

## شناسایی و تحلیل فیلوژنتیکی و بیانی چپرون‌های هیستونی کلاس NAP در ذرت (*Zea mays*)

امین عابدی<sup>۱</sup>، رضا شیرزادیان خرم‌آباد<sup>۲\*</sup>، محمد مهدی سوهانی<sup>۳</sup>

۱. دانش‌آموخته دکترای بیوتکنولوژی کشاورزی-گیاهی، دانشکده علوم کشاورزی دانشگاه گیلان، رشت، ایران

۲. استادیار گروه بیوتکنولوژی کشاورزی دانشکده علوم کشاورزی دانشگاه گیلان، رشت، ایران

۳. دانشیار گروه بیوتکنولوژی کشاورزی دانشکده علوم کشاورزی دانشگاه گیلان، رشت، ایران

(تاریخ دریافت: ۱۳۹۶/۳/۷ - تاریخ تصویب: ۱۳۹۶/۶/۲۶)

## Identification, phylogeny and expression analysis of NAP-family histone chaperones in maize (*Zea mays*)

Amin Abedi<sup>1</sup>, Reza Shirzadian-Khoramabad<sup>2\*</sup>, Mohammad-Mehdi Sohani<sup>3</sup>

1. Graduate Ph.D. Department of Agricultural Biotechnology, Faculty of Agricultural Sciences, University of Guilan, Rasht, Iran

2. Assistant Professor, Department of Agricultural Biotechnology, Faculty of Agricultural Sciences, University of Guilan, Rasht, Iran

3. Associate Professor, Department of Agricultural Biotechnology, Faculty of Agricultural Sciences, University of Guilan, Rasht, Iran

(Received: May 28, 2017 - Accepted: )

### Abstract

In eukaryotes cells, genomic DNA in combination with histone proteins is formed the chromatin. Histone chaperones affect the gene transcription via altering in DNA accessibility. In contrast to their animal and yeast counterparts, not much is known about plant histone chaperones. Nucleosome assembly protein (NAP) family histone chaperones are conserved throughout eukaryotic genomics. NAP is an integral component in the establishment, maintenance, and dynamics of eukaryotic chromatin. They transfer histones into the nucleus, assemble nucleosomes, and promote chromatin fluidity, thereby, affecting the transcription of many genes. In this study, by applying some bioinformatics analysis approaches, six putative NAP genes (*ZmNAPL1-ZmNAPL6*) were identified in maize (*Zea mays*) using the released maize genomic sequences. Phylogenetic analysis showed that these *ZmNAPLs* are classified into two subgroups as found in *Arabidopsis* and rice. Moreover, it was found that maize NAPL proteins are more closely related to rice. The *ZmNAPL* genes contained three to eleven introns and were distributed across 5 out of 20 chromosomes in maize. Microarray-based expression analysis of *ZmNAPLs* showed that there is a tight transcriptional regulation on *ZmNAPL* genes during the plant development in maize suggesting that they may play a role in genetic reprogramming in association with the developmental process. This study is the first report about NAPL gene family in maize and obtained results provide basic information for future research on the functions of NAPL genes in maize.

**Keywords:** Bioinformatics, Gene Expression, Nucleosome, Phylogenetic Analysis.

### چکیده

در یوکاریوت‌ها DNA ژنومی در ترکیب با پروتئین‌های هیستونی کروماتین را ایجاد می‌کند. چپرون‌های هیستونی از طریق تغییر دسترسی به DNA بر میزان رونویسی ژن‌ها تأثیر می‌گذارند. بر خلاف مخمر و جانوران، در مورد چپرون‌های هیستونی گیاهی اطلاعات کمی وجود دارد. در این رابطه، خانواده nucleosome assembly protein (NAP) در تمام یوکاریوت‌ها حفاظت شده بوده و جزء جدایی ناپذیر در پایداری، حفظ و پویایی کروماتین یوکاریوتی می‌باشد. این پروتئین‌ها در انتقال هیستون‌ها به هسته، تشکیل نوکلئوزوم و القاء سیالیت کروماتین نقش داشته و لذا رونویسی بسیاری از ژن‌ها را تحت تأثیر قرار می‌دهد. در این مطالعه با استفاده از روش‌های بیوانفورماتیکی، ۶ ژن شبه *NAP* (*ZmNPL1*) تا *ZmNAPL6* در ذرت شناسایی شد. آنالیز فیلوژنتیکی نشان داد که این ژن‌های *NAPL* همانند ژن‌های *NAPL* آراییدوسیس و برنج به دو زیر گروه تقسیم شده و رابطه تکاملی نزدیکتری با ژن‌های *NAPL* برنج داشتند. این ژن‌ها دارای ۳ تا ۱۱ اینترون بوده و بر روی ۵ کروموزوم از ۱۰ کروموزوم ذرت قرار گرفته‌اند. آنالیز بیانی بر پایه ریزآرایه نشان دهنده تنظیم دقیق رونویسی ژن‌های *ZmNAPL* در طول نمو ذرت می‌باشد. این امر حاکی از نقش مهم این ژن‌ها در برنامه ریزی مرتبط با فرآیندهای نمو ذرت بود. این مطالعه اولین گزارش در مورد شناسایی و بررسی روابط تکاملی، ساختاری و بیانی ژن‌های *NAPL* ذرت بوده و نتایج بدست آمده از آن اطلاعات پایه برای تحقیقات آتی در مورد کاربرد ژن‌های *NAPL* ذرت را مهیا می‌سازد.

**واژه‌های کلیدی:** آنالیز فیلوژنتیکی، بیان ژن، بیوانفورماتیک، نوکلئوزوم.

## مقدمه

نوکلئوزوم واحد ساختاری کروماتین بوده و شامل ۱۴۷ جفت باز DNA دو رشته‌ای می‌باشد که دور یک اوکتامر متشکل از تترامر<sub>2</sub> (H3-H4) و دو دایمر H2A-H2B پیچیده شده است (Valieva *et al.*, 2016). جاذبه ذاتی بین بار منفی گروه فسفات مولکول DNA و بار مثبت پروتئین‌های هیستونی غنی از آرژینین و لیزین عامل اصلی ایجاد نوکلئوزوم است. با این حال، در شرایط آزمایشگاهی ترکیب این دو مولکول رسوب‌های غیر محلول تولید کرده و نوکلئوزوم ایجاد نمی‌شود. در نتیجه فاکتورهای دیگری برای میان‌کنش صحیح و کنترل شده هیستون و DNA و ایجاد نوکلئوزوم نیاز است (Dennehey and Tyler, 2014). عوامل اتصال به DNA مانند فاکتورهای رونویسی و عوامل درگیر با هیستون‌های کروماتینی قابلیت تغییر ساختار نوکلئوزوم را دارند (Eitoku *et al.*, 2008). عوامل درگیر با هیستون‌های کروماتینی شامل آنزیم‌های کاتالیز کننده تغییرات کووالانسی هیستونی نظیر هیستون استیل ترانسفراز، هیستون داستیلاز و هیستون ترانسفراز می‌باشند. همچنین فاکتورهای بازسازی کروماتینی وابسته به ATP<sup>۱</sup> و چپرون‌های هیستونی نیز جزء این گروه هستند (De Koning *et al.*, 2007). چپرون‌های هیستونی توانایی اتصال به هیستون‌های H2A/H2B یا H3/H4 و نیز هیستون لینکر H1 را دارند (Hondele and Ladurner, 2011). علاوه بر این، کلاس‌های خاصی از این چپرون‌ها توانایی اتصال به بیش از یک نوع هیستون را دارند. همچنین چندین نوع چپرون هیستونی قادر به اتصال به یک نوع هیستون هستند اما نقش آن‌ها در فرآیندهای بیولوژیکی سلول متفاوت است (Dennehey and Tyler, 2014). از نظر تکاملی

بسیاری از خانواده‌های چپرون‌های هیستونی در بین یوکاریوت‌ها حفاظت شده است (Hondele and Ladurner, 2011). بر اساس مطالعات انجام شده در مورد چپرون‌های هیستونی انسان و مخمر، این پروتئین‌ها شامل خانواده‌های NAP، NPM، CAF1، ASF1، HIRA، FACT، NASP و SPT6 می‌باشند (Tripathi *et al.*, 2015). کارکرد فیزیولوژیکی چپرون‌های هیستونی در موجودات مختلف به صورت کامل مشخص نشده است. با این حال مطالعات نشان داده است که جهش در چپرون‌های هیستونی موجب نقص در پایداری ژنوم و بیان ژن می‌شود (Burgess and Zhang, 2013).

پروتئین NAP1 نقش اصلی را در انتقال هیستون‌های سنتز شده به داخل هسته دارد که پیش نیاز تشکیل نوکلئوزوم می‌باشد (Mosammamarast *et al.*, 2002). علاوه بر نقش NAP1 در ایجاد نوکلئوزوم و اتصال به هیستون‌ها، مشخص شده است که این پروتئین‌ها در رونویسی و نیز سیکل سلولی نقش دارند (Shimizu *et al.*, 2000). سرکوب ژن *NAP1-2* موجب مرگ در مرحله جنینی موش می‌شود. در مخمر آنالیز ژن‌های *NAP* نشان داد که ده درصد کل ژن‌ها در اثر سرکوب ژن *NAP-1* تحت تأثیر قرار می‌گیرند (Park and Luger, 2006). همچنین پروتئین NAP1 مخمر با پروتئین‌های سیکلین، کیناز G4 و NBP میان‌کنش داشته و در کنترل میتوز نقش دارد (Kellogg and Murray, 1995). بررسی جایگاه درون سلولی پروتئین NAP1 با روش برچسب زنی فلورسنتی<sup>۱۰</sup>

2. Nucleoplasmin/Nucleophosmin

3. Nucleosome assembly protein

4. Chromatin assembly factor

5. Anti-silencing factor 1

6. Histone regulatory homolog A

7. Facilitates chromatin transcription

8. Nuclear Autoantigenic Sperm Protein

9. Suppressor of Ty element 6

10. Fluorescent Taging

1. ATP-dependent chromatin remodeling factors

هدف شناسایی خانواده‌های ژنی چپرون‌های هیستونی برنج و آراییدوپسیس انجام گرفت به ترتیب ۷ و ۶ ژن *NAP* شناسایی و ویژگی‌های تکاملی، ساختاری و بیانی آن‌ها بررسی شد (Tripathi et al., 2015). تاکنون بررسی جامع و کاملی از روابط تکاملی و کارکردی خانواده ژنی *NAP* ذرت انجام نشده است. در این مطالعه ۶ عضو خانواده ژنی *NAP* در ذرت شناسایی و خصوصیات بیوشیمیایی، جایگاه سلولی، روابط فیلوژنتیکی، موتیف‌های حفاظت شده و ساختار ژنی آن‌ها با استفاده از روش‌های بیوانفورماتیکی تعیین شد. در نهایت الگوی بیانی ژن‌های *NAP* ذرت در بافت‌ها و مراحل نموی مختلف با استفاده از داده‌های ریزآرایه موجود در پایگاه داده‌های بیانی بررسی شد. این مطالعه اولین گزارش در مورد در مورد شناسایی و بررسی روابط تکاملی، ساختاری و بیانی خانواده ژنی *NAP* ذرت بوده و نتایج آن می‌تواند اطلاعات پایه برای درک بهتر کارکرد و روابط تکاملی خانواده ژنی *NAP* ذرت را فراهم کرده و استفاده از آن‌ها در تولید گیاهان ذرت تراریخته مقاوم به تنش‌های زیستی و غیر زیستی و نیز افزایش عملکرد این گیاه را تسهیل نماید.

## مواد و روش‌ها

### شناسایی ژن‌های *NAP* ذرت

برای شناسایی اعضاء خانواده ژنی *NAP* در ذرت، توالی پروتئینی ژن‌های *NAP* برنج و آراییدوپسیس (Tripathi et al., 2015) به ترتیب با استفاده از ۲ پایگاه داده RGAP7 (Kawahara et al., 2013) و TAIR10 (Lamesch et al., 2012) دریافت و از روش tBLASTn (McGinnis and Maden, 2004) جهت جستجوی ژن‌های *NAP* در پایگاه داده ژنوم ذرت در سایت فایتوزوم استفاده شد (Goodstein et al., 2012). پس از دریافت تمام توالی‌ها، توالی‌های تکراری حذف و در نهایت تأیید

نشان داد که این پروتئین در فاز S چرخه سلولی در هسته و در فاز G2 در سیتوپلاسم قرار دارد. این امر نشان دهنده درگیری *NAP1* در انتقال هیستون بین سیتوپلاسم و هسته می‌باشد (Marheineke and Krude, 1998). همچنین مشخص شده است که *NAP1* با پروتئین مسئول انتقال H2A و H2B به هسته (ایمپورتین Kap114p) دارای میانکنش بوده و نقش این پروتئین در انتقال هیستون‌ها به هسته را تأیید می‌کند (Park and Luger, 2006).

مطالعه انجام شده بر روی سه ژن *NAP* آراییدوپسیس نشان داد که سرکوب همزمان این سه ژن موجب فوق حساسیت آراییدوپسیس به UV-C می‌شود (Liu et al., 2009). پروتئین *OsNAPL1* برنج چپرون اختصاصی H3/H4 بوده و در هسته قرار دارد. با استفاده از لاین‌های سرکوب و فرابیان شده این ژن یک رابطه مثبت بین سطوح بیان این ژن و مقاومت برنج به تنش‌های غیر زیستی مشخص شد. این پروتئین موجب القاء بیان ژن‌های پاسخ به تنش‌های غیرزیستی نظیر *OsRad51* شده و از این طریق مقاومت برنج به تنش‌های غیرزیستی را افزایش می‌دهد (Tripathi et al., 2016). آگاهی ما در مورد چپرون‌های هیستونی در گیاهان و نقش فیزیولوژیکی آن‌ها زیاد نیست. علاوه بر این مکانیسم تنظیمی تشکیل و یا تغییر ساختار نوکلئوزوم در پاسخ به نیازهای متعدد سلولی در گیاهان و نیز نقش کلاس‌های مختلف چپرون‌های هیستونی در مراحل نمو و یا شرایط تنش تا حدود زیادی مبهم باقی مانده است. امروزه اولین قدم در شناسایی و مطالعه کارکرد ژن‌های مختلف استفاده از داده‌های موجود در پایگاه‌های اطلاعاتی با استفاده از ابزار بیوانفورماتیکی می‌باشد. بیوانفورماتیک علم دسته‌بندی، سازمان‌دهی، آنالیز، تفسیر و استفاده از اطلاعات مولکول‌ها و توالی‌های زیستی است. هدف اصلی بیوانفورماتیک افزایش درک و فهم از فرآیندهای زیستی می‌باشد (Hogeweg, 2011). در مطالعه *in silico* که با

در نامگذاری ژن‌های *NAP* ذرت، ابتدا پسوند Zm از *Zea mays* و سپس NAP و در نهایت پسوند L از Like بعلت تشابه پایین آن‌ها به ژن‌های *NAP* انسان به آن اضافه شد (Tripathi *et al.*, 2015). شماره‌گذاری ژن‌ها بر اساس جایگاه آن‌ها بر روی کروموزوم‌های ۱ تا ۱۰ و از بالا به پایین انجام گرفت. آدرس پایگاه داده‌های استفاده شده در این تحقیق در جدول ۱ ارائه شده است.

وجود دومین *NAP* (PF00956) در توالی‌های شناسایی شده با استفاده از پایگاه داده Pfam انجام گرفت (Finn *et al.*, 2013). وزن مولکولی و نقطه ایزوالکتریک (pI) تئوریتیکال پروتئین‌های NAP با استفاده از ابزار ProtParam سایت ExpASY محاسبه شد (Artimo *et al.*, 2012). برای شناسایی جایگاه سلولی پروتئین‌ها از برنامه برخط CELLO استفاده شد (Yu *et al.*, 2006).

جدول ۱. اسامی و آدرس پایگاه داده‌های استفاده شده در این پژوهش

پایگاه داده	URL	رفرنس
RGAP7	<a href="http://rice.plantbiology.msu.edu/">http://rice.plantbiology.msu.edu/</a>	Kawahara <i>et al.</i> , 2013
TAIR10	<a href="http://www.arabidopsis.org/">http://www.arabidopsis.org/</a>	Lamesch <i>et al.</i> , 2012
Phytozome	<a href="https://phytozome.jgi.doe.gov/">https://phytozome.jgi.doe.gov/</a>	Goodstein <i>et al.</i> , 2012
CELLO	<a href="http://cello.life.nctu.edu.tw/">http://cello.life.nctu.edu.tw/</a>	Yu <i>et al.</i> , 2006
Conserved domain database	<a href="https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi">https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi</a>	Marchler-Bauer <i>et al.</i> , 2011
Expasy	<a href="http://web.expasy.org/compute_pi/">http://web.expasy.org/compute_pi/</a>	Artimo <i>et al.</i> , 2012
GSDS	<a href="http://gsds.cbi.pku.edu.cn/">http://gsds.cbi.pku.edu.cn/</a>	Hu <i>et al.</i> , 2015
MEME	<a href="http://meme-suite.org/">http://meme-suite.org/</a>	Bailey <i>et al.</i> , 2006
Pfam	<a href="http://pfam.xfam.org/">http://pfam.xfam.org/</a>	Finn <i>et al.</i> , 2013
SMART	<a href="http://smart.embl-heidelberg.de/">http://smart.embl-heidelberg.de/</a>	Letunic <i>et al.</i> , 2012
MeV4.0	<a href="http://www.tm4.org/mev.html/">http://www.tm4.org/mev.html/</a>	-
PLEXdb	<a href="http://www.plexdb.org/">http://www.plexdb.org/</a>	Dash <i>et al.</i> , 2012
STRING v.10	<a href="http://string-db.org/">http://string-db.org/</a>	Szklarczyk <i>et al.</i> , 2016
DAVID 6.8	<a href="https://david.ncifcrf.gov/">https://david.ncifcrf.gov/</a>	Huang <i>et al.</i> , 2009

شناسایی موتیف‌های اختصاصی این خانواده از نرم‌افزار Multiple Em for Motif (MEME Elicitation) استفاده شد (Bailey *et al.*, 2006). پارامترهای مورد استفاده شامل شناسایی ۵ موتیف و حداقل و حداکثر طول موتیف‌ها به ترتیب ۶ و ۵۰ اسید آمینه بود. پایگاه داده Pfam و برنامه SMART برای بررسی موتیف‌های شناسایی شده استفاده شد (Finn *et al.*, 2012; Letunic *et al.*, 2013).

#### بررسی دیجیتالی بیان ژن‌های *NAP* ذرت

آنالیز بیان ژن‌های *NAP* در مراحل مختلف نمو ذرت با استفاده از داده‌های ترانسکریپتوم اطلس بیانی اینبرد لاین B73 ذرت (تکنولوژی ریزآرایه NimbleGen) انجام شد. این داده‌ها شامل پروفایل بیانی ۶۰ بافت متفاوت مربوط به ۱۱ اندام ذرت شامل بذر درحال جوانه‌زنی، ریشه اولیه، گیاهچه کامل، ساقه، مریستم انتهایی ساقه، چوب بلال،

#### روابط تکاملی خانواده ژنی *NAP* در ذرت

برای رسم درخت فیلوژنتیکی، ابتدا هم‌ردیف‌سازی طول کامل پروتئین‌های NAP ذرت، برنج، آراییدوپسیس و انسان با استفاده از نرم‌افزار ClustalX 2.0.8 (Larkin *et al.*, 2007) انجام و سپس رسم درخت فیلوژنتیکی با نرم‌افزار MEGA7 (Kumar *et al.*, 2016) بر اساس روش اتصال همسایه‌ها [Neighbor Joining (N.J)] رسم و سطح اطمینان از صحت گروه‌بندی با آزمون بوت‌استرپ (Felsenstein, 1985) با تکرار ۱۰۰۰ تخمین زده شد.

#### تجزیه و تحلیل ساختار اگزون- اینترونی و شناسایی موتیف‌های حفاظت شده

الگوی توزیع اینترون‌ها از طریق مقایسه توالی CDS نسبت به DNA ژنومی مربوط به هر ژن با استفاده از سرور GSDS انجام گرفت (Hu *et al.*, 2015). برای

انواع چپرون‌های هیستونی را تشکیل داده بودند ولی به علت تشابه پایین آن‌ها به ژن‌های شناخته شده NAP انسان از پسوند Like برای نام‌گذاری آن‌ها استفاده کردند (Tripathi *et al.*, 2015). لذا شش ژن *NAP* شناسایی شده در ژنوم ذرت پسوند Like گرفته و بر اساس جایگاه کروموزومی از *ZmNAPL1* تا *ZmNAPL6* نامگذاری شدند (جدول ۲). این شیوه شماره‌دهی به ژن‌ها بیشتر در مورد پروتئین‌های NAC سبب زمینی، WRKY ذرت، NCX آرآبیدوپسیس و برنج نیز استفاده شده است (Singh *et al.*, 2013; Wei *et al.*, 2012; Singh *et al.*, 2015).

از شش ژن *NAPL* شناسایی شده ذرت ژن‌های *ZmNAPL1*، *ZmNAPL2*، *ZmNAPL5* و *ZmNAPL6* به ترتیب بر روی کروموزوم‌های ۴، ۵، ۸ و ۹ قرار دارند. همچنین جایگاه ژن‌های *ZmNAPL1* و *ZmNAPL1* بر روی کروموزوم ۶ می‌باشد (جدول ۲). مطالعات فیزیکی‌شیمیایی این خانواده ژنی با استفاده از ابزار تجزیه و تحلیل پروتئین ExpASY (Artimo *et al.*, 2012) نشان داد که این ژن‌ها از نظر تعداد آمینواسید، وزن مولکولی و نقطه ایزوالکتریک باهم تفاوت دارند. طول پروتئین‌های این خانواده از ۱۵۹ تا ۳۸۰ اسید آمینه و وزن مولکولی آن‌ها از ۱۸/۳ تا ۴۳/۰۹ کیلو دالتون است.

دامنه نقطه ایزوالکتریک این پروتئین‌ها اسیدی بوده و کمترین و بیشترین نقطه ایزوالکتریک با ۴/۰۸ در *ZmNAPL2* و ۵/۲۵ در *ZmNAPL4* است. نکته جالب توجه آنکه تنها ویژگی مشترک انواع کلاس‌های چپرون‌های هیستونی ماهیت اسیدی آن‌ها می‌باشد. مکان‌یابی پروتئین‌ها *ZmNAPL* مشخص کرد که این پروتئین‌ها در هسته فعال هستند که با کارکرد آن‌ها مطابقت دارد (جدول ۲).

میانگه، ابریشم، تاسل و دانه گرده، برگ، غلاف بلال و بذر می‌باشد (Sekhon *et al.*, 2011). داده‌های نرمال‌شده با روش Robust Multi-array Average (RMA) و بر مبنای Log2 به شماره دسترسی ZM29 از پایگاه داده بیانی گیاهی PLEXdb دریافت (Dash *et al.*, 2012) و از نرم افزار Mev4.0 برای خوشه‌بندی داده‌های بیانی بر اساس روش اقلیدسی و الگوریتم Complete Linkage و رسم نقشه حرارتی (Heat Map) استفاده شد.

**بررسی شبکه تعامل پروتئین-پروتئین (PPI) و هستی‌شناختی (GO) خانواده ژنی NAP ذرت**  
تعامل پروتئین-پروتئین [Protein-Protein Interaction Network (PPI)] سیستم‌های زیستی بوده و آنالیز الگوهای میان‌کنش پروتئین‌ها اطلاعات با ارزشی را در مورد کارکرد پروتئین‌ها ارائه می‌دهد (Braun *et al.*, 2013). به منظور بررسی شبکه تعامل پروتئین-پروتئین ژن‌های *NAP* ذرت از پایگاه داده STRING v.10 بر اساس پارامترهای پیش فرض استفاده شد (Szkarczyk *et al.*, 2016). مطالعه هستی‌شناختی [Gene Ontology (GO)] ژن‌های *ZmNAPL* با استفاده از پایگاه داده DAVID 6.8 و بر اساس پارامترهای پیش فرض انجام گرفت (Huang *et al.*, 2009).

## نتایج و بحث

### شناسایی و بررسی ویژگی‌های خانواده ژنی *NAPL* ذرت

از طریق جستجوی tBLASTn توالی‌های شناخته شده *NAP* در آرآبیدوپسیس و برنج، ۶ ژن کدکننده *NAP* در ژنوم ذرت شناسایی شد. در مطالعه‌ای که برای شناسای چپرون‌های هیستونی برنج و آرآبیدوپسیس انجام گرفت به ترتیب ۷ و ۶ ژن *NAPL* شناسایی شد که بزرگ‌ترین گروه در بین

### بررسی روابط تکاملی ژن‌های *ZmNAPL*

برای درک و شناخت رابطه تکاملی ژن‌های *NAPL* ذرت، درخت فیلوژنتیکی بر مبنای الگوریتم N.J. با استفاده از توالی پروتئینی ژن‌های *NAPL* آراییدوپسیس، برنج، ذرت و انسان رسم شد (شکل ۱). در درخت رسم شده کلاستر یک که تنها از ژن‌های گیاهی تشکیل شده است بیشترین عضو را دارد. ژن‌های *NAP* انسان به استثنای ژن *HsSET* در کلاستر دو قرار گرفته‌اند که نشان‌دهنده همولوژی پایین و مسیر تکاملی متفاوت این ژن‌ها با ژن‌های *NAPL* گیاهی است (شکل ۱).

بر اساس مطالعات کارکرد ژن *HsSET* از سایر ژن‌های *NAP* انسان متفاوت بوده و بیشتر در تنظیم رونویسی نقش دارد (Gamble *et al.*, 2005). به نظر می‌رسد کارکرد ژن‌های اورتولوگ گونه‌های مختلف در هر کلاستر به احتمال زیاد و نه حتماً مشابه باشد در حالی که ژن‌های پارالوگ شاید برای نقش جدید نسبت به ژن جد خود تکامل یابند (Guo *et al.*, 2008). لذا به نظر می‌رسد ژن‌های گیاهی که همراه ژن *HsSET* در کلاستر سه قرار گرفته‌اند احتمالاً کارکردی شبیه به این ژن انسانی داشته

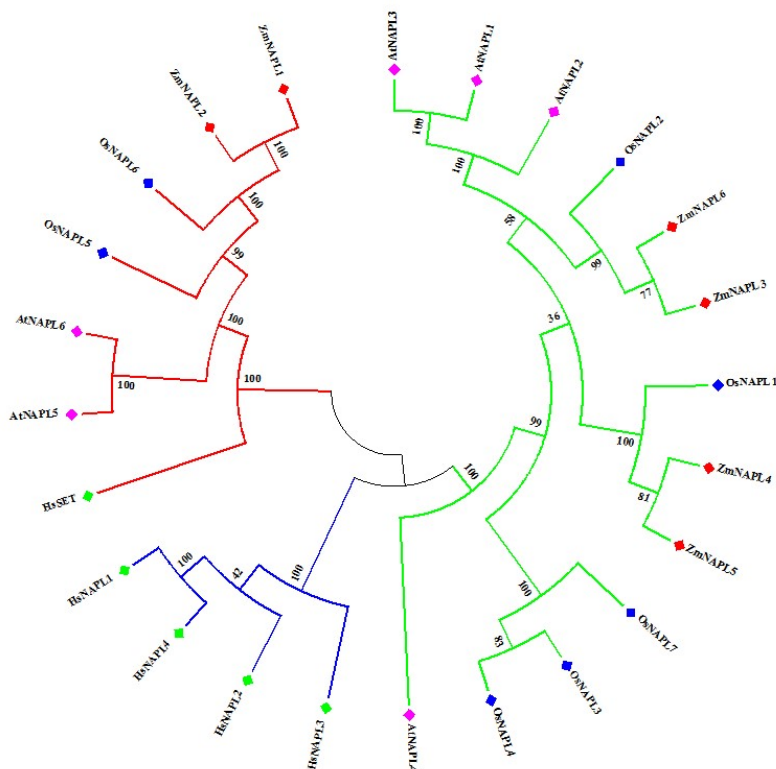
باشند. باید توجه کرد در کلاسترهای یک و سه که از ژن‌های *NAPL* گیاهی ایجاد شده‌اند، ژن‌های *NAPL* ذرت و برنج با توجه به تک لپه‌ای بودن این دو گیاه و دارا بودن رابطه تکاملی نزدیک‌تر نسبت به آراییدوپسیس دو لپه‌ای در زیرگروه‌های مشترک قرار گرفته‌اند (شکل ۱).

### بررسی ساختار ژنی و شناسایی موتیف‌های حفاظت‌شده

بررسی ساختار اگزون-اینترونی می‌تواند شواهد بیشتری در تأیید گروه بندی درخت فیلوژنتیکی و روابط تکاملی ارائه نماید زیرا این نوع ساختاری اغلب نقش مهمی در تکامل خانواده‌های ژنی دارد (Zhang *et al.*, 2012). باید توجه داشت اعضاء بسیار نزدیک در هر گروه فیلوژنی از نظر ساختار اگزون-اینترونی یعنی تعداد اینترون و طول اگزون شبیه هستند. علاوه بر این، توزیع موتیف‌ها در پروتئین‌های هر گروه نیز مشابه است که به‌عنوان یک دلیل در مورد رابطه تکاملی نزدیک بین آن‌ها می‌باشد.

جدول ۲. ویژگی‌های خانواده ژنی *NAPL* در ذرت

لوکوس ژن	اسم ژن	اسم پروتئین	وزن مولکولی	pI	طول پروتئین	جایگاه سلولی	کروموزوم
GRMZM2G157019	<i>ZmNAPL1</i>	ZmNAPL1	28.44	4.16	251	nuclear	4
GRMZM2G075637	<i>ZmNAPL2</i>	ZmNAPL2	29.33	4.08	258	nuclear	5
GRMZM6G790713	<i>ZmNAPL3</i>	ZmNAPL3	18.3	4.13	159	nuclear	6
GRMZM2G121186	<i>ZmNAPL4</i>	ZmNAPL4	42.8	5.25	372	nuclear	6
GRMZM2G176707	<i>ZmNAPL5</i>	ZmNAPL5	42.39	4.34	373	nuclear	8
GRMZM2G140051	<i>ZmNAPL6</i>	ZmNAPL6.1	43.09	4.31	380	nuclear	9
GRMZM2G140051	<i>ZmNAPL6</i>	ZmNAPL6.2	41.96	4.32	369	nuclear	9



شکل ۱. رابطه فیلوژنتیکی پروتئین‌های NAPL ذرت، برنج، آرابیدوپسیس و انسان. هم‌ردیف‌سازی چندگانه با نرم‌افزار ClustalX انجام و درخت به کمک نرم‌افزار MEGA7 بر اساس روش اتصال همسایه‌ها با بوت‌استرپ هزار رسم شد. گروه‌های ۱، ۲ و ۳ به ترتیب با رنگ‌های سبز، آبی و قرمز مشخص شده‌اند. پروتئین‌های NAPL ذرت، برنج، آرابیدوپسیس و انسان به ترتیب با رنگ‌های قرمز، آبی، صورتی و سبز نشان داده شده است.

*ZmNAPL6* دو نوع پروتئین را کد می‌کند (شکل ۲).

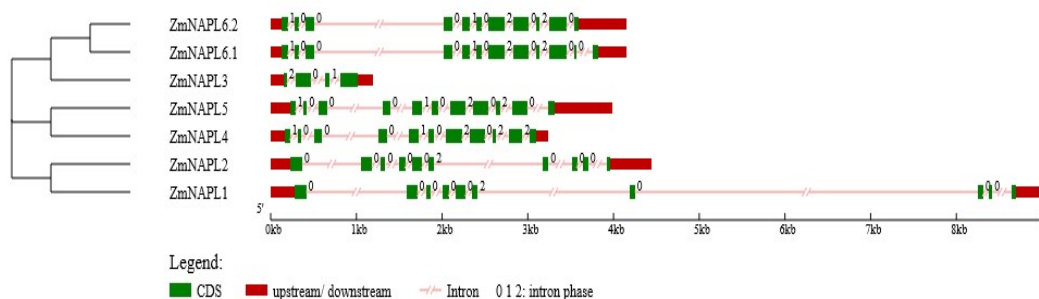
اسپلایسینگ اینترون بر اساس جایگاه انجام آن به سه فاز تقسیم می‌شود. فاز ۱ که در آن اسپلایسینگ در اولین نوکلئوتید بعد از کدون انجام می‌شود. فاز دو در دومین نوکلئوتید بعد از کدون و فاز صفر نیز در سومین نوکلئوتید بعد از کدون ایجاد می‌شود (Sharp, 1981). هر سه فاز اسپلایسینگ اینترونی در خانواده ژنی *NAPL* ذرت مشاهده می‌شود. مقایسه ساختار ژنی و درخت فیلوژنتیکی نشان داد که بجز ژن *ZmNAPL3*، ساختار ژنی و فازهای اسپلایسینگ و نیز الگوی این فازها در هر کلاستر ژنی مشابه می‌باشد (شکل ۲).

برای ارزیابی وجود موتیف‌های حفاظت شده در داخل این خانواده ژنی، توالی پروتئینی ژن‌های

همچنین جایگاه موتیف‌های مختلف بر روی توالی پروتئین مربوطه احتمالاً در بررسی اختلاف عملکرد ژن‌ها در زیر خانواده‌های مختلف کمک می‌کند (Liao *et al.*, 2016). بررسی ساختار اگزون اینترونی خانواده ژنی *ZmNAPL* نشان داد که تعداد اینترون این ژن‌ها به استثنای ژن *ZmNAPL3* بالا بوده و ۹ تا ۱۱ اینترون دارند. ژن *ZmNAPL3* دارای ۳ اینترون است. به دلیل طول زیاد اینترون‌ها در ژن *ZmNAPL1*، این ژن بیشترین طول را در سطح ژنومی دارد. ژن *ZmNAPL6* دارای دو سیستم اسپلایسینگ است. با توجه به اینکه تفاوت بین دو نوع اسپلایسینگ این ژن بر روی اگزون‌ها می‌باشد، لذا موجب تغییر توالی کدکننده شده و لذا ژن

گروه‌بندی آن‌ها در درخت فیلوژنتیکی مطابقت دارد. به نظر می‌رسد تفاوت ساختاری و نیز نوع قرار گیری موتیف‌های حفاظت شده ژن *ZmNAPL3* به دلیل مسیر تکاملی و نقش کارکردی متفاوت آن باشد که برای نتیجه‌گیری صحیح نیاز به تحقیقات بیشتر دارد (شکل ۳).

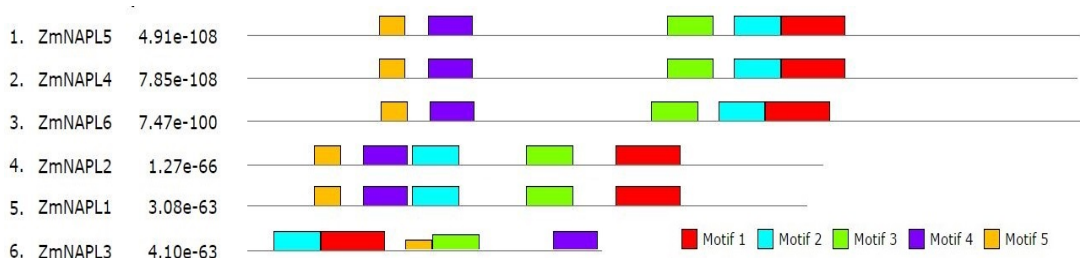
*ZmNAPL* به‌وسیله نرم‌افزار MEME مورد ارزیابی قرار گرفت. ارزیابی موتیف‌ها با Pfam و SMART نشان داد که به غیر از موتیف شماره پنج سایر موتیف‌ها مربوط به دومین *NAP* (PF00956) می‌باشد (جدول ۳). به‌غیر از ژن *ZmNAPL3* در سایر ژن‌ها، موتیف‌های حفاظت شده و الگوی قرار گیری آن‌ها بر روی هر پروتئین *ZmNAPL*



شکل ۲. ساختار اگزون- اینترونی ژن‌های NAPL ذرت مطابق رابطه فیلوژنتیکی آن‌ها، در شکل بخش‌های سبز، صورتی و قرمز به ترتیب معرف ناحیه کد کننده ژن، اینترون و ناحیه‌های 5' UTR و 3' UTR می‌باشد.

جدول ۳. موتیف‌های حفاظت شده پروتئین‌های NAPL ذرت

کارکرد موتیف	توالی	طول	امتیاز	فراوانی	شماره موتیف
NAP domain	PRKGSKNTKPITKTEDCESFFNFFSPPQV	29	7.60E-54	5	1
NAP domain	IGTEIEWYPGKCLTQKILKKK	21	1.80E-32	5	2
NAP domain	HFGTNPYFRNSVLTKTYHMVD	21	5.50E-24	5	3
NAP domain	RMALEAKYQKLYRPLYMKRY	20	9.90E-24	5	4
-	RVEKLRLEIQGEH	12	4.80E-07	5	5



شکل ۳. موتیف‌های حفاظت شده پروتئین‌های NAPL ذرت که با MEME شناسایی شده‌اند. موتیف‌های ۱ تا ۵ با رنگ‌های مختلف مشخص شده‌اند. هر موتیف نشانگر بخشی از دامین کارکردی NAP بوده (به استثناء موتیف ۵) و توالی‌ها، امتیاز موتیف و طول آن‌ها در جدول ۳ ارائه شده است.

آن ژن باشد. برای ارزیابی پروفایل بیانی ژن‌های NAPL ذرت، بیان این ژن‌ها در بافت‌ها و مراحل نموی مختلف ذرت با استفاده از داده‌های منتشر شده در

بررسی بیان ژن‌های *ZmNAPL* در مراحل نموی ذرت الگوی بیان ژن‌ها می‌تواند سرنخ مهمی در مورد عملکرد



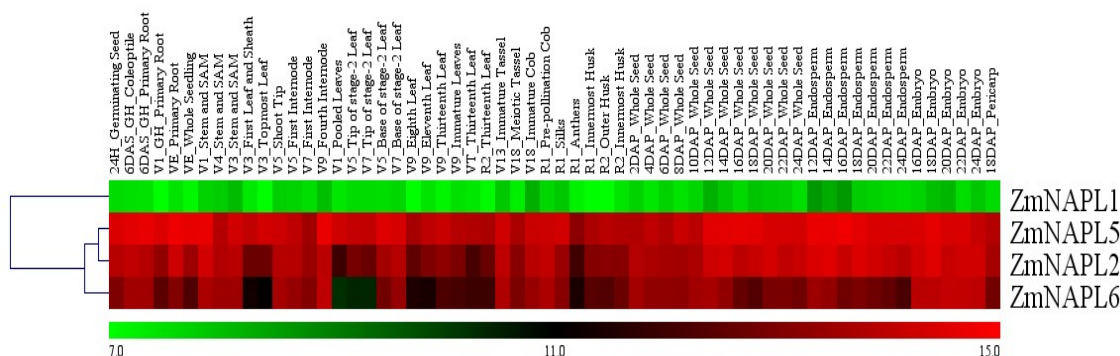
### بررسی شبکه تعامل پروتئین-پروتئین و هستی شناسی ژن‌های *ZmNAPL*

پروتئین‌ها به‌عنوان عوامل اصلی فرآیندهای بیولوژیکی اغلب نیازمند میان‌کنش با سایر پروتئین‌ها هستند تا کارکرد بیولوژیکی خود را ایفا کنند. تعامل پروتئین-پروتئین در تمام فرآیندهای زیستی نظیر همانندسازی DNA، رونویسی، پیام‌رسانی، تخریب پروتئین و تنظیم چرخه سلولی مشاهده می‌شود. لذا، شناسایی شبکه تعامل پروتئین-پروتئین برای درک فرآیندهای بیولوژیکی بسیار با ارزش است (Zhu et al., 2011).

از شش پروتئین *ZmNAPL* تنها برای پروتئین *ZmNAPL3* در پایگاه داده STRING داده‌های شبکه تعامل پروتئین-پروتئین شناسایی نشد. بررسی سایر پروتئین‌های *ZmNAPL* نشان داد که پروتئین‌های درگیر در شبکه تعامل پروتئین-پروتئین سه پروتئین *ZmNAPL4*، *ZmNAPL5* و *ZmNAPL6* کاملاً مشابه بوده و لذا ده ژن با بالاترین امتیاز میان‌کنش برای این سه پروتئین مشترک هستند (جدول ۴). همچنین ۹ پروتئین از پروتئین‌هایی که بالاترین امتیاز میان‌کنش را برای شبکه تعامل پروتئین-پروتئین *ZmNAPL1* و *ZmNAPL2* دارند مشابه بوده و تنها یک پروتئین در شبکه تعامل پروتئین-پروتئین این دو پروتئین باهم متفاوت است (جدول ۴).

پایگاه‌های داده عمومی بررسی شد. برای ژن‌های *ZmNAPL3* و *ZmNAPL4* پروب‌ست ID یافت نشد. نقشه حرارتی مربوط به پروفایل بیانی سایر ژن‌های *ZmNAPL* در شکل ۴ ارائه شده است. الگوی بیان این چهار ژن را می‌توان به سه گروه تقسیم کرد. گروه اول شامل ژن *ZmNAPL1* است که در تمامی بافت‌ها و مراحل نمو بیان پایدار و پایین دارد. گروه دوم شامل ژن‌های *ZmNAPL2* و *ZmNAPL5* است که تقریباً بیان ثابت و بالایی را در تمامی بافت‌ها و مراحل نمو دارد. ژن *ZmNAPL6* گروه سوم را تشکیل داده است که بیان آن در بافت‌ها پایین تا متوسط است. به‌عنوان مثال بیان این ژن در دانه گرده تغییر نکرده است در حالی که بیان آن در برگ‌ها پایین یا بدون تغییر است (شکل ۴).

مطالعه الگوی بیان هیستون‌های چپرونی در آراییدوپسیس و برنج نشان داد که ژن‌های *NAPL* این دو گیاه نیز در بافت‌ها و مراحل نمو الگوی بیانی مشابه‌ای با ژن‌های *NAPL* ذرت دارند (Tripathi et al., 2015). با توجه به نقش ژن‌های *NAPL* در تغییرات کروماتینی و تأثیر آن بر روی بیان بسیاری از ژن‌ها در مراحل مختلف نمو و نیز بافت‌ها به نظر می‌رسد که دو ژن *ZmNAPL5* و *ZmNAPL2* نقش مهمی در فرآیندهای بیولوژیکی ذرت داشته باشند. در مورد نقش ژن‌های *ZmNAPL* در پاسخ به تنش‌های زیستی و غیرزیستی مطالعات اندکی انجام گرفته است. مطالعه پروفایل بیانی گیاهچه‌های ذرت تحت تنش حرارتی نشان داد که بیان ژن *ZmNAPL1* در پاسخ به این تنش بصورت معنی‌داری افزایش می‌یابد که می‌تواند حاکی از نقش مهم این ژن در پاسخ ذرت به تنش حرارتی باشد (Frey et al. 2015).



شکل ۴. پروفایل بیانی ژن‌های *NAPL* ذرت بر اساس داده‌های ریزآرایه در مراحل مختلف نمو و بافت‌ها. نقشه حرارتی تغییر بیان ژن‌های *NAPL* ذرت را نشان می‌دهد. ارزش سیگنال به صورت نوار رنگی در انتهای نقشه حرارتی ارائه شده است که در آن رنگ سبز نشان دهنده سرکوب بیان، رنگ سیاه نشان دهنده عدم تغییر بیان و رنگ قرمز نشان دهنده القاء بیان است.

داد که ۵ ژن *ZmNAPL* (به استثناء *ZmNAPL3*) در یک کلاستر قرار گرفته‌اند. این کلاستر شامل پایگاه داده INTERPRO و ۱۲ گروه GO برای فرآیندهای بیولوژیکی شامل DNA packaging، chromatin assembly، nucleosome assembly، protein-DNA، nucleosome organization، chromatin assembly or، complex assembly، chromatin organization، disassembly، cellular، chromosome organization، macromolecular complex assembly، cellular macromolecular complex subunit، macromolecular complex organization و assembly می‌باشد که به صورت معنی‌دار در این کلاستر قرار گرفته‌اند. برای ژن *ZmNAPL3* اطلاعاتی در این پایگاه داده شناسایی نشد. در جدول ۵ گروه‌های شناسایی شده، شماره GO، توصیف GO و P-Value مربوط به هر یک از گروه‌های هستی شناسی ارائه شده است. این نتایج نشان‌دهنده مشترک بودن فعالیت‌های بیولوژیکی ژن‌های *NAPL* ذرت در فرآیندهای مربوط به نوکلئوزوم نظیر تشکیل نوکلئوزوم و سیالیت کروماتین می‌باشد.

همچنین بر اساس این داده‌ها دو پروتئین *ZmNAPL1* و *ZmNAPL2* با پروتئین‌های *ZmNAPL4*، *ZmNAPL5* و *ZmNAPL6* میانکنش دارند. مطالعه میان‌کنش پروتئین‌های NAP در موش با استفاده از تکنیک two-hybrid نشان داد که پروتئین‌های NAP1L1، NAP1L2 و NAP1L4 دارای میان‌کنش بسیار شدید بوده و در بسیاری از فرآیندهای بیولوژیکی نظیر آستیل‌اسیون هیستون، تنظیم در دسترس بودن کروماتین و تسهیل اتصال سایر پروتئین‌ها به کروماتین در موش به صورت مشترک دخالت دارند (Attia *et al.*, 2015). بررسی ژن‌های درگیر در ایجاد شبکه پروتئین-پروتئین نشان داد که کارکرد برخی از این ژن‌ها نامشخص است. اما با بررسی زمین‌های موجود می‌توان گفت که کارکردی مشابه به پروتئین‌های هیستونی و یا ایمپورتین‌ها دارند. مشخصات، کارکرد و زمین ژن‌های درگیر در شبکه تعامل پروتئین-پروتئین ژن‌های *ZmNAPL* در جدول ۴ ارائه شده است. از ابزار Functional clustering پایگاه داده DAVID 6.8 برای مطالعه هستی‌شناسی ژن‌های *ZmNAPL* و نیز گروه‌بندی آن‌ها بر اساس کارکردهای بیولوژیکی استفاده شد که به صورت معنی‌دار enrich شده بودند. آنالیز انجام گرفته نشان

جدول ۴. ژن‌های درگیر در شبکه تعامل پروتئین- پروتئین

ژن‌های شبکه	کارکرد	دمین	امتیاز
GRMZM2G152819_P01	annotation not available (112 aa)	-	0.957
nfa102	ZmNAPL5		0.919
GCN5	histone acetyl transferase GNAT/MYST 101		0.914
GRMZM2G473891_P01	annotation not available (612 aa)	RNA recognition motif	0.913
nfa101	ZmNAPL6		0.901
GRMZM2G121186_P01	ZmNAPL4		0.882
GRMZM2G360455_P01	protein phosphatase 2C isoform gamma		0.877
GRMZM2G110622_P01	annotation not available (923 aa)	Histone 2A	0.856
GRMZM2G348319_P01	Histone H2A (305 aa)		0.844
GRMZM2G096686_P01	annotation not available (236 aa)	Histone 2A	0.844
GRMZM2G152819_P01	annotation not available (112 aa)	-	0.957
nfa101	ZmNAPL6		0.926
GCN5	histone acetyl transferase GNAT/MYST 101 (515 aa)		0.919
GRMZM2G473891_P01	annotation not available (612 aa)	RNA recognition motif	0.913
mop1	RNA-dependent RNA polymerase (1127 aa)		0.896
nfa102	ZmNAPL5		0.888
GRMZM2G121186_P01	ZmNAPL4		0.881
GRMZM2G360455_P01	protein phosphatase 2C isoform gamma (907 aa)		0.877
GRMZM2G110622_P01	annotation not available (923 aa)	Histone 2A	0.856
GRMZM2G348319_P01	Histone H2A (305 aa)		0.844
GRMZM2G457415_P01	hypothetical protein LOC100384150 (975 aa)	Importin	0.986
GRMZM2G348319_P01	Histone H2A (305 aa)		0.959
GRMZM2G096686_P01	annotation not available (236 aa)	Histone 2A	0.959
AC207678.4_FGP002	annotation not available (296 aa)	Histone 2A	0.959
ZmNAPL4 (naf102)		Histone 2A	0.957
ZmNAPL5 (naf102)		Histone 2A	0.957
ZmNAPL6 (naf101)		Histone 2A	0.957
GRMZM5G845388_P01	annotation not available (116 aa)	Histone 2A	0.957
GRMZM5G835036_P01	annotation not available (116 aa)	Histone 2A	0.957
GRMZM5G801589_P01	annotation not available (288 aa)	Histone 2A	0.957
GRMZM2G416544_P01	annotation not available (257 aa)	Histone 2A	0.957
GRMZM2G181497_P01	annotation not available (116 aa)	Histone 2A	0.957
GRMZM2G158298_P01	annotation not available (132 aa)	Histone 2A	0.957

جدول ۵. اطلاعات مربوط به Functional clustering ژن‌های ZmNAPL و هستی شناسی آن‌ها

گروه	شماره GO	توصیف GO	P Value
INTERPRO	IPR002164	Nucleosome assembly protein (NAP)	3.86E-15
	GO:0006323	DNA packaging	3.42E-09
	GO:0006334	nucleosome assembly	3.42E-09
	GO:0031497	chromatin assembly	3.42E-09
	GO:0034728	nucleosome organization	3.42E-09
	GO:0065004	protein-DNA complex assembly	3.42E-09
	GO:0006333	chromatin assembly or disassembly	6.21E-09
	GO:0006325	chromatin organization	2.64E-08
	GO:0051276	chromosome organization	2.94E-08
	GO:0034622	cellular macromolecular complex assembly	4.87E-08
	GO:0034621	cellular macromolecular complex subunit organization	7.95E-08
	GO:0065003	macromolecular complex assembly	1.59E-07
	GO:0043933	macromolecular complex subunit organization	2.30E-07

آن‌ها کمتر شناخته شده است. این مطالعه اولین گزارش در مورد چپرون‌های هیستونی کلاس NAPL ذرت می‌باشد که طی آن ۶ ژن کد کننده NAPL شناسایی شد. بررسی رابطه تکاملی این ژن‌ها نشان داد که ژن‌های ZmNAPL همانند ژن‌های برنج و آرابیدوپسیس همولوژی پایینی با ژن‌های NAP انسان داشته و نزدیکترین رابطه تکاملی را با برنج دارند. همچنین ساختار ژنی و

چپرون‌های هیستونی با تأثیر بر روی دسترسی به DNA در بسیاری از فرآیندهای بیولوژیکی نظیر ترمیم DNA آسیب دیده، تقسیم سلولی و مقاومت به تنش‌های غیرزیستی نقش داشته و به‌عنوان یک کلاس مهم از تنظیم‌کننده‌های کلیدی معرفی شده‌اند که توانایی تغییر بیان بسیاری از ژن‌ها را دارند. با این حال این پروتئین‌ها در گیاهان به خوبی مطالعه نشده و تعداد دقیق، ساختار، روابط تکاملی و تنظیم بیان

سایر پروتئین‌های درگیر در هم‌گذاری نوکلئوزوم میانکنش دارند. شناسایی سایر کلاس‌های چپرون هیستونی و نیز بررسی بیان ژن‌های *ZmNAPL* در پاسخ به تنش‌ها و نیز استفاده از لاین‌های سرکوب و فرابیان شده این ژن‌ها می‌تواند در درک بیشتر کارکرد ژن‌های *ZmNAPL* مفید باشد.

موتیف‌های حفاظت شده ژن‌های *ZmNAPL* در هر کلاستر فیلوژنتیکی تا حدود زیادی حفاظت شده می‌باشد. علاوه بر این، الگوی بیان این ژن‌ها حاکی از نقش این ژن‌ها در بسیاری از فرآیندهای نمو ذرت می‌باشد. در نهایت شبکه تعامل پروتئین-پروتئین نشان داد که این پروتئین‌ها با یکدیگر و نیز

## REFERENCES

- Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, De Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E (2012) ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* 40: W957-W603.
- Attia M, Förster A, Rachez C, Freemont P, Avner P, Rogner UC (2011) Interaction between nucleosome assembly protein 1-like family members. *J. Mol. Biol.* 407(5): 647-660.
- Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 34: W369-W373.
- Braun P, Aubourg S, Van Leene J, De Jaeger G, Lurin C (2013) Plant protein interactomes. *Annu. Rev. Plant Biol.* 64: 161-187.
- Burgess RJ, Zhang Z (2013) Histone chaperones in nucleosome assembly and human disease. *Nat. Struct. Mol. Biol.* 20(1): 14-22.
- Dash S, Van Hemert J, Hong L, Wise RP, Dickerson JA (2012) PLEXdb: gene expression resources for plants and plant pathogens. *Nucleic Acids Res.* 40: D1194-D1201.
- De Koning L, Corpet A, Haber JE, Almouzni, G (2007) Histone chaperones: an escort network regulating histone traffic. *Nat. Struct. Mol. Biol.* 14(11): 997-1007.
- Dennehey BK, Tyler J (2014) Histone chaperones in the assembly and disassembly of chromatin. In: Workman JA, Abmayr SM (Eds) *Fundamentals of chromatin*, Springer, New York, pp 29-67.
- Eitoku M, Sato L, Senda T, Horikoshi M (2008) Histone chaperones: 30 years from isolation to elucidation of the mechanisms of nucleosome assembly and disassembly. *Cell. Mol. Life Sci.* 65(3): 414-444.
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution.* 39(4): 783-791.
- Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J (2013) Pfam: the protein families database. *Nucleic Acids Res.* 42: D222-D230.
- Frey FP, Urbany C, Hüttel B, Richard Reinhardt R, Stich Email B (2015) Genome-wide expression profiling and phenotypic evaluation of European maize inbreds at seedling stage in response to heat stress. *BMC Genomics.* 16:123.
- Gamble MJ, Erdjument-Bromage H, Tempst P, Freedman LP, Fisher RP (2005) The histone chaperone TAF-I/SET/INHAT is required for transcription in vitro of chromatin templates. *Mol. Cell. Biol.* 25(2): 797-807.
- Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40: D1178-D1186.
- Guo J, Wu J, Ji Q, Wang C, Luo L, Yuan Y, Wang Y, Wang J (2008) Genome-wide analysis of heat shock

- transcription factor families in rice and Arabidopsis. *J. Genet. Genomics*. 35(2): 105-118.
- Hogeweg P (2011) The roots of bioinformatics in theoretical biology. *PLoS Comput. Biol.* 7: e1002021.
- Hondele M, Ladurner AG (2011) The chaperone-histone partnership: for the greater good of histone traffic and chromatin plasticity. *Curr. Opin. Struct. Biol.* 21(6): 698-708.
- Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G (2014) GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*. 31(8):1296-1297.
- Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4(1): 44-57.
- Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, Schwartz DC, Tanaka T, Wu J, Zhou S, Childs KL, Davidson RM, Lin H, Quesada-Ocampo L, Vaillancourt B, Sakai H, Lee S S, Kim J, Numa H, Itoh T, Buell CR, Matsumoto T (2013) Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice*. 6:4.
- Kellogg DR, Murray AW (1995) NAP1 acts with Clb1 to perform mitotic functions and to suppress polar bud growth in budding yeast. *J. Cell. Biol.* 130(3): 675-685.
- Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33(7):1870-1874.
- Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, Karthikeyan AS, Lee CH, Nelson WD, Ploetz L, Singh S, Wensel A, Huala E (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* 40: D1202-D1210.
- Larkin MA, Blackshields G, Brown N, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*. 23(21): 2947-2948.
- Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res.* 40: D302-D305.
- Liao Y, Liu S, Jiang Y, Hu C, Zhang X, Cao X, Xu Z, Gao X, Li L, Zhu J (2017) Genome-wide analysis and environmental response profiling of dirigent family genes in rice (*Oryza sativa*). *Genes and Genomics*. 39(1): 47-62.
- Liu Z, Zhu Y, Gao J, Yu F, Dong A, Shen WH (2009) Molecular and reverse genetic characterization of NUCLEOSOME ASSEMBLY PROTEIN1 (NAP1) genes unravels their function in transcription and nucleotide excision repair in Arabidopsis thaliana. *Plant J.* 59(1): 27-38.
- Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, eWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* 39: D225-D229.
- Marheineke K, Krude T (1998) Nucleosome assembly activity and intracellular localization of human CAF-1 changes during the cell division cycle. *J. Biol. Chem.* 273(24): 15279-15286.
- McGinnis S, Madden TL (2004) BLAST: at the core of a powerful and diverse

- set of sequence analysis tools. *Nucleic Acids Res.* 32: W20-W25.
- Mosammaparast N, Ewart CS, Pemberton LF (2002) A role for nucleosome assembly protein 1 in the nuclear transport of histones H2A and H2B. *EMBO J.* 21(23): 6527-6538.
- Park YJ, Luger K (2006) The structure of nucleosome assembly protein 1. *Proc. Natl. Acad. Sci. U. S. A.* 103(5): 1248-1253.
- Sekhon RS, Lin H, Childs KL, Hansey CN, Buell CR, de Leon N, Kaeppler SM (2011) Genome-wide atlas of transcription during maize development. *Plant J.* 66(4): 553-563.
- Sharp PA (1981). Speculations on RNA splicing (minireview). *Cell.* 23(3): 643-646.
- Shimizu Y, Akashi T, Okuda A, Kikuchi A, Fukui K (2000) NBP1 (Nap1 binding protein 1), an essential gene for G2/M transition of *Saccharomyces cerevisiae*, encodes a protein of distinct sub-nuclear localization. *Gene.* 246(1-2): 395-404.
- Singh AK, Kumar R, Tripathi AK, Gupta BK, Pareek A, Singla-Pareek SL (2015) Genome-wide investigation and expression analysis of Sodium/Calcium exchanger gene family in rice and *Arabidopsis*. *Rice.* 8: 21.
- Singh AK, Sharma V, Pal AK, Acharya V, Ahuja PS (2013) Genome-wide organization and expression profiling of the NAC transcription factor family in potato (*Solanum tuberosum* L.). *DNA Res.* 20(4): 403-423.
- Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P (2016) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 45(D1): D362-D368.
- Tripathi AK, Pareek A, Singla-Pareek SL (2016) A NAP-Family histone chaperone functions in abiotic stress response and adaptation. *Plant Physiol.* 171(4): 2854-2868.
- Tripathi AK, Singh K, Pareek A, Singla-Pareek SL (2015) Histone chaperones in *Arabidopsis* and rice: genome-wide identification, phylogeny, architecture and transcriptional regulation. *BMC Plant Biol.* 15: 42.
- Valieva M, Feofanov A, Studitsky V (2016) Histone chaperones: Variety and functions. *Moscow Univ. Biol.Sci. Bull.* 71(3): 165-169.
- Wei KF, Chen J, Chen YF, Wu LJ, Xie DX (2012) Molecular phylogenetic and expression analysis of the complete WRKY transcription factor family in maize. *DNA Res.* 19(2): 153-64.
- Zhang Y, Gao M, Singer SD, Fei Z, Wang H, Wang X (2012) Genome-wide identification and analysis of the TIFY gene family in grape. *PLoS One.* 7(9): e44465.
- Zhu P, Gu H, Jiao Y, Huang D, Chen M (2011) Computational identification of protein-protein interactions in rice based on the predicted rice interactome network. *Genomics, proteomics and Bioinformatics.* 9(4-5): 128-137.
- Yu CS, Chen YC, Lu CH, Hwang JK (2006) Prediction of protein subcellular localization. *Proteins.* 64(3): 643-651.