



半滑舌鳎3D-BAC池构建及性别连锁标记的物理定位

张俊杰^{1,2,3} 张文婷¹ 高峰涛¹ 董忠典¹ 张丽燕¹ 刘琨¹ 邵长伟¹ 陈松林^{1*}

1 中国水产科学研究院 黄海水产研究所, 青岛266071; 2 上海海洋大学 水产与生命学院, 上海201306; 3 新疆农业大学 动物科学学院, 乌鲁木齐830052

*通讯作者, chensl@ysfri.ac.cn

摘要 半滑舌鳎(*Cynoglossus semilaevis*)属于雌配子异形染色体性别决定机制(ZW/ZZ),其W染色体包含的大量重复序列阻碍了W染色体序列的精确组装。目前只有物理图谱可以有效地解决由重复序列所造成的组装难题。本研究在已有物理图谱和遗传图谱的基础上,利用3D(three-dimensional)-BAC(bacterial artificial chromosome)池构建和PCR筛选策略进行了性别连锁标记的物理定位。首先将所选44个384孔板的BAC克隆进行接种和预培养,再接种到4个96孔深孔板中。将每两个384孔微孔板的BAC克隆分别汇集获得相应的板池、行池和列池,在提取DNA后,再将相应板池的部分DNA混合得到超级池。最终构建的3D-BAC池,共包括22个超级池和440个基池,覆盖半滑舌鳎基因组4.2倍,完成一个阳性超级池的基池筛选只需24个反应(20个BAC基池和4个对照)。然后选用159个半滑舌鳎性别连锁标记对这些克隆池进行两步PCR筛选。在对超级池的筛选中,有142个标记获得阳性扩增。对这些阳性超级池进行基池筛选后,最终有100个标记得到准确定位。定位到半滑舌鳎物理图谱的84个重叠群(contigs)和20个单个克隆(singletons)上,共计1760个克隆,包含有21927个共享条带,物理长度为42.4 Mb,覆盖半滑舌鳎基因组约5.3%;确定一个定位关系(contig与标记之间的)平均使用了2.2个克隆。在这些标记中,有17个同时定位到了两个或两个以上的contigs上,有11个标记同时定位到contigs和singletons上。另外有19个contigs定位了2个或2个以上标记。这些性别连锁标记的物理定位将有助于半滑舌鳎性染色体基因组序列的精确组装。

关键词 半滑舌鳎,物理图谱,3D-BAC池,性别连锁标记,物理定位

Construction of 3D-BAC Pools and Physical Location of Sex-linked Markers in Half-smooth Tongue-sole (*Cynoglossus semilaevis*)

ZHANG Jun-Jie^{1,2,3} ZHANG Wen-Ting¹ GAO Feng-Tao¹ DONG Zhong-Dian¹ ZHANG Li-Yan¹
LIU Kun¹ SHAO Chang-Wei¹ CHEN Song-Lin^{1*}

1 Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, China; 2 College of Fisheries and Life Science, Shanghai Ocean University, Shanghai 201306, China; 3 College of Animal Science, Xinjiang Agricultural University, Urumqi 830052, China

* Corresponding author, chensl@ysfri.ac.cn

Abstract The chromosome sex determining mechanism of half-smooth tongue sole (*Cynoglossus semilaevis*) is heteromorphism in female gamete, and the large amount of repetitive sequences, W chromosome contains, impede the accurate assembly of W chromosome sequence, which can only be solved by physical map effectively, at present. Based on the available physical map and genetic map, physical location of sex linked

基金项目: 国家自然科学基金重点项目(No. 31130057)、国家高技术研究发展计划(863)项目(No. 2012AA092203 和 No. 2012AA10A403-2)和山东省泰山学者项目

收稿日期: 2014-06-03 接受日期: 2014-07-17

markers was performed in this study using the strategies of 3D(three-dimensional)-BAC(bacterial artificial chromosome) pool construction and PCR screening. Firstly, the BAC clones of 44 384-well microtiter plates were inoculated and precultivated with new 384-well plates, and then inoculated with four 96 deep-well plates. The plate, row and column pools were obtained through pooling the BAC clone cultures of each two 384-well microtiter plates, and then DNA were extracted and the super pools were obtained by mixing some DNA of corresponding plate pools. Eventually, 3D-BAC pools including 22 BAC super pools and 440 corresponding matrix pools were constructed, which represented 4.2 genome coverage of *C. semilaevis*. Only 24 reactions (20 BAC matrix pools and 4 controls) were needed to complete the screening on matrix pools of a positive super pool. Then two-step PCR screening was performed on these pools with 159 sex-linked markers, and 142 markers obtained positive amplifications in the screening of super pools. Finally, a total of 100 markers were accurately located in the screening of matrix pools of these positive super pools, and they were located in 84 contigs and 20 singletons of the *C. semilaevis* physical map, which included a total of 1 760 clones and 21 927 consensus bands. Their physical length was 42.4 Mb, covering 5.3% of the genome of *C. semilaevis*. On average, 2.2 clones were used to determine a location relationship between a contig and a marker. Among these markers, 17 markers were simultaneously located in 2 or more than 2 contigs, and 11 markers were simultaneously located in contigs and singletons. In addition, 19 contigs were simultaneously located with 2 or more markers. The physical location of these sex-linked markers can promote accurate assembly of sex chromosomes sequences of *C. semilaevis*.

Keywords *Cynoglossus semilaevis*, Physical map, 3D (three-dimensional)-BAC pools, Sex-linked marker, Physical location

半滑舌鲷(*Cynoglossus semilaevis* Günther)属鲽形目(Pleuronectiformes)、舌鲷科(Cynoglossidae)、舌鲷属(*Cynoglossus*), 俗称龙利、鲷米、牛舌头等, 是分布于中国沿海的一种特有比目鱼类(孟庆闻等, 1995; 李思忠, 王惠民, 1995)。半滑舌鲷活动范围小、营养等级低、食性温和, 具有生长快、个体大等一系列优良性状, 而且肉嫩味佳, 市场经济价值很高, 有非常大的养殖推广前景(柳学周等, 2005)。目前半滑舌鲷已实现了全人工育苗和工厂化养殖(雷霖, 2005)。

半滑舌鲷属于雌配子异形染色体性别决定机制(ZW/ZZ)(周丽青等, 2005; Zhuang et al., 2006; Chen et al., 2009), 特别是其雌性生长比雄性快2~3倍, 不但可以开发具有重要应用前景的全雌苗种(Chen et al., 2009), 还可以作为研究硬骨鱼类性别决定机制的理想模型(Shao et al., 2010; Chen et al., 2014)。Chen等(2007; 2008; 2009)利用雌性特异标记发展了半滑舌鲷分子标记辅助性别控制方法, 刘洋等(2014)将筛选到的半滑舌鲷性别特异微卫星标记进行序列特异性扩增区域(sequence characterized amplified regions, SCAR)转化, 并应用于生产实践中伪雄鱼的快速检测。邓思平等(2008)、Deng和Chen(2008)和Deng等(2009)克隆了半滑舌鲷性

别决定区Y-框10(SRY (sex determining region Y)-box 10, *Sox10*)、细胞色素P450芳香化酶(cytochrome P450 aromatase, *P450arom*)和双性和Mab3相关转录因子1 α (doublesex and mab-3-related transcription factor 1 α , *Dmrt1 α*)等性别相关基因, 研究了其时空表达规律。胡乔木等(2013)通过注射*Dmrt1*基因原核表达蛋白的方法研究了其在半滑舌鲷性别决定与分化中的作用。

为了深入理解半滑舌鲷性别决定机制, 其性别相关的基因组学研究也得到不断发展。Liao等(2009)构建了半滑舌鲷第一代遗传连锁图谱, 并将一个雌性特异标记定位到该图谱上。Song等(2012)构建了半滑舌鲷高密度遗传连锁图谱, 包括21个连锁群, 并将筛选出的159个半滑舌鲷性别连锁标记全部定位在雌性图谱1号连锁群的同一位点。半滑舌鲷全基因组测序揭示了半滑舌鲷性染色体的结构和进化, 并发现半滑舌鲷W染色体包含大量的重复序列, 这些重复序列阻碍了W染色体序列的精确组装(Chen et al., 2014)。目前只有物理图谱可以有效地解决由重复序列所造成的组装难题(Lewin et al., 2009), Shao等(2010)构建了两个雌性半滑舌鲷细菌人工染色体(bacterial artificial chromosome, BAC)文

库, Zhang等(2014)利用该BAC文库构建了半滑舌鳎基因组物理图谱, 为半滑舌鳎基因组特别是W染色体序列的精确组装奠定了基础。

本研究采取改进3D-pooling(three-dimensional pooling)策略, 构建了半滑舌鳎3D-BAC池, 并利用半滑舌鳎性别连锁标记对这些克隆池进行了PCR筛选, 以通过这些标记的物理定位为半滑舌鳎性染色体序列的精确组装奠定基础。

1 材料与方法

1.1 实验材料

1.1.1 半滑舌鳎物理图谱和BAC文库

本研究采用Shao等(2010)分别在pECBAC1质粒的BamH I和Hind III酶切位点构建两个雌性半滑舌鳎(*Cynoglossus semilaevis* Günther)BAC文库(总计55 296个克隆, 分布于144个384孔板中, 保存于-80°C超低温冰箱中)和Zhang等(2014)利用上述文库中的88个384孔微孔板(共33 575个克隆)构建的半滑舌鳎第一代物理图谱(该图谱含有1 485个重叠群(contigs)和585个单个克隆(singletons))用于标记的物理定位。从该物理图谱构建所用88个384孔板中, 选取44个384孔板(共16 896个克隆)用于构建3D-BAC池, 覆盖半滑舌鳎基因组约4.2倍。

1.1.2 主要设备和试剂

所用设备主要有384针和96针微孔板复制器(Boekel, 美国)、高速冷冻离心机(Beckman, 美国)、PCR仪(AB, 美国)、电泳仪(君意东方, 北京)和凝胶成像系统(Bio-Rad, 美国)等; 实验试剂主要有胰蛋白酶和酵母提取物(Oxoid, 英国)、Taq DNA聚合酶(Fermentas, 美国)、琼脂糖G-10(Biowest, 西班牙)和GelRed核酸凝胶染料(Biotium, 美国)等。

1.2 实验方法

1.2.1 BAC克隆的培养

参照Zhang等(2014)的方法进行BAC克隆的培养。用384针微孔板复制器将BAC克隆从储备384孔板接种到新384孔板内, 每孔内含60 μ L 2 \times YT培养基(内含12.5 μ g/mL氯霉素)。经过24 h预培养后, 用96针微孔板复制器接种到96孔深孔板中, 每孔内含1.3 mL 2 \times YT培养基(内含12.5 μ g/mL氯霉素), 每个384孔板接种4个96孔板。用透气封板膜封板

后置于37°C摇床上振荡(300 r/min)培养26~28 h。

1.2.2 3-D BAC池的构建

3D-BAC池的构建参照Bouzidi等(2006)的方法并作适当修改, 所构建的3D-BAC池包括超级池和基池两部分。其中超级池共有22个, 每个超级池由2个384孔板组成, 共768个克隆, 覆盖舌鳎基因组约0.19倍, 全部超级池所用克隆的来源情况见表1。每个超级池含有20个基池, 包括2个板池、8行池和10个列池。这些基池的具体编排方式和克隆来源见表2。

采用碱裂解法提取各个基池的DNA(Russell, Sambrook, 2001), 用紫外分光光度计测定各基池DNA的浓度和纯度。最后取出各超级池内的两个板池的部分DNA合并, 作为该超级池的BAC DNA。

1.2.3 标记选择和引物合成

本研究所用的159个半滑舌鳎性别连锁微卫星标记由Song等(2012)鉴定, 并被区分为两种类

表1 半滑舌鳎3D-BAC池的超级池克隆来源

Table 1 Clone sources of super pools in half-smooth tongue-sole 3D-BAC pools

超级池编号	384孔板编号(P1和P2)
Super pool ID	384-well plate ID (P1 and P2, respectively)
S1	061, 062
S2	063, 064
S3	065, 066
S4	067, 068
S5	069, 070
S21	101, 102
S22	103, 104
S23	105, 106
S24	107, 108
S25	109, 110
S31	002, 122
S32	123, 124
S33	125, 126
S34	127, 128
S35	129, 130
S36	131, 132
S37	133, 134
S38	135, 136
S39	137, 138
S40	139, 140
S41	141, 142
S42	143, 144

型:一种为雌鱼双带、雄鱼单带;另一种为雌鱼单带、雄鱼无带,其在遗传作图所用群体中的基因型完全一致,全部定位在雌性图谱1号连锁群的同一点。标记引物由北京华大基因研究中心合成,引物浓度控制在10 $\mu\text{mol/L}$ 。

1.2.4 PCR筛选

采用两步PCR筛选法对159个性别连锁标记进行物理定位。第一步,以雌雄半滑舌鳎基因组为对照,对22个超级池进行筛选;第二步,以雌雄半滑舌鳎基因组和双蒸水(ddH_2O)为对照,对阳性超级池的20个基池进行筛选,如果阳性超级池为4个或4个以下,则对全部阳性超级池进行基池筛选,如果阳性超级池为4个以上,则一般只选其中阳性最为显著的4个进行基池筛选。PCR反应体系:10 \times buffer 2 μL , MgCl_2 (25 mmol/L) 0.8 μL , dNTPs (2.0 mmol/L) 1.3 μL , 上下游引物(10 $\mu\text{mol/L}$)各0.25 μL , *Taq* DNA聚合酶0.8 U, BAC DNA(约100 ng/ μL) 2 μL , 以 ddH_2O 调整至20 μL 。PCR反应程序:95 $^\circ\text{C}$ 预变性5 min,然后扩增35个循环(95 $^\circ\text{C}$ 变性30 s, 55 $^\circ\text{C}$ 退火30 s, 72 $^\circ\text{C}$ 延伸30 s),最后72 $^\circ\text{C}$ 再延伸5 min,置于4 $^\circ\text{C}$ 冰箱备用。

1.2.5 琼脂糖凝胶电泳鉴定扩增产物

对PCR扩增产物进行1.0%的琼脂糖凝胶电泳,利用凝胶成像系统拍照观察,以是否出现与阳性对照(雌鱼基因组)相同的条带来判断是否属于阳性池。

1.2.6 阳性克隆的推导与定位

根据对超级池和基池进行的两步PCR筛选结果,结合所采用的3D-BAC池构建方式,可推导出每个性别连锁标记的阳性克隆编号。在组装物理图谱所用FPC软件9.4版(Soderlund et al., 1997)的主界面,对所获得的各个阳性克隆进行搜索,确定其所在的contig,完成该标记的物理定位。

2 结果与分析

2.1 3D-BAC池构建

按照预定建池方案,将所选44个384孔板的BAC克隆接种培养,汇集获得各个超级池的行池、列池和板池,在提取DNA后,再将相应板池的部分DNA混合得到超级池。最终构建的3D-BAC池共包括22个超级池和440个基池。表2显示了一个超

级池(S37)内各基池DNA浓度(*OD*)和纯度(A_{260}/A_{280})的测定结果情况。本研究预实验PCR扩增中所要求的BAC DNA浓度最低约为0.3 ng/ μL ,所对应的基池、行池和列池的总浓度应该不低于115、58和48 ng/ μL 。从表2可以看出,本研究所建的3D-BAC池中各基池DNA浓度能够满足PCR筛选的需要。

2.2 两步PCR筛选反应

用性别连锁标记对22个超级池进行PCR筛选后,可得到含有该标记的阳性超级池。图1所示,性别连锁标记 scaffold2722_838 对22个超级池的进行筛选,共有3个超级池(S24、S31和S37)扩增出与阳性对照(雌性基因组,以F表示)一致的条带。同样作为对照的雄性基因组(以M表示)没有出现对应的阳性条带,说明该标记为雌鱼单带、雄鱼无带类型。

获得阳性超级池后,就可以对阳性超级池的基池进行筛选。如对标记 scaffold2722_838 的3个阳性超级池(S24、S31和S37)分别进行了筛选基池,并获得了比较理想的结果,其中对S31筛得的阳性基池依次为:P2、R4、R6、C4和C8(图2A);对S37筛得的阳性基池依次为:P1、R3、R6、C5和C8(图2B)。

2.3 阳性克隆的推导

基于获得的各超级池的阳性基池情况,根据本研究采用的3D-BAC池构建策略来推导各阳性超级池内的阳性克隆。如 scaffold2722_838 对S31的阳性基池为P2、R4、R6、C4和C8,通过P2得知此克隆来自于122号384孔板,从R4和R6得知此克隆来自于第N行,从C4和C8得知此克隆来自于第18列,所以该阳性克隆为122N18。同理可以根据 scaffold2722_838 对S37扩增的阳性基池(P1、R3、R6、C5和C8)推得阳性克隆为133J23。根据对S24筛选结果推得的阳性克隆是108B20。

2.4 阳性克隆的物理定位

在组装物理图谱所用FPC软件的主界面,将阳性克隆编号输入进行搜索,即可进行该标记的物理定位。通过搜索阳性克隆122N18得到一个重叠群(ctg853),他包含有该阳性克隆。而通过搜索133J23和108B20也得到了同一个重叠群(ctg853)。图3显示了利用这3个阳性克隆完成的 scaffold2722_838 物理定位结果。从图3中可以看出,这3个阳性克隆不但定位到了物理图谱中的同

表 2 一个超级池的基池构成和质粒DNA浓度

Table 2 The composition and plasmid DNA concentration of the matrix pools of one super pool

池类别	池编号	克隆来源	OD/(ng·μL ⁻¹)	A ₂₆₀ /A ₂₈₀
Pool type	Pools ID	Clone sources		
板池	P1	第一个384孔板 The first 384-well plate	236	1.778
Plate pools	P2	第二个384孔板 The second 384-well plate	231	1.791
行池	R1	A, B, C, D	127	1.752
Row pools	R2	E, F, G, H	111	1.729
	R3	I, J, K, L	107	1.698
	R4	M, N, O, P	158	1.727
	R5	A, E, I, M	116	1.731
	R6	B, F, J, N	148	1.772
	R7	C, G, K, O	99	1.707
	R8	D, H, L, P	180	1.782
列池	C1	1, 2, 3, 4, 5	115	1.750
Column pools	C2	6, 7, 8, 9, 10	80	1.713
	C3	11, 12, 13, 14, 15	95	1.727
	C4	16, 17, 18, 19, 20	91	1.701
	C5	21, 22, 23, 24	139	1.794
	C6	1, 6, 11, 16, 21	96	1.761
	C7	2, 7, 12, 17, 22	100	1.724
	C8	3, 8, 13, 18, 23	103	1.725
	C9	4, 9, 14, 19, 24	139	1.766
	C10	5, 10, 15, 20	101	1.781



图 1 scaffold2722_838对22个超级池进行PCR筛选

Figure 1 PCR screening of 22 super pools with scaffold2722_838

M1: 分子量标准 DL2000; F: 雌鱼基因组DNA(对照); M: 雄鱼基因组DNA(对照); 扩增阳性的BAC池用下划线注明, 其余为扩增阴性的BAC池; 下同

M1: DL2000 marker; F: Genomic DNA of females(control); M: Genomic DNA of males (control); BAC pools with positive amplification are underlined, others were negative ; The same below

一个重叠群中, 而且其在该重叠群中的位置也非常靠近。

2.5 半滑舌鲷性别连锁标记的物理定位

本研究用159个半滑舌鲷性别连锁微卫星标记分别对所构建的3D-BAC池进行了PCR筛选, 经过第一步对超级池的筛选后, 有142个标记获得阳性超级池, 对这些阳性超级池进行基池筛选, 共有100个标记得到准确定位(表3)。表4物理定位结果表明, 分别定位到84个contigs和20个singletons

上, 总计1760个克隆, 包括21927个共享条带, 物理长度为42.4 Mb, 覆盖半滑舌鲷基因组约5.3%。在这些标记中, 有17个同时定位到了两个或两个以上的contigs上, 有11个标记同时定位到contigs和singletons上去。另外有19个contigs定位了2个或2个以上标记。在定位过程中, 定位重叠群所用的阳性克隆共计247个, 重叠群与标记之间的定位关系总数为112, 由此可知确定一个定位关系(一个contig与一个标记之间)平均使用了2.2个克隆。

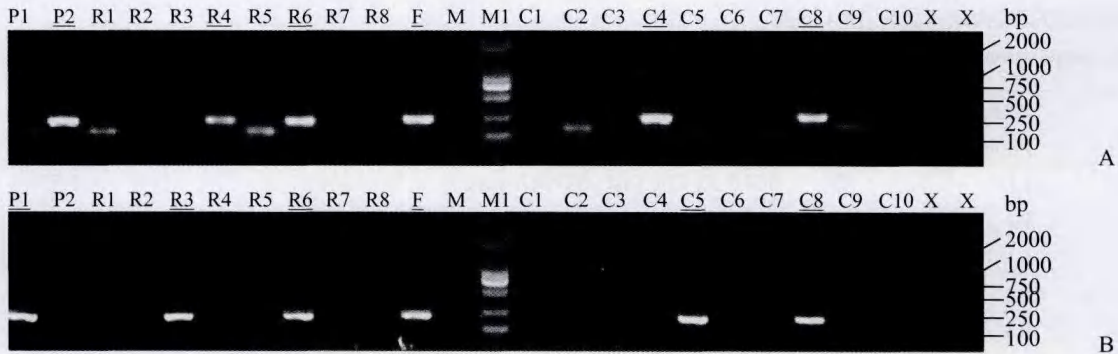


图2 用scaffold2722_838对S31(A)和S37(B)的基池进行PCR筛选

Figure 2 PCR screening of the matrix pools of S31(A) and S37(B) with scaffold2722_838

X: 双蒸水(空白对照) X: ddH₂O (blank control)

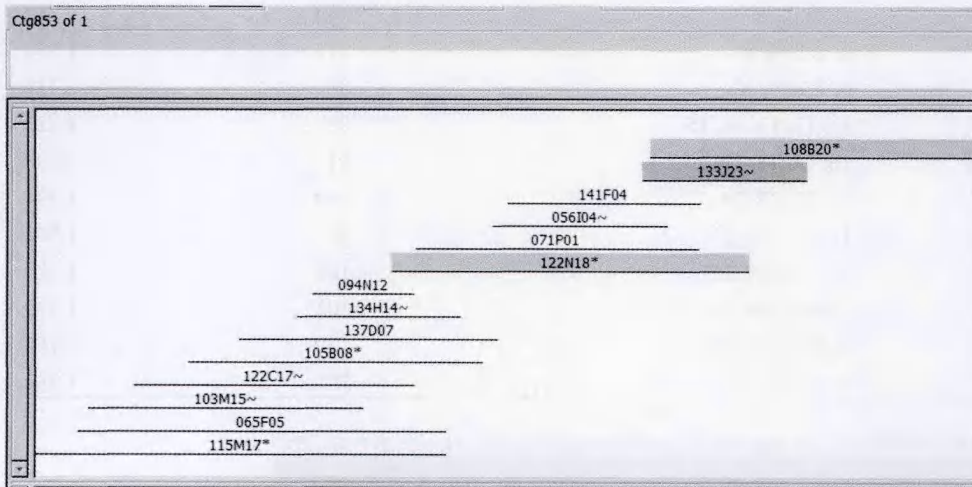


图3 scaffold2722_838通过3个阳性克隆(122N18、133J23和108B20)进行的物理定位

Figure 3 The physical location of scaffold2722_838 with 3 positive clones (122N18, 133J23 and 108B20)

3 讨论

3.1 多维BAC池的构建策略选择

DNA 标记在物理图谱中的定位可以通过对BAC文库的筛选来完成,筛选方法主要有滤膜杂交法和多维池PCR筛选法(Bouzidi et al., 2006; Madishetty et al., 2007)。由于滤膜杂交法假阳性信号多,且需要利用放射性同位素,使得多维池PCR筛选法优势明显并得到广泛应用(Ariyadasa, Stein, 2012)。

多维BAC池的构建是多维池PCR筛选法的基础,其构建策略也在不断地优化发展。简单三维BAC池是最早出现并经常使用的建池策略,其将一定数量的384孔板合并形成超级池,每个超级池的基池部分一般包括16个行池、24个列池和若干个板池,完成一个超级池的基池筛选所需要的PCR反应数较多(孙建, 2013; Bruno et al., 1995),而且使

用过程中也容易出现假阳性干扰(You et al., 2010; Ariyadasa, Stein, 2012)

近来提出一种改进3D-BAC建池策略,将若干行、列或板分别汇集1次成池后,再交错进行2次汇集,形成第2套基池,覆盖超级池2次,最终不但减少了基池数量,使基池筛选反应减少,而且用双阳性信号来减少了假阳性信号的干扰(Bouzidi et al., 2006)。本研究即参考这种策略构建了半滑舌鲷3D-BAC池。

3.2 本研究3D-BAC建池策略的效果分析

利用本研究所构建的3D-BAC池,完成一个标记对覆盖基因组4.2倍的超级池筛选需要24个反应(22个BAC基池和2个对照),完成一个标记对一个阳性超级池的基池筛选也只需24个反应(20个BAC基池和4个对照),而且依靠双信号来断推基

表3 在半滑舌鳎物理图谱中得到准确定位的100个性别特异标记

Table 3 100 sex-specific markers accurately located in half-smooth tongue sole physical map

标记名称	物理定位	标记名称	物理定位	标记名称	物理定位
Marker name	Physical location	Marker name	Physical location	Marker name	Physical location
scaffold256_49	ctg220 [#]	scaffold2081_663	ctg119	scaffold3982_1057	ctg1762
scaffold2999_952	ctg783 [#]	scaffold2133_705	ctg1836 [#]	scaffold4113_1070	ctg3604 [#]
scaffold3067_958	ctg137 [#]	scaffold2148_678	ctg4076 [#]	scaffold4119_1071	ctg837
scaffold3391_1010	ctg1939	scaffold2164_693	ctg755	scaffold4172_1073	ctg161
scaffold1024_286 [*]	ctg3202 [#] /061M19 [#]	scaffold2360_734 [*]	ctg3924/ctg385 [#]	scaffold4195_1075	ctg2766 [#]
scaffold1030_289	ctg1025	scaffold2431_777	ctg2636	scaffold422_81	130C02
scaffold1038_293 [*]	ctg996/ctg1616	scaffold2453_780	125C02	scaffold4321_1086	ctg433
scaffold1111_312	ctg915	scaffold2484_785	ctg1276	scaffold4466_1096	ctg104 [#]
scaffold1113_315 [*]	ctg1347/ctg1359 [#]	scaffold2527_791 [*]	ctg853 [#] /ctg322 [#]	scaffold4665_1103	ctg322 [#]
scaffold1128_343	ctg3559	scaffold2532_794 [*]	133G17/002J14	scaffold484_90 [*]	ctg1527/ctg2884
scaffold1220_378	ctg244	scaffold2565_799	ctg1669	scaffold513_107	069G19
scaffold1246_388	ctg1329	scaffold2566_800	ctg145	scaffold522_110	ctg4064
scaffold1332_411 [*]	ctg1560 [#] /ctg584	scaffold259_54	136I14	scaffold543_118	ctg4076 [#]
scaffold1339_414	ctg3	scaffold2607_804 [*]	ctg165/ctg104 [#] /ctg2065	scaffold552_122	ctg3382 [#]
scaffold1354_430 [*]	ctg3382 [#] /136O23	scaffold2618_811	ctg287 [#]	scaffold56_4	ctg323
scaffold1358_427 [*]	ctg472/062L22	scaffold2719_837	ctg1488	scaffold571_123	128D20
scaffold1392_450 [*]	ctg937/ctg2596	scaffold2722_838	ctg853 [#]	scaffold757_179	ctg1722
scaffold1423_459	ctg783 [#]	scaffold2744_845 [*]	ctg139/ctg1436	scaffold804_188	ctg2305
scaffold1447_469 [*]	ctg592/ctg1206	scaffold2746_846 [*]	141F24/064A06	scaffold813_200 [*]	ctg2345 [#] /104A06 [#]
scaffold1500_508	070E12	scaffold2950_942 [*]	ctg615/ctg93/ctg1241	scaffold868_211	ctg853 [#]
scaffold1503_484	ctg385 [#]	scaffold2957_945	ctg581 [#]	scaffold869_212 [*]	ctg1083/ctg267
scaffold1545_522	ctg532	scaffold3003_954	ctg2345 [#]	scaffold885_215 [*]	ctg3604 [#] /142C07
scaffold1573_527	ctg83	scaffold3032_956 [*]	ctg1359 [#] /ctg2094	scaffold899_221	ctg1384 [#]
scaffold1611_534 [*]	ctg2402/ctg360/ctg287 [#]	scaffold3064_962	130N09	scaffold1283_399 [*]	126K04/134O17
scaffold1614_544	ctg2482	scaffold3068_959 [*]	ctg1560 [#] /ctg104 [#]	scaffold1590_532	ctg153
scaffold1653_550 [*]	ctg2345 [#] /104A06 [#]	scaffold3080_963	ctg744	scaffold1890_606	ctg1384 [#]
scaffold1657_558	ctg220 [#]	scaffold3360_1002 [*]	ctg1934/ctg271	scaffold1951_619 [*]	ctg3604 [#] /121J14
scaffold1658_561 [*]	ctg137 [#] /ctg2766 [#]	scaffold3405_1014	ctg3940	scaffold2195_702	ctg144
scaffold1676_567	ctg3123	scaffold3424_1020 [*]	ctg994/ctg385 [#]	scaffold2211_718	ctg3446
scaffold1905_607	ctg3382 [#]	scaffold3487_1025	ctg1836 [#]	scaffold2322_727	ctg423
scaffold1941_615	ctg116	scaffold3726_1044	ctg403	scaffold2372_736	ctg1684
scaffold1985_624 [*]	ctg581 [#] /ctg656	scaffold386_75	125F03	scaffold2491_788	ctg1384 [#]
scaffold1991_626 [*]	ctg3202 [#] /122O11/061M19 [#]	scaffold3922_1053	ctg104 [#]		
scaffold1993_631	ctg833	scaffold393_77	ctg2259		

*: 一个标记定位到不同的 contigs 或 singletons; #: 该 contigs 或 singletons 上有多个标记定位

*: One marker mapped to different contigs or singletons; #: Several markers mapped to the contigs or singletons

池的阳性克隆。而用传统3D-BAC池方法进行同样尺度的筛选,则可能需要46个反应完成板池筛选,44个反应完成基池筛选,共需要90个反应(孙建,2013)。本研究所采用的改进3D-BAC建池策略从操作效率和可靠性两个方面提高了筛选效率。

但在本研究进行3D-BAC池的筛选过程中,仍

然存在假阳性和假阴性问题,使得实际的筛选效率仍然受到影响。例如用标记scaffold1503_484对超级池筛选得到了7个阳性,而基池筛选时有2个超级池只有对照(雌鱼基因组)获得阳性,说明这两个超级池为假阳性。在另外5个阳性超级池的基池筛选中,只有2个得到了唯一阳性克隆,能够准确

表 4 半滑舌鲷性别连锁标记的物理定位

Table 4 Physical location of half-smooth tongue sole sex-linked markers

项目 Items	数量 Numbers
性别连锁标记数 The number of sex-linked markers	159
第一次 PCR 筛选的阳性标记数 The number of positive markers in the first PCR screening	142
得到准确定位的标记数 The number of markers accurately located	100
共享条带数量(CB)The number of consensus bands (CB)	21927
物理长度/MbPhysical length	42.4
占总基因组比例 The proportion in total genome	5.3%
已定位的 contig 数 The number of located contigs	84
已定位的克隆数 The number of located clones	1760
已定位的 singleton 数 The number of located singletons	20
用于定位的克隆数 The number of clones used to locate	247
Contig 和标记之间定位关系系数 The number of location relationship between contigs and markers	112
确定一个 contig 和标记之间的定位关系所用平均克隆数 Average clones used for every location relationship between contigs and markers	2.2
定位多个标记的 contig 数 The number of contigs located with multiple markers	19
定位到多个 contig 或 singleton 的标记数 The number of markers located in multiple contigs or singletons	28

地定位;有 2 个出现假阴性,各缺少 1 个阳性条带,形成有 5 个可能的克隆;另 1 个多了 1 个阳性结果,有 2 个可能的克隆。在准确定位的两个克隆所在的 contig 中刚好能够找到这 3 个超级池的各自几种可能克隆中的 1 个,从而确定其阳性克隆。

由于本研究进行基池筛选时,如果阳性超级池少于 4 个,就全部进行筛选;如果阳性超级池多于 4 个,则一般只选其中 4 个进行筛选,再加上基池筛选中假阳性和假阴性结果的干扰,本研究最后对 100 个标记进行了准确定位,共形成了 112 个定位关系,确定一个定位关系平均使用了 2.2 个克隆,用于定位的阳性克隆有 247 个,而没有达到理论上根据基因组覆盖率所应该有的 420 阳性克隆。

在 3D-BAC 池的构建和筛选过程中可能存在一些影响建池质量和筛选效果的因素,包括:(1)在文库正式接种前虽然经过了一次预培养,但得到的 BAC 细菌浓度仍然存在一定的差别;(2)在 BAC 文库保存、接种、汇集以及质粒提取过程中可能出现相互污染或者缺失情况;(3)引物特异性、扩增效率和循环次数等因素也会影响阳性结果的出现情况。

3.3 物理定位区域的覆盖率

关于半滑舌鲷 W 染色体序列的大小,不同学者利用不同的实验手段也得到的结果也不尽相同。Wang 等(2009)采用染色体显微切割和 DOP(degener-

ate oligonucleotide primed)-PCR 方法分离和扩增 DNA 后,构建出了 W 染色体特异 DNA 文库,测序获得 29.4 Mb 序列。Chen 等(2014)用全基因组测序得到的半滑舌鲷雌鱼基因组序列扣除雄鱼基因组序列后得到 16.4 MB 的雌性特异基因组序列。但 W 染色体作为半滑舌鲷 21 对染色体中最大的一条,其基因组序列据估计为 70 Mb 左右,约为占总基因组的 10% 以上(周丽青等, 2005; Zhuang et al., 2006; Chen et al., 2014)。本研究中有 100 个标记得到了准确定位,其所在的 84 个 contigs 及 20 个 singletons 总物理长度为 42.4 Mb,覆盖基因组约 5.3%,应该还有部分属于 W 染色体的 contigs 尚未通过这种方法得到确认,这可能需要扩大 3D-BAC 池的覆盖率或者增加更多的标记来进行筛选和定位。

3.4 标记定位与物理图谱验证

由于本研究采用的定位方法操作方便,所以一般选用 4 个阳性超级池进行基池筛选。scaffold2722_838 作为本研究所用的标记之一,用 PCR 方法筛选到的 3 个克隆都定位到 ctg853 上(图 3),而且 3 个克隆在该 contig 中的位置靠近。同一个标记在不同的克隆中得到阳性扩增,表明这两个克隆具有相同的片段(Zhang et al., 2014)。这样可以通过定位于同一 contig 内不同的克隆来对其定位结果进行相互印证,也进一步证明了物理图谱组装的可靠性。

3.5 标记定位与物理图谱长度

在定位的标记中,有17个定位到了两个或两个以上的contigs上,另有11个标记同时定位到contig和singleton上去。这可能是因为经过FPC软件自动组装得到的1485个contigs及586个singletons之间仍然存在着一定的重叠关系,只是在组装时为保证可靠性而没有采用更低的严谨度进行合并,并由此导致物理图谱的组装长度明显大于基因组实际长度(Shao et al., 2010; Zhang et al., 2014)。另外本研究所选用的性别连锁标记全部为SSR标记(Song et al., 2012),而W染色体本身就存在大量重复序列(Chen et al., 2014),可能部分标记在基因组上本身就有不只1个位置,从而导致这些标记定位到不同的contigs上。这些有待进一步通过contigs的人工合并和这些标记在染色体上定位来确定。

3.6 物理定位与遗传定位的不相关性

本研究所选的159个性别连锁标记全部定位到半滑舌鳎高密度遗传连锁图谱中W染色体的同一个位点(Song et al., 2012),说明这些标记所在的基因组区域在细胞分裂时几乎没有发生重组交换,导致在这个区域上的物理图谱与遗传图谱之间呈现出无关性。而这也正说明这些区域可能是异染色质区域,非常缺少基因,由相对简单、高度重复的DNA序列组成(翟中和等, 2007),因交换频率减少而导致遗传距离缩短(徐晋麟等, 2011)。另外鉴于许多物种的性染色体完全是由异染色质形成的异染色体(韩贻仁, 2007),可进一步推测这100个标记所在的84个contigs及20个singletons处在半滑舌鳎的性染色体(Z/W)上。未来将通过BAC-荧光原位杂交(fluorescence *in situ* hybridization, FISH)等手段进一步验证这些contigs在W染色体上的定位,并最终将这些定位结果应用于半滑舌鳎性染色体序列的精确组装工作中。

4 结论

本研究在已有物理图谱和遗传图谱的基础上,首先利用44个384孔板的BAC克隆,构建了覆盖半滑舌鳎基因组4.2倍的3D-BAC池,该3D-BAC池包括22个超级池,每个超级池有20个基池,共440个基池。然后对这些克隆池进行了性别连锁标记的PCR筛选,最终准确定位了100个标记,定位到半滑舌鳎物理图谱的84个contigs和20个sin-

gletons上,共计1760个克隆,物理长度为42.4 Mb,覆盖半滑舌鳎基因组约5.3%。这些性别连锁标记的物理定位将有助于半滑舌鳎性染色体基因组序列的精确组装。

参考文献

- 邓思平,陈松林. 2008. 半滑舌鳎*Dmrt1 α* 基因的cDNA克隆及其表达[J]. 中国水产科学, 15(4): 577-584. (Deng S P, Chen S L. 2008. Molecular cloning, characterization and RT-PCR expression analysis of *Dmrt1 α* from half-smooth tongue-sole, *Cynoglossus semilaevis*[J]. Journal of Fishery Sciences of China, 15(4): 577-584.)
- 韩贻仁, 主编. 2007. 分子细胞生物学(第三版)[M]. 高等教育出版社, 中国, 北京, pp. 298-343. (Han Y R, ed. 2007. Molecular Cell Biology (3rd ed.)[M]. Higher Education Press, Beijing, China, pp. 298-343.)
- 胡乔木,王凯琳,陈松林. 2013. 半滑舌鳎*Dmrt1*蛋白表达、纯化及功能[J]. 中国水产科学, 20(6): 1132-1138. (Hu Q M, Wang K L, Chen S L, 2013. Protein expression, purification, and elementary function of *Dmrt1* in half-smooth tongue sole (*Cynoglossus semilaevis*) [J]. Journal of Fishery Sciences of China, 20(6): 1132-1138.)
- 雷霖霖, 主编. 2005. 海水鱼类养殖理论与技术[M]. 中国农业出版社, 中国, 北京, pp. 647-665. (Lei Q L, ed. 2005. Theory and Technology of Marine Fish Culture[M]. China Agriculture Press, Beijing, China, pp. 647-665.)
- 李思忠,王惠民. 1995. 中国动物志: 硬骨鱼纲鲷形目[M]. 科学出版社, 中国, 北京, pp. 334-366. (Li S Z, Wang H M, 1995. Fauna Sinica, Osteichthyes, Pleuronectiformes [M]. Science Press, Beijing, China, pp. 334-366.)
- 柳学周,庄志猛,马爱军,等. 2005. 半滑舌鳎繁殖生物学及繁育技术研究[J]. 海洋水产研究, 26(5): 7-14. (Liu X Z, Zhuang Z M, Ma A J, et al. 2005. Reproductive biology and breeding technology of *Cynoglossus semilaevis* Günther[J]. Marine Fisheries Research, 26(5): 7-14.)
- 刘洋,陈松林,高峰涛,等. 2014. 半滑舌鳎(*Cynoglossus semilaevis*)性别特异微卫星标记的SCAR转化及其在伪雄亲鱼快速检测中的应用[J]. 农业生物技术学报, 22(6): 787-792. (Liu Y, Chen S L, Gao F T, et al. 2014. SCAR-transformation of sex-specific SSR marker and its application in neo-male parent half-smooth tongue sole(*Cynoglossus semilaevis*)fast test[J]. Journal of Agricultural Biotechnology, 22(6): 787-792.)
- 孟庆闻,苏绵祥,缪学祖. 1995. 鱼类分类学[M]. 中国农业出版社, 中国, 北京, pp. 979-981. (Men Q W, Su J X, Miao X Z, 1995. Systematics of Fishes[M]. China Agriculture Press, Beijing, China, pp. 979-981.)

- 孙建. 2013. 半滑舌鳎BAC文库筛选及相关基因的FISH定位[D]. 硕士学位论文, 大连海洋大学, 导师: 陈松林, pp. 14-27. (Sun J. 2013. Screening of half-smooth tongue sole BAC library (*Cynoglossus semilaevis*) and FISH locating of related genes[D]. Thesis for M.S., Dalian Ocean University, Supervisor: Chen S L, pp. 14-27.)
- 徐晋麟, 徐沁, 陈淳. 2011. 现代遗传学原理(第三版)[M]. 科学出版社, 中国, 北京, pp. 39-55. (Xu J L, Xu Q, Chen C. 2011. Principles of Modern Genetics (3rd ed.)[M]. Science Press, Beijing, China, pp. 39-55.)
- 翟中和, 王喜忠, 丁明孝, 主编. 2007. 细胞生物学(第三版)[M]. 高等教育出版社, 中国, 北京, pp. 307-365. (Zhai Z H, Wang X Z, Ding M X, ed. 2007. Cell Biology (3rd ed.)[M]. Higher Education Press, Beijing, China, pp. 307-365.)
- 周丽青, 杨爱国, 柳学周, 等. 2005. 半滑舌鳎染色体核型分析[J]. 水产学报, 29(3): 417-419. (Zhou L Q, Yang A G, Liu X Z, et al. 2005. The karyotype of the tonguefish *Cynoglossus semilaevis* [J]. Journal of Fisheries of China, 29(3): 417-419.)
- Ariyadasa R, Stein N. 2012. Advances in BAC-based physical mapping and map integration strategies in plants[J]. Journal of Biomedicine and Biotechnology, 2012: 184854.
- Bouzidi M F, Franchel J, Tao Q, et al. 2006. A sunflower BAC library suitable for PCR screening and physical mapping of targeted genomic regions[J]. Theoretical and Applied Genetics, 113(1): 81-89.
- Bruno W J, Knill E, Balding D J, et al. 1995. Efficient pooling designs for library screening[J]. Genomics, 26(1): 21-30.
- Chen S L, Deng S P, Ma H Y, et al. 2008. Molecular marker-assisted sex control in half-smooth tongue sole (*Cynoglossus semilaevis*)[J]. Aquaculture, 283(1): 7-12.
- Chen S L, Li J, Deng S P, et al. 2007. Isolation of female-specific AFLP markers and molecular identification of genetic sex in half-smooth tongue sole (*Cynoglossus semilaevis*)[J]. Marine Biotechnology, 9(2): 273-280.
- Chen S L, Tian Y S, Yang J F, et al. 2009. Artificial gynogenesis and sex determination in half-smooth tongue sole (*Cynoglossus semilaevis*)[J]. Marine Biotechnology, 11(2): 243-251.
- Chen S L, Zhang G J, Shao C W, et al. 2014. Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle[J]. Nature Genetics, 46: 253-260.
- Deng S P, Chen S L. 2008. cDNA cloning, tissues, embryos and larvae expression analysis of Sox10 in half-smooth tongue-sole, *Cynoglossus semilaevis*[J]. Marine Genomics, 1(3): 109-114.
- Deng S P, Chen S L, Xu J Y, et al. 2009. Molecular cloning, characterization and expression analysis of gonadal P450 aromatase in the half-smooth tongue-sole, *Cynoglossus semilaevis*[J]. Aquaculture, 287(1): 211-218.
- Lewin H A, Larkin D M, Pontius J, et al. 2009. Every genome sequence needs a good map[J]. Genome Research, 19(11): 1925-1928.
- Liao X L, Ma H Y, Xu G B, et al. 2009. Construction of a genetic linkage map and mapping of a female-specific DNA marker in half-smooth tongue sole (*Cynoglossus semilaevis*)[J]. Marine Biotechnology, 11(6): 699-709.
- Madishetty K, Condamine P, Svensson J T, et al. 2007. An improved method to identify BAC clones using pooled overgos[J]. Nucleic Acids Research, 35(1): e5.
- Russell D W, Sambrook J. 2001. Molecular Cloning: A Laboratory Manual (3rd ed.)[M]. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA, pp. 1-68.
- Shao C W, Chen S L, Scheuring C F, et al. 2010. Construction of two BAC libraries from half-smooth tongue sole *Cynoglossus semilaevis* and identification of clones containing candidate sex-determination genes[J]. Marine Biotechnology, 12(5): 558-568.
- Soderlund C, Longden I, Mott R. 1997. FPC: A system for building contigs from restriction fingerprinted clones[J]. Computer Applications in the Biosciences: CABIOS, 13(5): 523-535.
- Song W T, Li Y Z, Zhao Y W, et al. 2012. Construction of a high-density microsatellite genetic linkage map and mapping of sexual and growth-related traits in half-smooth tongue sole (*Cynoglossus semilaevis*) [J]. PloS one, 7(12): e52097.
- You F M, Luo M C, K Xu, et al. 2010. A new implementation of high-throughput five-dimensional clone pooling strategy for BAC library screening[J]. BMC genomics, 11: 692.
- Wang X B, Zhang Q Q, Ren J, et al. 2009. The preparation of sex-chromosome-specific painting probes and construction of sex chromosome DNA library in half-smooth tongue sole (*Cynoglossus semilaevis*)[J]. Aquaculture, 297(1): 78-84.
- Zhang J J, Shao C W, Zhang L Y, et al. 2014. A first generation BAC-based physical map of the half-smooth tongue sole (*Cynoglossus semilaevis*) genome[J]. BMC genomics, 15: 215.
- Zhuang Z M, Wu D, Zhang S C, et al. 2006. G-banding patterns of the chromosomes of tonguefish *Cynoglossus semilaevis* Günther, 1873[J]. Journal of Applied Ichthyology, 22(5): 437-440.