

海带3个着丝粒蛋白编码序列的克隆及特征分析

徐梅, 周志刚*

(上海海洋大学水产与生命学院, 上海 201306)

摘要: 本研究根据海带基因组及糖海带转录组数据库中获得的编码着丝粒相关蛋白CENH3、Nuf2和Ndc80的contig序列设计特异性引物, 利用PCR技术自海带中克隆得到它们的开放阅读框(ORF)序列, 分别命名为*SjCENH3*、*SjNuf2*和*SjNdc80*, 大小分别为432、1365和2172 bp, 相应地编码由143、454和723个氨基酸组成的蛋白。同源序列比对表明, *SjCENH3*由氨基端尾巴和羧基端的组蛋白折叠域组成, 后者包括4个 α 螺旋(α N、 α 1、 α 2和 α 3)和2个环(Loop1和Loop2); 但与组蛋白H3相比, *SjCENH3*的氨基端略长且序列不相似, 同时组蛋白折叠域的 α 1、 α 2螺旋及Loop1序列更多变; *SjNuf2*主要的结构域有Nuf2 domain、Spc7和SPT2; *SjNdc80*主要的结构域除了具有与*SjNuf2*相似的SPT2和Spc7结构域以外, 还有Ndc80_Hec、FoP_duplication和Kin17_mid。基于海带和其他物种的CENH3、H3、Nuf2和NCD80蛋白序列所构建的Neighbor-Joining系统进化树显示, *SjCENH3*和*SjH3*分别被显著地聚类在CENH3和H3两支中, H3在物种间是高度保守的, 而CENH3在物种间的差异相对较大; 来自不同生物的Nuf2和NCD80均可显著地分为被子植物和藻类2支, 它们可能存在着共同演化的关系。本研究首次报道了海带着丝粒蛋白*SjCENH3*、*SjNuf2*和*SjNCD80*编码序列的特征, 为制备特异抗体通过免疫荧光技术来定位海带染色体着丝粒的具体位置, 通过染色质免疫共沉淀获得海带染色体着丝粒DNA序列及进行染色体核型分析打下了良好的基础。

关键词: 海带; CENH3; 动粒蛋白; 基因克隆; 生物信息学

中图分类号: Q 785; S 968.4

文献标志码: A

海带(*Saccharina japonica*)等大型褐藻染色体由于其数目多而体积小, 多呈短棒状甚至点状, 且形态相似^[1-3], 使得人们难以通过传统的细胞学方法获得其染色体着丝粒(centromere)位置、长短臂的臂比等核型分析所需的相关数据, 因而到目前为止还没有海带等大型褐藻染色体的核型分析报道。刘宇等^[4]通过改进的染色体制备技术和染色方法, 获得了清晰的海带染色体图片, 但仍无法分清其着丝粒的具体位置。最近, Liu等^[5]和Gu等^[6]又运用荧光原位杂交(FISH)技术, 将海带雌配子体特异的分子标记定位于染色体上。这些技术和方法为通过FISH技术定位海带染色体的着丝粒及深入开展其核型

分析提供借鉴。

着丝粒是有丝分裂中将姐妹染色单体连接在一起形成染色体的主缢痕并将染色单体分为两臂(长臂和短臂)的部位^[7]。着丝粒主要由着丝粒DNA(centromere DNA, CEN)和着丝粒蛋白(centromere specific proteins, CENPs)组成。染色体的着丝粒功能与位置较保守, 但其DNA的序列并非保守^[8]。目前, 虽然不清楚海带染色体着丝粒的DNA序列, 但借助其保守的着丝粒蛋白序列, 可用免疫荧光技术对其进行定位。着丝粒蛋白由特异组蛋白CENH3(组蛋白H3的一个变体)和动粒蛋白(kinetochore protein)组成, 其中后者至少有65种, 它们通过蛋白互作形成一个庞大

收稿日期: 2015-10-10 修回日期: 2016-05-05

资助项目: 国家高技术研究发展计划(2012AA10A406); 国家自然科学基金(41376136)

通信作者: 周志刚, E-mail: zgzhou@shou.edu.cn

而复杂的结构网络^[9-12]。目前,常用的植物染色体着丝粒标记蛋白是CENH3,此外还有CENP-B、CENP-C、CENP-E、ZW10、Nuf2和Ndc80等^[13-14]。

本研究在获得编码糖海带(*S. latissima*)着丝粒特异组蛋白CENH3的contig序列以及从海带基因组数据库中筛选到的编码动粒蛋白Nuf2和Ndc80的cDNA序列基础上,通过cDNA末端快速扩增(RACE)和反转录PCR技术克隆海带CENH3、Nuf2和Ndc80等3个着丝粒蛋白编码基因的cDNA序列;通过生物信息学分析、蛋白序列比对、系统进化树构建等来分析这3个基因的序列特征。为后续的着丝粒蛋白表达载体构建、重组蛋白分离纯化与特异抗体制备,并用于海带染色体着丝粒的免疫荧光原位检测提供基础。

1 材料与方法

1.1 海带配子体无性繁殖体系培养

本实验中所用的海带雌、雄配子体无性繁

殖系为“荣福”品系,按已报道的方法分离和保藏^[15]。在温度(17±1) °C、光周期16L:8D(光/暗)等条件下,于PES培养基中培养^[16],每2个星期换1次培养基。

1.2 海带着丝粒相关蛋白编码基因cDNA全长序列的克隆

采用TRIzol试剂(Invitrogen公司)并按该产品说明提取海带总RNA。使用PrimeScript™ RT反转录试剂盒(TaKaRa公司),按其操作步骤合成cDNA第一链,作为后续PCR反应的模板。

在本研究中,编码海带H3的contig序列由刘福利博士馈赠;编码糖海带CENH3的contig序列来自糖海带转录组库^[17];编码Nuf2(gene ID: SJ19509)和Ndc80(gene ID: SJ20066)的contig序列来自于海带基因组文库^[18]。根据这些序列并利用Primer Premier 5软件设计引物(表1),以分别扩增这4个基因的cDNA片段。PCR反应体系为25 mL,包含9.5 mL灭菌双蒸水、12.5 mL的2×Pfu PCR

表1 引物及其序列

Tab. 1 The sequence of primers

| 引物 primer | 序列(5'-3') sequence (5'-3') | 退火温度/°C annealing temperature | 产物大小/bp size of amplified product |
|--------------|-------------------------------|----------------------------------|--------------------------------------|
| CENH3-U | ATGGCGCGCCGGAAGCAT | 66.5 | 425 |
| CENH3-D | CGACTCTGTCCTCGAATC | | |
| H3-U | ATGGCTCGACCAAGCAGAC | 68.2 | 405 |
| H3-D | GTTGCGCTCACCGCGGAG | | |
| Nuf2-U1 | AGGCCCGTTCTATCTACGTACC | 66.5 | 541 |
| Nuf2-D1 | AGGCAATGTGTGCTGTAATGG | | |
| Nuf2-U2 | CCTCTCTGCCCCACAAAGTCT | 68 | 737 |
| Nuf2-D2 | TGACACGCTCACAAACCCAG | | |
| Nuf2-U3 | GTCCTGTGACCCAGAGCA | 68 | 1362 |
| Nuf2-D3 | ACCTTGCTCAGAGCGTGTCC | | |
| Ndc80-U1 | GAAAGTATCCCCTCGTCTGTGGACATCA | 66 | 883 |
| Ndc80-D1 | GAGCAGGTCGACTGGTTTGTCTTCACG | | |
| Ndc80-U2 | AGTGCGTCCCTCAACTCGC | 65 | 552 |
| Ndc80-D2 | CATAGTTGGGTCGATCTG | | |
| Ndc80-U3 | ATGGCGCAGGAAGAGGTGTCTA | 68 | 784 |
| Ndc80-D3 | CGACCCTGTGTCCGGCAT | | |
| Ndc80-U4 | ATGGCGGCTCGCAGGC | 66.7 | 2158 |
| Ndc80-D4 | CAACTGAAGGTAAGTTCAGCGT | | |

Master Mix(天根生化科技有限公司)、1 mL模板cDNA、上下游引物(10 mmol/L)各1.0 mL。反应程序: 95 °C预变性5 min; 94 °C变性45 s, 在各自退火温度(表1)退火1 min, 72 °C延伸2 min, 35个循环; 最后72 °C延伸10 min。

利用1%琼脂糖凝胶电泳检测PCR扩增反应的产物, 并按照琼脂糖凝胶纯化试剂盒(北京艾德莱生物科技有限公司)的说明回收PCR产物。将回收的产物与pMD-19T载体(TaKaRa公司)于16 °C连接过夜, 再转化到大肠杆菌(*Escherichia coli*) DH5a感受态细胞(天根生化科技公司)中, 涂平板, 将蓝白斑筛选后获得的阳性克隆送上海生工生物工程技术有限公司进行测序。

1.3 海带着丝粒相关蛋白编码基因序列的生物信息学及聚类分析

利用NCBI中的BLAST(<http://www.ncbi.nlm.nih.gov/BLAST/>)对获得的海带*SjCENH3*、*SjH3*、*SjNuf2*和*SjNdc80*基因的ORF序列进行核酸、蛋白质、氨基酸序列同源性搜索, 并利用Expasy中工具ProtParam(<http://www.expasy.org/tool/protparam.html>)在线分析蛋白质氨基酸组成、分子量、等电点。使用COILS软件(http://www.ch.embnet.org/software/COILS_form.html)进行内环结构预测^[19], 使用SMART软件(http://smart.embl-heidelberg.de/smart/set_mode.cgi?NORMAL=1)进行功能结构域预测。将所有序列进行比对, 并手工删除缺口(gap), 利用MEGA6.0软件^[20]构建Neighbor-Joining(NJ)系统进化树。

2 结果与分析

2.1 海带3个着丝粒蛋白和组蛋白H3编码基因的ORF序列

由于CENH3是H3的一个变体, 为了展现克隆得到的海带CENH3编码基因序列不同于H3, 在克隆*CENH3*、*Nuf2*和*Ndc80*的同时, 本研究还克隆了H3的编码基因。根据所获得的*CENH3*、*H3*、*Nuf2*和*Ndc80*的contig序列设计正、反向引物(表1), 经PCR扩增(图1)、克隆、测序, 再通过ORF Finder在线预测, 发现这些基因的开放阅读框架(ORF)序列是完整的, 其长度分别为432、411、1365和2172 bp。经过BLASTp搜索发现, 自海带中所克隆到的*CENH3*、*H3*、*Nuf2*以及*Ndc80*所编码的蛋白序列与糖海带CENH3^[17]、人

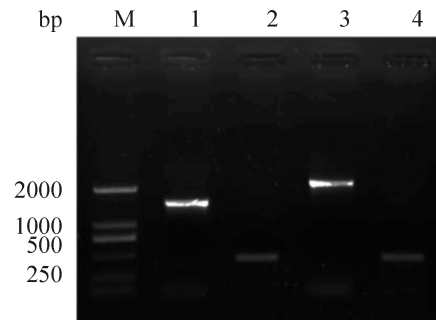


图1 海带着丝粒相关蛋白编码基因的全长ORF序列克隆电泳图

M. 2000 bp DNA分子量标准; 泳道1~4分别为从cDNA中扩增出的*SjNuf2*、*SjCENH3*、*SjNdc80*和*SjH3*的开放阅读框(ORF)

Fig. 1 Electrophoresis patterns of full-length cDNA sequence of centromere-associated proteins genes cloned from *S. japonica*

M. 2000 bp DNA marker; lanes 1 to 4. electrophoresis patterns of amplified product of *SjNuf2*, *SjCENH3*, *SjNdc80* and *SjH3*

类(*Homo sapiens*)H3(GenBank登录号: NP066403)、长囊水云(*Ectocarpus siliculosus*)Nuf2(GenBank登录号: CBJ32075)以及Ndc80 (GenBank登录号: CBN77150)的同源性分别为94%、96%、58%和62%, 因而将这4个所克隆到的基因分别命名为*SjCENH3*、*SjH3*、*SjNuf2*和*SjNdc80*。

2.2 着丝粒蛋白SjCENH3的基因特征分析

预测海带*SjCENH3*、*SjH3*基因编码的蛋白分别由143和136个氨基酸组成, 分子量分别为16.2和15.3 ku, 等电点分别为11.6和11.3。BLASTp搜索发现*SjCENH3*与糖海带^[17]和长囊水云(GenBank登录号: CBJ2646)CENH3的相似度分别为94%和78%, 它们的主要差别在于氨基端的长度和序列(图2-a)。充分表明CENH3在物种间是高度变异的, 即便在不同的亚种间也会表现出一定的适应性进化^[14]。与CENH3的高度变化形成鲜明的对照, 组蛋白H3则是高度保守的, 甚至自海带中所克隆的*SjH3*与人类的*H3*, 二者的氨基酸序列相似性也高达96%(图2-a)。在目前所知的真核生物中, H3均由136个氨基酸组成, 分子量均在15 ku左右。但通过比对发现, 自海带中所克隆的*SjCENH3*与*SjH3*, 它们二者所编码的氨基酸序列的相似性仅为58%(图2-a), 表明应是不同的基因。

通过同源序列比对和结构域分析可知*SjCENH3*和*SjH3*的相同点: 二者均由氨基端尾巴(N-tail)和羧基端的组蛋白折叠域(histone-fold domain, HFD)

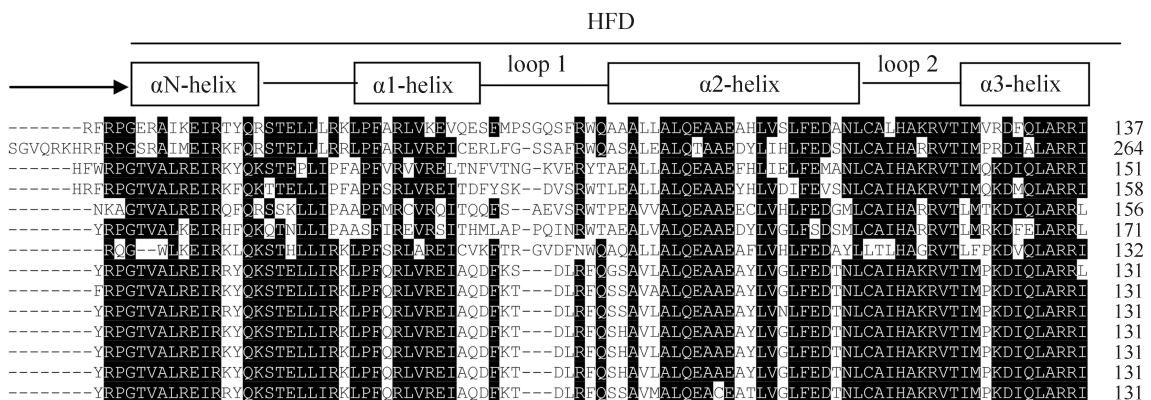
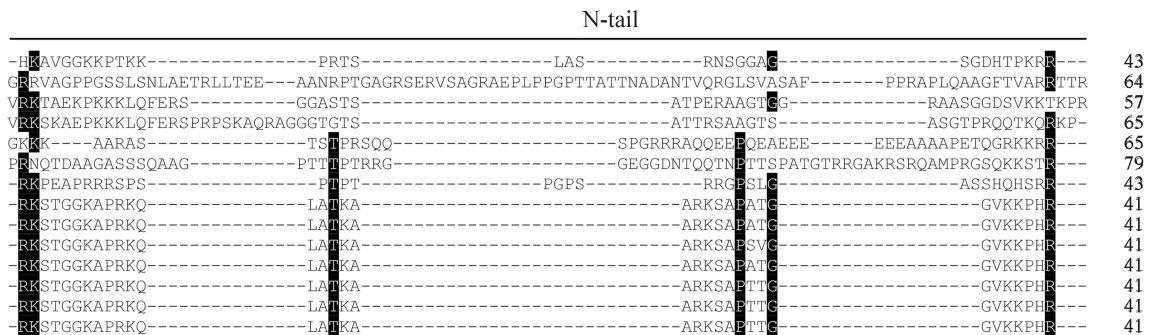
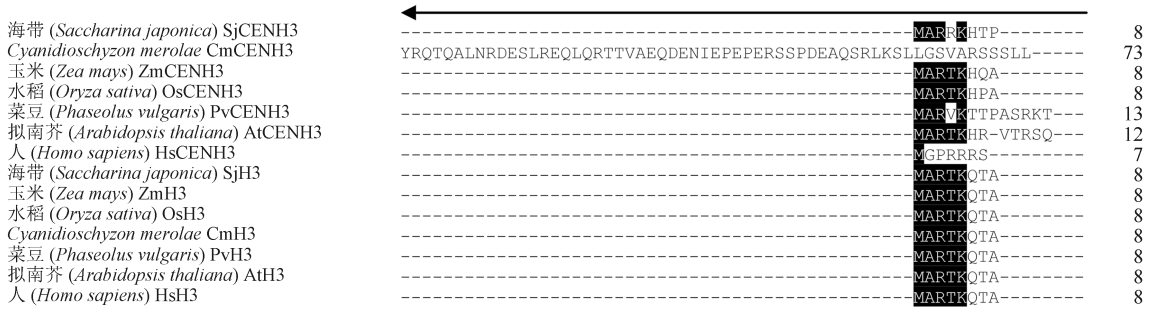
组成;羧基端的组蛋白折叠域均包括4个 α 螺旋(α N、 α 1、 α 2和 α 3)和2个环(Loop1和Loop2);还具有相同的功能结构域CENPs(位于SjCENH3的71Phe~139Gly区域)(图2-a)。CENPs是一种具有双重功能的蛋白结构域,可参与DNA的损伤修复并维持基因组的完整,同时可促进有丝分裂过程中染色体的分离^[21-22]。SjCENH3和SjH3的不同点:SjCENH3的氨基端略长于SjH3,且二者之间没有相似性;SjCENH3组蛋白折叠域的 α 1、 α 2螺旋及Loop1序列比SjH3显得更多变(图2-a),从而导致了它们在羧基端仅有65%的同源性。这与之前有关CENH3与H3的报道相符^[14, 23-25]。

上述分析结果表明SjCENH3既具备其他物种CENH3的主要特征,也与SjH3具有共同的特征。因而推测SjCENH3是来自SjH3家族但属于不同于SjH3的序列。

2.3 着丝粒蛋白SjNuf2的基因特征分析

预测海带SjNuf2基因编码的蛋白由454个氨基酸组成,分子量为50.7ku,等电点为7.6。BLASTp搜索结果显示,SjNuf2及其他物种Nuf2之间的相似性都较低,与同属于褐藻的长囊水云的Nuf2也仅有58%的相似性。

经同源序列比对和结构域预测,发现SjNuf2的氨基端序列相对保守,含有Nuf2结构域(位于SjNuf2的9Thr-135Glu区域)(图2-b);而羧基端的序列则变化很大,含有的主要功能结构域有Spc7(位于SjNuf2的103Ile~411Asp区域)和SPT2(位于SjNuf2的147Lys~239Glu区域)。其中,Spc7是动粒蛋白结构域,存在于动粒—纺锤体相关的细胞分裂蛋白中^[26],是NMS(Ndc80-MIND-Spc7)超级复合体的组成成分(图2-c),在细胞的减数分裂



(a)

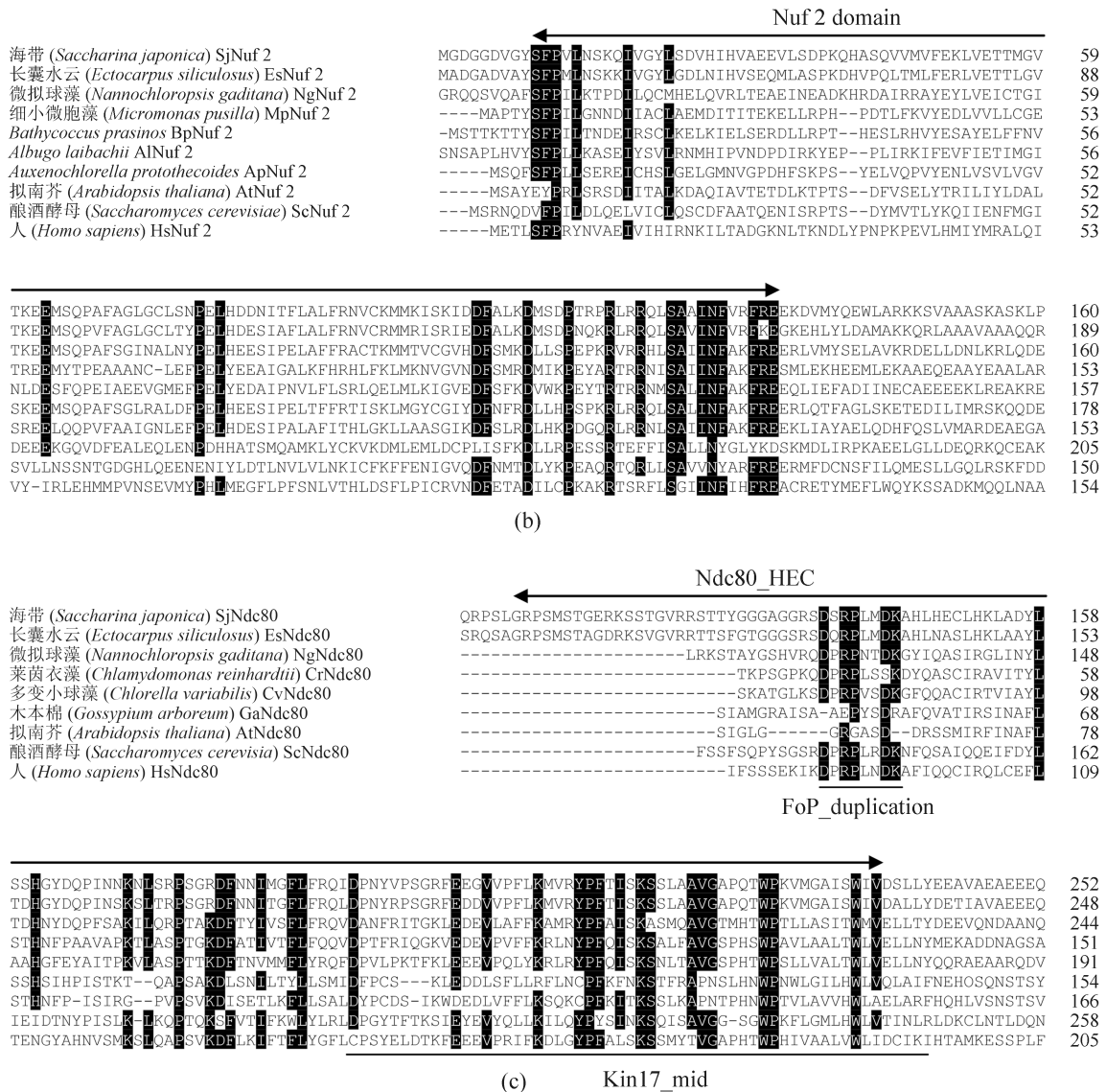


图2 来自不同物种着丝粒蛋白CENH3(a)、Nuf2(b)和Ndc80(c)的序列比对

Fig. 2 Alignment of deduced amino acid sequences of CENH3 (a), Nuf2 (b) and Ndc80 (c) genes from various organisms

后期和有丝分裂整个过程中都起着重要作用^[27]；SPT2是染色质蛋白结构域之一，参与了组蛋白H3转录水平的调节^[28]。

虽然SjNuf2与其他物种Nuf2氨基酸序列的相似性低，但因其均具有Nuf2、Spc7和SPT2等特异的功能结构域，说明其结构和功能比较保守^[9]。本研究所克隆到的*SjNuf2*应属于*Nuf2*家族中的一员。

2.4 着丝粒蛋白SjNdc80的基因特征分析

预测海带*SjNdc80*基因编码的蛋白由723个氨基酸组成，分子量为80.3 ku，等电点为5.4。BLASTp搜索结果显示，SjNdc80与长囊水云Ndc80的氨基

酸序列相似性为62%。

经同源序列比对和功能结构域预测，发现SjNdc80氨基端的序列相对保守，主要结构域有Ndc80_HEC(位于SjNdc80的101Gly~254Arg区域)、FoP_duplication(位于SjNdc80的131Asp~138Lys区域)和Kin17_mid(位于SjNdc80的186Asp~242Tyr区域)(图2-c)；而羧基端则不保守，主要结构域有SPT2(位于SjNdc80的247Ala~327Leu区域)和Spc7(位于SjNdc80的182Phe~500Asp区域)。在这些结构域中，Fop_duplication是染色质相关蛋白所特有的结构域，与兼性异染色质共定位在一起^[29]；Kin17_mid是一种氨基端具有锌指结构的可与DNA结合的蛋白结构域^[30]。Ndc80是Ndc80

蛋白复合体的一部分^[31],与染色体的结构保持蛋白(structural maintenance of chromosomes)相互作用,通过与纺锤体微管结合从而参与染色体的分离^[32-33]。此外,通过COILS软件预测出SjNdc80具有一个内环结构(位于SjNdc80的480Leu~534Thr区域),这与之前报道的其他物种Ndc80的特征相符^[34-37],可将Ska、Dis1和Dam1等动粒蛋白募集到动粒上,是动粒微管结合的最佳位点。

从上述的分析可知,SjNdc80与其他物种Ndc80的氨基酸相似度虽然很低,但它们都具有Ndc80_HEC、FoP_duplication、Kin17_mid、SPT2和Spc7等特异的功能结构域,还具有Ndc80家族特有的内环结构,因而其结构和功能却是相对保守的^[9],表明本研究自海带中所克隆到的SjNdc80应是Ndc80这个家族的成员。

2.5 不同物种3个着丝粒蛋白氨基酸序列的聚类分析

通过同源搜索,从NCBI中获得其他物种的9个CENH3基因和9个H3基因、15个Nuf2基因和16个Ndc80基因,结合本研究自海带中所克隆到的SjCENH3、SjH3、SjNuf2和SjNdc80,基于它们的蛋白序列,利用MEGA6.0软件构建NJ系统进化树。

CENH3及H3的聚类结果(图3-a)显示,来自不同物种的H3包括SjH3被显著地(bootstrap为100%)聚为一支,从而区别于CENH3所组成的分支,也暗示SjCENH3和SjH3之间的不同;H3在物种间特别在高等植物之间是高度保守的,以致难以在所构建的NJ系统进化树中来分辨它们的差异;与H3的高度保守形成鲜明的对照,CENH3在物种间的差异相对较大;海带的SjCENH3与长囊水云的CENH3亲缘关系最近(bootstrap为100%),并与Cyanidioschyzon merolae这种单细胞红藻的CENH3(GenBank登录号:XP005535607)组成一个亚支(bootstrap为93%),从而区别于由高等植物CENH3所组成的亚支。

来自不同生物的Nuf2可分为2支,即被子植物(bootstrap为100%)和藻类(图3-b)。海带SjNuf2与长囊水云Nuf2亲缘关系最近(bootstrap为100%),并与微拟球藻(Nannochloropsis gaditana)Nuf2(GenBank登录号:EWM29090)组成显著(bootstrap为95%)不同于绿藻和红藻这些一次内吞

作用所衍生物种的亚支(bootstrap为85%),暗示绿藻和红藻的Nuf2可能来自共同的祖先,而海带、长囊水云与微拟球藻等这些二次内吞作用所衍生物种的Nuf2则可能来自另一共同的祖先。当然,这还需要更多物种的Nuf2序列证据来进一步证实。

与Nuf2的聚类结果类似,来自不同生物的Ndc80也可分为被子植物(bootstrap为100%)和藻类这2支(图3-c)。海带SjNdc80与长囊水云Ndc80及微拟球藻Ndc80(GenBank登录号:EWM24208)等这些二次内吞作用所衍生物种显著(bootstrap为100%)不同于绿藻类Ndc80所组成的亚支(bootstrap为91%)。鉴于Nuf2与Ndc80都是着丝粒相关蛋白以及图3-b和3-c所显示的类似演化趋势,推测来自不同生物的Nuf2与Ndc80可能存在共同演化(co-evolution)的关系,从而暗示它们是细胞所必需的蛋白并在细胞分裂过程中可能起着相同的作用。

3 讨论

CENH3是位于染色体着丝粒内层替代核小体组蛋白H3的一种基本蛋白^[38-39],第一个CENH3(即CENP-A)是从人类的活性着丝粒中鉴定出来的^[40],在已研究的所有真核生物中,它只存在于染色体的着丝粒处^[41-42]。有关CENH3的研究多集中在高等植物的禾本科(Poaceae)、十字花科(Cruciferae)及常见的几种模式生物中,而对低等植物如藻类CENH3的研究,目前仅在C. merolae这种单细胞红藻中有报道^[23, 43]。同源序列比对(图2)及聚类分析(图3)的结果表明,SjCENH3与SjH3都具有相同的组蛋白折叠域和功能结构域CENPs,因而起源相同;但其氨基端序列的差异导致SjCENH3明显不同于SjH3(图3-a)。但这样的差异为制备SjCENH3特异性抗体以荧光免疫技术确定海带染色体着丝粒位置提供了便利。这也是目前利用CENH3进行物种着丝粒免疫定位所经常采用的策略^[23, 25, 44]。

Nuf2和Ndc80都是Ndc80蛋白复合体(包括Ndc80、Nuf2、Spc24和Spc25等4种蛋白)的重要组成成分^[45]。Ndc80复合体首次发现于酿酒酵母(Saccharomyces cerevisiae)中^[46],在细胞的有丝分裂中可与着丝粒相互作用,参与酵母染色体的分离和移动^[47-48]。目前有关Nuf2和Ndc80的研究

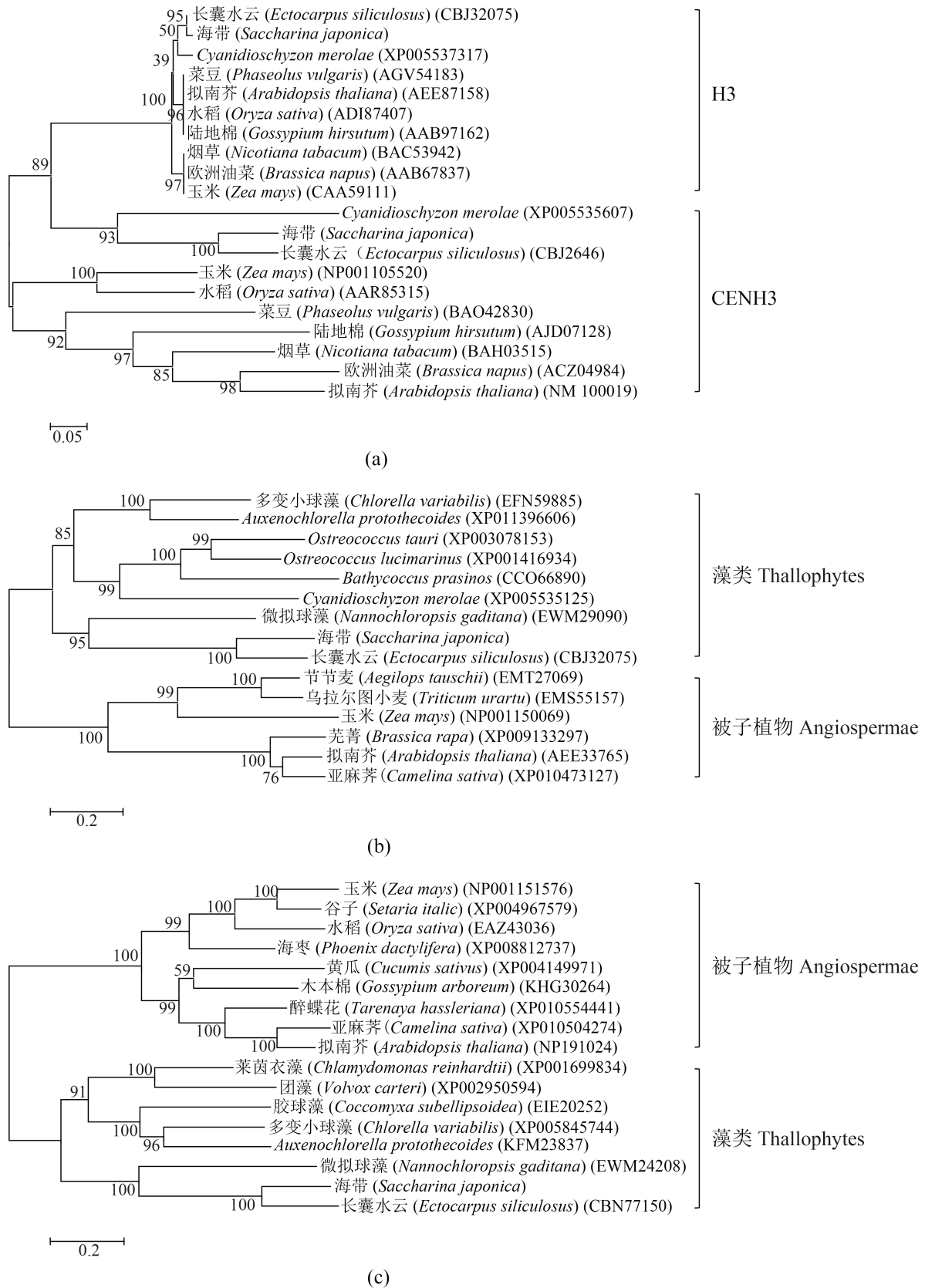


图 3 基于不同物种CENH3(a)、Nuf2(b)和Ndc80(c)的蛋白序列利用邻接法构建的聚类图
 Fig. 3 Neighbour-Joining phylogenetic tree inferred from the deduced amino acid sequences of CENH3 (a), Nuf2 (b) and Ndc80 (c) genes from various species of organisms

主要集中在酵母^[46-48]、人类^[49-50]和小鼠^[51]中, 在植物中相关的研究很少, 仅在拟南芥(*Arabidopsis thaliana*)这种模式生物中有报道^[8]。SjNuf2和SjNdc80与各自的同系物虽然序列差异比较大, 但具有各自相同的功能结构域(图2-b, 2-c); 系统发育分析进一步表明SjNuf2和SjNdc80分别属于Nuf2家族和Ndc80家族(图3-b, 3-c)。鉴于这2个蛋白序列在不同物种之间存在着较大的差异, 在利用免疫荧光技术进行海带染色体着丝粒定位时, 就必须制备该物种的SjNuf2和SjNdc80特异性抗体。

本研究首次报道并分析了海带3个着丝粒蛋白SjCENH3、SjNuf2和SjNdc80编码序列的特征, 为这些蛋白的重组表达、特异性抗体制备, 以利用免疫荧光技术定位海带染色体着丝粒位置; 或通过染色质免疫共沉淀技术以获取海带着丝粒DNA序列, 并利用FISH技术将其定位于染色体中打下基础。

参考文献:

- [1] Evans L V. A large chromosome in the laminarian nucleus[J]. *Nature*, 1963, 198(4876): 215.
- [2] Yabu H, Yasui H. Chromosome number in four species of *Laminaria* (Phaeophyta)[J]. *Japanese Journal of Phycology*, 1991, 39(2): 185-187.
- [3] Lewis R J. Chromosomes of the brown algae[J]. *Phycologia*, 1996, 35(1): 19-40.
- [4] 刘宇, 毕燕会, 周志刚. 海带染色体的DAPI染色及核型初步分析[J]. *水产学报*, 2012, 36(1): 50-54.
- Liu Y, Bi Y H, Zhou Z G. Karyological observation on *Saccharina japonica* chromosomes stained with DAPI[J]. *Journal of Fisheries of China*, 2012, 36(1): 50-54 (in Chinese).
- [5] Liu Y, Bi Y H, Gu J G, *et al.* Localization of a female-specific marker on the chromosomes of the brown seaweed *Saccharina japonica* using fluorescence *in situ* hybridization[J]. *PLoS One*, 2012, 7(11): e48784.
- [6] Gu J G, Sun Y P, Liu Y, *et al.* Sex identification and genetic variation of *Saccharina* (Phaeophyta) gametophytes as revealed by inter-simple sequence repeat (ISSR) markers[J]. *Journal of Applied Phycology*, 2014, 26(1): 635-646.
- [7] 翟中和, 王喜忠, 丁明孝. 细胞生物学[M]. 北京: 高等教育出版社, 2007.
- Zhai Z H, Wang X Z, Ding M X. *Cell Biology*[M]. Beijing: Higher Education Press, 2007 (in Chinese).
- [8] Ramahi J S. Quantitative analysis of CENH3 and NDC80 complexes in *Arabidopsis* kinetochores during mitosis and meiosis[D]. California: University of California, 2012.
- [9] Cheeseman I M, Desai A. Molecular architecture of the kinetochore-microtubule interface[J]. *Nature Reviews Molecular Cell Biology*, 2008, 9(1): 33-46.
- [10] Maddox P S, Oegema K, Desai A. "Holo"er than thou: Chromosome segregation and kinetochore function in *C. elegans*[J]. *Chromosome Research*, 2004, 12(6): 641-653.
- [11] Westermann S, Drubin D G, Barnes G. Structures and functions of yeast kinetochore complexes[J]. *Annual Review of Biochemistry*, 2007, 76: 563-591.
- [12] Santaguida S, Musacchio A. The life and miracles of kinetochores[J]. *The EMBO Journal*, 2009, 28(17): 2511-2531.
- [13] Copenhaver G P, Zieler H, Preuss D. Identification of centromere sequences using centromere associated proteins and uses thereof: United States, 20120115132[P]. 2012-05-10.
- [14] Nagaki K, Yamamoto M, Yamaji N, *et al.* Chromosome dynamics visualized with an anti-centromeric histone H3 antibody in *Allium*[J]. *PLoS One*, 2012, 7(12): e51315.
- [15] 周志刚, 吴超元. 海带无性繁殖系的形成及孢子体诱导[J]. *生物工程学报*, 1998, 14(1): 109-111.
- Zhou Z G, Wu C Y. Clone culture of *Laminaria japonica* and induction of its sporophytes[J]. *Chinese Journal of Biotechnology*, 1998, 14(1): 109-111 (in Chinese).
- [16] Starr R C, Zeikus J A. UTEX-The culture collection of algae at the University of Texas at Austin 1993 list of cultures[J]. *Journal of Phycology*, 1993, 29(S2): 1-106.
- [17] Heinrich S, Valentin K, Frickenhaus S, *et al.* Transcriptomic analysis of acclimation to temperature and light stress in *Saccharina latissima* (Phaeophyceae)[J]. *PLoS One*, 2012, 7(8): e44342.
- [18] Ye N H, Zhang X W, Miao M, *et al.* *Saccharina* genomes provide novel insight into kelp biology[J].

- Nature Communications, 2015, 6: 6986.
- [19] Lupas A, Van Dyke M, Stock J. Predicting coiled coils from protein sequences[J]. Science, 1991, 252(5009): 1162-1164.
- [20] Tamura K, Peterson D, Peterson N, *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods[J]. Molecular Biology and Evolution, 2011, 28(10): 2731-2739.
- [21] Osman F, Whitby M C. Emerging roles for centromere-associated proteins in DNA repair and genetic recombination[J]. Biochemical Society Transactions, 2013, 41(6): 1726-1730.
- [22] Bhattacharjee S, Osman F, Feeney L, *et al.* MHF1-2/CENP-S-X performs distinct roles in centromere metabolism and genetic recombination[J]. Open Biology, 2013, 3(9): 130102.
- [23] Maruyama S, Kuroiwa H, Miyagishima S Y. *et al.* Centromere dynamics in the primitive red alga *Cyanidioschyzon merolae*[J]. The Plant Journal, 2007, 49(6): 1122-1129.
- [24] Ravi M, Shibata F, Ramahi J S, *et al.* Meiosis-specific loading of the centromere-specific histone CENH3 in *Arabidopsis thaliana*[J]. PLoS Genetics, 2011, 7(6): e1002121.
- [25] Iwata A, Tek A L, Richard M M S, *et al.* Identification and characterization of functional centromeres of the common bean[J]. The Plant Journal, 2013, 76(1): 47-60.
- [26] Kerres A, Vietmeier-Decker C, Ortiz J, *et al.* The fission yeast kinetochore component Spc7 associates with the EB1 family member Mal3 and is required for kinetochore-spindle association[J]. Molecular Biology of the Cell, 2004, 15(12): 5255-5267.
- [27] Jakopec V, Topolski B, Fleig U. Sos7, an essential component of the conserved *Schizosaccharomyces pombe* Ndc80-MIND-Spc7 complex, identifies a new family of fungal kinetochore proteins[J]. Molecular and Cellular Biology, 2012, 32(16): 3308-3320.
- [28] Nourani A, Robert F, Winston F. Evidence that Spt2/Sin1, an HMG-like factor, plays roles in transcription elongation, chromatin structure, and genome stability in *Saccharomyces cerevisiae*[J]. Molecular and Cellular Biology, 2006, 26(4): 1496-1509.
- [29] van Dijk T B, Gillemans N, Stein C, *et al.* Friend of Prmt1, a novel chromatin target of protein arginine methyltransferases[J]. Molecular and Cellular Biology, 2010, 30(1): 260-272.
- [30] Kannouche P, Mauffrey P, Pinon-Lataillade G, *et al.* Molecular cloning and characterization of the human KIN17 cDNA encoding a component of the UVC response that is conserved among metazoans[J]. Carcinogenesis, 2000, 21(9): 1701-1710.
- [31] Wigge P A, Kilmartin J V. The Ndc80p complex from *Saccharomyces cerevisiae* contains conserved centromere components and has a function in chromosome segregation[J]. The Journal of Cell Biology, 2001, 152(2): 349-360.
- [32] Hirano M, Hirano T. Hinge-mediated dimerization of SMC protein is essential for its dynamic interaction with DNA[J]. The EMBO Journal, 2002, 21(21): 5733-5744.
- [33] Jensen R B, Shapiro L. Cell-cycle-regulated expression and subcellular localization of the *Caulobacter crescentus* SMC chromosome structural protein[J]. Journal of Bacteriology, 2003, 185(10): 3068-3075.
- [34] Nilsson J. Looping in on Ndc80-How does a protein loop at the kinetochore control chromosome segregation?[J]. BioEssays, 2012, 34(12): 1070-1077.
- [35] Zhang G, Kelstrup C D, Hu X W, *et al.* The Ndc80 internal loop is required for recruitment of the Ska complex to establish end-on microtubule attachment to kinetochores[J]. Journal of Cell Science, 2012, 125(13): 3243-3253.
- [36] Maure J F, Komoto S, Oku Y, *et al.* The Ndc80 loop region facilitates formation of kinetochore attachment to the dynamic microtubule plus end[J]. Current Biology, 2011, 21(3): 207-213.
- [37] Varma D, Chandrasekaran S, Sundin L J R, *et al.* Recruitment of the human Cdt1 replication licensing protein by the loop domain of Hec1 is required for stable kinetochore-microtubule attachment[J]. Nature Cell Biology, 2012, 14(6): 593-603.
- [38] Panchenko T, Sorensen T C, Woodcock C L, *et al.* Replacement of histone H3 with CENP-a directs global

- nucleosome array condensation and loosening of nucleosome superhelical termini[J]. Proceedings of the National Academy of Sciences of the United States of America, 2011, 108(40): 16588-16593.
- [39] Tachiwana H, Kagawa W, Kurumizaka H. Comparison between the CENP-a and histone H3 structures in nucleosomes[J]. Nucleus, 2012, 3(1): 6-11.
- [40] Palmer D K, O'day K, Wener M H, *et al.* A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones[J]. The Journal of Cell Biology, 1987, 104(4): 805-815.
- [41] Nagaki K, Terada K, Wakimoto M, *et al.* Centromere targeting of alien CENH3s in *Arabidopsis* and tobacco cells[J]. Chromosome Research, 2010, 18(2): 203-211.
- [42] Sullivan B A, Blower M D, Karpen G H. Determining centromere identity: Cyclical stories and forking paths[J]. Nature Reviews Genetics, 2001, 2(8): 584-596.
- [43] Kanasaki Y, Imamura S, Matsuzaki M, *et al.* Identification of centromere regions in chromosomes of a unicellular red alga, *Cyanidioschyzon merolae*[J]. FEBS Letters, 2015, 589(11): 1219-1224.
- [44] Nagaki K, Kashihara K, Murata M. A centromeric DNA sequence colocalized with a centromere-specific histone H3 in tobacco[J]. Chromosoma, 2009, 118(2): 249-257.
- [45] Wei R R, Sorger P K, Harrison S C. Molecular organization of the Ndc80 complex, an essential kinetochore component[J]. Proceedings of the National Academy of Sciences of the United States of America, 2005, 102(15): 5363-5367.
- [46] Wigge P A, Jensen O N, Holmes S, *et al.* Analysis of the *Saccharomyces* spindle pole by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry[J]. The Journal of Cell Biology, 1998, 141(4): 967-977.
- [47] Janke C, Ortiz J, Lechner J, *et al.* The budding yeast proteins Spc24p and Spc25p interact with Ndc80p and Nuf2p at the kinetochore and are important for kinetochore clustering and checkpoint control[J]. The EMBO Journal, 2001, 20(4): 777-791.
- [48] He X W, Rines D R, Espelin C W, *et al.* Molecular analysis of kinetochore-microtubule attachment in budding yeast[J]. Cell, 2001, 106(2): 195-206.
- [49] Meng Q C, Wang H C, Song Z L, *et al.* Overexpression of NDC80 is correlated with prognosis of pancreatic cancer and regulates cell proliferation[J]. American Journal of Cancer Research, 2015, 5(5): 1730-1740.
- [50] Sugimasa H, Taniue K, Kurimoto A, *et al.* Heterogeneous nuclear ribonucleoprotein K upregulates the kinetochore complex component NUF2 and promotes the tumorigenicity of colon cancer cells[J]. Biochemical and Biophysical Research Communications, 2015, 459(1): 29-35.
- [51] Zhang T, Zhou Y, Qi S T, *et al.* Nuf2 is required for chromosome segregation during mouse oocyte meiotic maturation[J]. Cell Cycle, 2015, 14(16): 2701-2710.

Molecular cloning and characterization of 3 centromere specific proteins (CENPs) genes from *Saccharina japonica*

XU Mei, ZHOU Zhigang*

(College of Fisheries and Life Science, Shanghai Ocean University, Shanghai 201306, China)

Abstract: Based upon contigs coding for centromere specific proteins (CENPs) from a pyrosequencing transcriptome of *Saccharina japonica*, gene-specific primers were designed for the three centromere specific protein gene cloning from *S. japonica*. Three CENPs genes, named as *SjCENH3*, *SjNuf2* and *SjNCD80* were cloned by using the technique of Polymerase Chain Reaction. The open reading frame of the cloned full-length cDNAs were 432, 1365 and 2172 bp in length, encoding putative proteins composed of 143, 454 and 723 amino acids, respectively. Homologous sequence alignments show: *SjCENH3* include the diversity of the N-terminal region and C-terminal histone fold domain (HFD), and HFD including four α helixes (α N helix, α 1 helix, α 2 helix and α 3 helix) and two loops (Loop1 and Loop2). But compared to H3, *SjCENH3* amino terminus is slightly longer and there is no similarity between them, while the sequence of histone fold domain α 1-helix, α 2-helix and Loop1 are more variable. The main domains of *SjNuf2* are Nuf2 domain, Spc7 and SPT2. The main domains of *SjNdc80* are SPT2, Spc7, Ndc80_Hec, FoP_duplication and Kin17_mid. Neighbor-Joining (NJ) phylogenetic tree inferred from the putative proteins of *CENH3*, *Nuf2* and *NCD80* genes of *S. japonica* and other species indicated that *SjCENH3* and *SjH3* genes were divided two clades: CENH3 and H3 clade. H3 is highly conserved across species, while CENH3 differences among species are relatively large. *Nuf2* and *Ndc80* from different organisms could be significantly divided into two clades, angiosperms and algae. They both may have relations of co-evolution. This study first reported the characteristics of three centromere protein-coding genes *SjCENH3*, *SjNuf2* and *SjNdc80*, providing a good basis to locate the centromere position by immunofluorescence technique and to get the centromeric DNA sequence by chromatin immunoprecipitation and to analyze chromosome karyotype using preparation of antibodies specific for the corresponding proteins of *S. japonica*.

Key words: *Saccharina japonica*; CENH3; kinetochore protein; gene cloning; bioinformatics

Corresponding author: ZHOU Zhigang. Email: zgzhou@shou.edu.cn

Funding projects: National High-Tech Research and Development Projects (2012AA10A406); National Natural Science Foundation of China (41376136)