome size of *B. schlosseri* from Monterey Bay and found it to be approximately 725 Mbp, or about 25% the size of human.

B. schlosseri represents an excellent protochordate for genetic studies. These animals can be raised in the laboratory, are fast growing, reproduce weekly and have a short generation time (5–7 days gestation, 3–4 wk to sexual maturity). Because they are colonial hermaphrodites, pieces of the colony can be isolated and self-crossed, and one genetic individual can be used in several different experiments concurrently.

However, one of the most difficult tasks in working with this organism is dealing with severe inbreeding depression. As previously stated, we believe that there are at least 14 high penetrance recessive lethal genes which we have not yet been able to purge from our

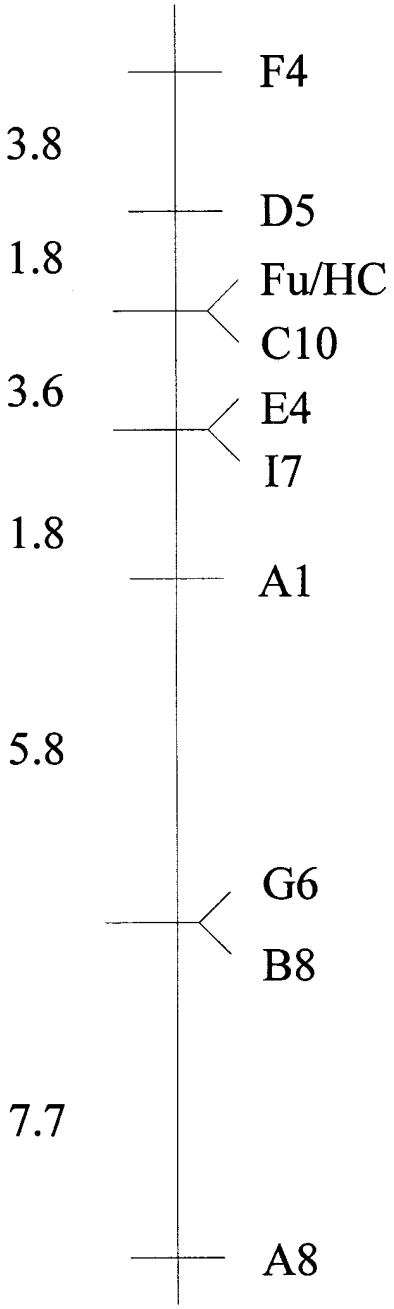

Figure 6.—A low resolution genetic map encompassing the Fu/HC locus. Eleven putative linked loci revealed by bulk segregant analysis were scored on the entire mapping population (n = 75). Linkage and ordering were done using maximum likelihood estimates in the computer program MAPMAKER. Loci were considered linked with a threshold LOD score of >3.0, and ordered with a LOD score of >2. Nine markers met this criteria. Recombination frequencies were converted to centimorgans using the Kosambi mapping function and are shown on the left. Genetic markers are represented as a letter (the *Eco*RI primer) and a number (the *Mse*I primer) as described in materials and methods.

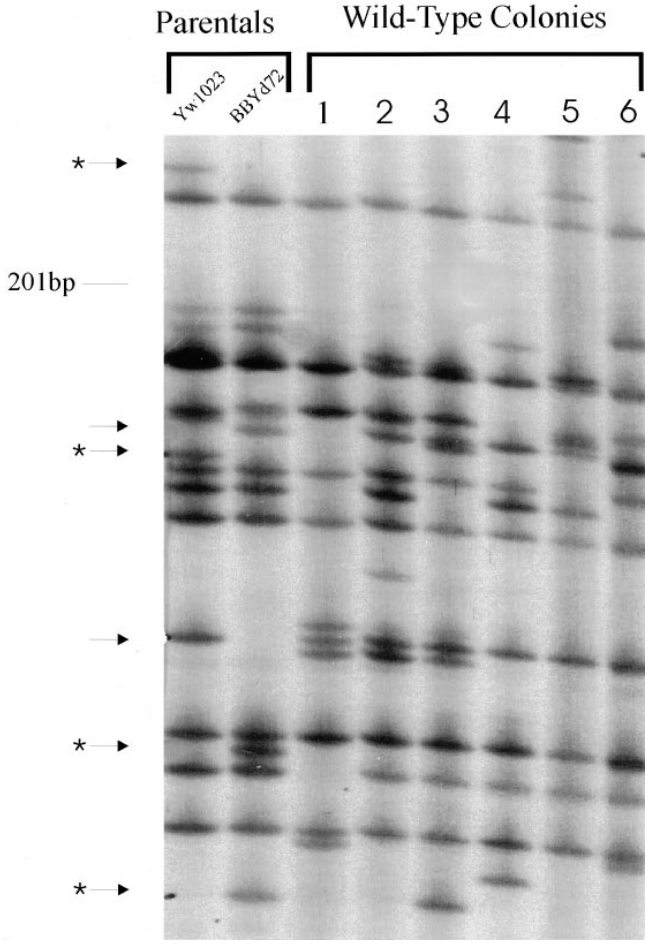

Figure 7.—Outcrosses allow independent verification of the segregation of many of the AFLP markers. A section of an AFLP fingerprint gel comparing the parents of the main cross with six different wild-type Fu/HC heterozygote lab strains. Six loci which are polymorphic between the Yw1023 and BBYD72 parents are shown (arrows). Four of these loci (*) are also polymorphic with the wild-type colonies. This allows independent verification of the segregation of individual AFLP loci in different crosses.

laboratory population. In another laboratory an animal originally collected from Monterey Bay has been self-crossed through four generations, and the frequency of viable offspring has decreased in each generation (B. Rinkevich, personal communication). Thus much work remains to be done before we have bred past these genes and have true inbred lines.

The size of the Monterey Bay *B. schlosseri* genome (725 Mbp) was larger than expected. Previous studies, while both on solitary tunicates, had shown that the genome sizes ranged from ~5–10% that of human, corresponding to sizes of ~150 Mbp for *Ciona intestinalis*, considered one of the most primitive ascidians (Lambert and Laird 1971) to approximately 300 Mbp for *Styela plicata*, a more advanced species (Taylor 1967). The complexity of the Ciona genome was also analyzed by *Cot* analysis and the kinetics suggested that approximately 70% of the DNA was single copy (Lambert and Laird 1971). As the earliest aquatic animals in the chordate line, it was suggested that the ascidians might

have simple, compact genomes. Further evidence for this idea was provided by Colombera (1974), who karyotyped over 30 species of ascidians at different localities and found no indication that speciation in this class had been accompanied by polyploidization. Thus we were not expecting to find that the Botryllus genome was almost twice the size of Styela, as well as of a compact chordate genome like *Fugu rubripes* (400 Mbp; Elgar 1996). In our analysis we looked at total dissociated cells by FACS analysis of propridium iodide staining and saw very little variation among different animals, regardless of the asexual budding stage of the colonies. We could also see a probable 4N peak of correct size on the log scale (Figure 2), all of which suggest that our analysis is correct. Although there is no relationship between the C-value and the evolutionary position of an organism (Li and Graur 1991), this did unfortunately indicate that there was more work to be done than expected.

The use of AFLPs to identify DNA polymorphisms has been invaluable for this study (Vos *et al.* 1995). This technique is fast, reliable, provides an almost inexhaustible supply of genetic markers, and requires very little input DNA. This latter point is particularly important as Botryllus is a small animal; each asexually derived individual in a colony is only about 4 mm^2 and could not provide the starting material for anything but a PCR-based assay. In this study, we averaged twelve polymorphic markers per AFLP primer set, and coupled with a bulk segregant strategy we were able to screen through over 200 polymorphisms / day searching for linked markers. Furthermore, since we used primers with 3-bp extensions, we have 4096 (2 primers with 4^3 sequences each) possible primers to screen through, and a single restriction site change provides 4096 potential new primers. In terms of reliability, over 90% of the polymorphic loci we followed were inherited in a Mendelian fashion when examined both as single loci and in pairs of unlinked loci. Additionally there was only one individual data point which may have been a PCR-based artifact in this entire study. So far, it appears that the AFLP technique is an excellent way to map an uncharacterized genome.

The ability to screen through so many polymorphisms will be a tremendous advantage when we get to the physical mapping phase of positionally cloning the Fu/HC locus. We are currently making BAC libraries (Schizua *et al.* 1992) of the genome, using the Fu/HC^{AA} parent, Yw1023, as the source of the DNA. This will allow us to screen the library using the AFLP primers. Furthermore, as the mapping population grows to 200–300 individuals, we will be able to reliably find markers <1-cM distance by using the strategy already implemented. This will greatly reduce the amount of physical mapping required, as each AFLP marker is already a sequence tagged site (STS; Olson *et al.* 1989), as distinguished by the primer sets and fragment size, and does not require sequencing or new primer synthesis.

A long-term goal of our lab is to complete a genome-wide linkage map. Besides Fu/HC-based allorecognition, we have observed genetic components to a secondary allorecognition phenomenon called resorbtion (Rinkevich *et al.* 1993), the life span of the colonies (Rinkevich *et al.* 1992), and the ability of two fused colonies to parasitize each other (Stoner and Weissman 1996). Botryllus also has two developmental pathways (sexual and asexual) which give rise to an indistinguishable adult; we have already observed several common mutations (discussed above). *B. schlosseri* represents a potential link between the most complex of the invertebrate model organisms (*C. elegans* and Drosophila), and the most primitive of the vertebrates (zebrafish), and in addition to the biological phenomena described above, a model protochordate such as *B. schlosseri* will be an important resource in the approaching era of large-scale comparative genomics.

We thank Peter Oeffner for introducing us to the AFLP technique, and Shauna Sommerville, Iain Wilson, Joe Ogas, Jennifer Weller and Paul Levine for helpful discussions. Sam Chesire, David Travers, Doug Wright and Jos Domen assisted with the FACS analysis and helpful discussions. This work was supported by a National Institutes of Health Immunology Postdoctoral fellowship to A.W.D., and a Sandoz/Systemix grant to I.L.W.

LITERATURE CITED

Allard, R. W., 1956 Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia **24:** 235–278.

Bancroft, F. W., 1903 Variation and fusion of colonies in compound ascidians. Proc. Calif. Acad. Sci. **3:** 138–186.

Boyd, H. C., S. K. Brown, J. A. Harp and I. L. Weissman, 1986 Growth and sexual maturation of laboratory cultured monterey *Botryllus schlosseri*. Biol. Bull. **170:** 91–109.

Boyse, E. A., G. K. Beauchamp and K. Yamazaki, 1987 The genetics of body scent. Trends Genet. **3:** 97–101.

Colombera, D., 1974 Chromosome number within the class Ascidiacea. Mar. Biol. **26:** 63–68.

Elgar, G., 1996 Quality not quantity: the pufferfish genome. Hum. Mol. Genet. **5:** 1437–1442.

Galbraith, D. W., K. R. Harkins, J. M. Madox, N. M. Ayres, D. P. Sharma *et al.*, 1983 Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science **220:** 1049–1051.

Grosberg, R. K., and J. F. Quinn, 1986 The genetic control and consequences of kin recognition by the larvae of a colonial marine invertebrate. Nature **322:** 456–459.

Grosberg, R. K., 1987 Limited dispersal and proximity-dependent mating success in the colonial ascidian, Botryllus schlosseri. Evolution **41:** 372–384.

Lambert, C. C., and C. D. Laird, 1971 Molecular properties of tunicate DNA. Biochim. Biophys. Acta **240:** 39–45.

Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly *et al.*, 1987 MAPMAKER, an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181.

Li, W.-H., and D. Graur, 1991 *Fundamentals of Molecular Evolution*. Sinauer, Sunderland, MA.

Michelmore, R. W., I. Paran and R. V. Kesseli, 1991 Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA **88:** 9828–9832.

Milkman, R., 1967 Genetic and developmental studies on Botryllus schlosseri. Biol. Bull. **132:** 229–243.

Oka, H., and H. Watanabe, 1957 Colony specificity in compound ascidians as tested by fusion experiments. Proc. Japan Acad. Sci. **33:** 657–664.

Oka, H., and H. Watanabe, 1960 Problems of colony specificity in compound ascidians. Bull. Biol. Stn. Asamushi **10:** 153–155.

Olson, M., L. Hood, C. Cantor and D. Botstein, 1989 A common language for physical mapping of the human genome. Science **245:** 1434–35.

Rinkevich, B., R. J. Lauzon, B. W. M. Brown and I. L. Weissman, 1992 Evidence for a programmed life span in a colonial protochordate. Proc. Natl. Acad. Sci. USA **89:** 3546–3550.

Rinkevich, B., Y. Saito and I. L. Weissman, 1993 A colonial invertebrate species that displays a hierarchy of allorecognition responses. Biol. Bull. **184:** 79–86.

Rinkevich, B., R. Porat and M. Goren, 1995 Allorecognition elements on a protochordate histocompatibility locus indicate unprecedented extensive polymorphism. Proc. R. Soc. Lond. B Biol. Sci. **259:** 319–324.

Sabbadin, A., 1962 Le basi genetiche della capacita di fusione fra colonie in *Botryllus schlosseri* (Ascidiacea). Atti Accad. Naz. Lincei. Rend. **32:** 1031–1035.

Sabbadin, A., 1971 Self- and cross-fertilization in the compound ascidian *Botryllus schlosseri*. Dev. Biol. **24:** 379–391.

Saito, Y., E. Hirose and H. Watanabe, 1994 Allorecognition in compound ascidians. Int. J. Dev. Biol. **38:** 237–247.

Satoh, N., and W. R. Jeffrey, 1995 Chasing tails in ascidians: developmental insights into the origin and evolution of chordates. Trends Genet. **11:** 354–359.

Scofield, V. L., J. M. Schlumberger, L. A. West and I. L. Weissman, 1982 Protochordate allorecognition is controlled by a MHC-like gene system. Nature **295:** 499–502.

Schizua, H., B. Birren, U.-J. Kim, V. Mancino, T. Slepak *et al.*, 1992 Cloning and stable maintenance of 300 kilobase pair fragments of human DNA in Escherichhia coli using an F-factor based vector. Proc. Natl. Acad. Sci. USA **89:** 8794–8797.

Stoner, D. S., and I. L. Weissman, 1996 Somatic and germ cell parasitism in a colonial ascidian: possible role for a highly polymorphic allorecognition system. Proc. Natl. Acad. Sci. USA **93:** 15254–15259.

Taylor, K. M., 1967 The chromosomes of some lower chordates. Chromosoma **21:** 181–188.

Tiersch, T. R., R. W. Chandler, S. S. Wachtel and S. Elias, 1989 Reference standards for flow cytometry and application in comparative studies of nuclear DNA content. Cytometry **10:** 706–710.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van De Lee *et al.*, 1995 AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. **23:** 4407–4414.

Watanabe, H., 1962 Further studies on the regulation in fused colonies in *Botryllus primigenus (Ascidiae Compositae).* Sci. Rep. Tokyo Kyoiku Daiqagu Sect. B **10:** 253–259.

Communicating editor: Z-B. Zeng