Санкт-Петербургский Государственный Университет

Биологический факультет

Кафедра генетики и биотехнологии

Дворникова Кристина Алексеевна

Полиморфизм генов *CLE*, регулирующих развитие симбиотических клубеньков у люцерны *Medicago truncatula*

Выпускная квалификационная работа магистра

по направлению подготовки «Биология»

основная образовательная программа магистратуры «Молекулярная биология и агробиотехнология растений»

Работа выполнена на кафедре генетики и биотехнологии СПбГУ в лаборатории генной и

клеточной инженерии растений

(зав. лаб. - д.б.н., проф. Лутова Л.А.)

Научный руководитель:

к.б.н. Лебедева М.А., ст.н.с. кафедры генетики и биотехнологии СПбГУ в лаборатории генной и клеточной инженерии растений

Рецензент:

к.б.н. Романюк Д.С., н.с. лаборатория генетики растительно-микробных взаимодействий, Отдел биотехнологии ФГБНУ ВНИИ сельскохозяйственной микробиологии

Санкт-Петербург

2021

**ОГЛАВЛЕНИЕ**

|  |  |
| --- | --- |
| СПИСОК СОКРАЩЕНИЙ  ВВЕДЕНИЕ | 3  4 |
| 1. ОБЗОР ЛИТЕРАТУРЫ    1. Пептиды CLE и их роль в развитии растений    2. Пептиды CLE в регуляции развития клубеньков    3. Полиморфизм и естественный отбор    4. Методы оценки полиморфизма генов у люцерны | 6  6  8  11  17 |
| 1. МАТЕРИАЛЫ И МЕТОДЫ | 20 |
| * 1. Растительный материал | 20 |
| * 1. Выделение растительной ДНК | 20 |
| * 1. Амплификация фрагментов генов *CLE*   2. Трансформация растений с помощью *A*. *rhizogenes*   3. Выделение РНК и обратная транскрипция   4. Проведение ПЦР в реальном времени   5. Биоинформатический анализ | 20  21  22  22  23 |
| 1. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ | 25 |
| * 1. Филогенетический анализ генов *CLE* | 25 |
| * 1. Распределение генов *MtCLE* по хромосомам | 27 |
| * 1. Оценка экспрессии генов *MtCLE* с помощью доступных транскриптомных данных | 29 |
| * 1. Оценка полиморфизма генов *MtCLE* люцерны | 33 |
| * 1. Определение последовательностей фрагментов гена *MtCLE34* с помощью секвенирования   2. Определение влияния отбора в участках генов *MtCLE*   3. Поиск ассоциаций полиморфизма генов *MtCLE* с количеством образующихся клубеньков   4. Изучение функциональной значимости гена *MtCLE34* в регуляции развития симбиотических клубеньков у *M. truncatula* | 34  35  45  47 |
| ЗАКЛЮЧЕНИЕ  ВЫВОДЫ  СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ  ПРИЛОЖЕНИЯ  БЛАГОДАРНОСТИ | 51  52  53  61  63 |

**СПИСОК СОКРАЩЕНИЙ**

AON (Autoregulation Of Nodulation) - система авторегуляции клубенькообразования

ТФ – транскрипционный фактор

SNP (Single Nucleotide Polymorphism) - однонуклеотидный полиморфизм

LRR-RK (Leucine-Rich Repeat Receptor Kinase) – рецепторная киназа с лейцин-богатыми повторами

RLK (Receptor-Like Kinases) - рецептор-подобные киназы

GWAS (Genome-wide association study) - полногеномный поиск ассоциаций

QTL (Quantitative Trait Loci) - локус количественных признаков

SDA (Strand Displacement Amplification) - одиночные аллельные SNP

ThetaPerBP – «Тэта» оценка Уоттерсона

PiPerBP – «Пи» оценка Уоттерсона

ЦТАБ - цетилтриметиламмонийбромид

ПЦР - Полимеразная цепная реакция

ТАЕ - трис-ацетатный буфер

GFP (Green Fluorescent Protein) - зелёный флуоресцентный белок

ORF (Open Reading Frame) – открытая рамка считывания

CDS (Coding Sequence) - кодирующая последовательность

IGR (Intergenic Region) – межгенная область

nsy (nonsynonymous) - несинонимичные замены

syn (synonymous) – синонимичные замены

GUS (β-glucuronidase) - бета-глюкуронидаза

**ВВЕДЕНИЕ**

Одними из наиболее изученных пептидных фитогормонов являются пептиды CLE (CLAVATA3/EMBRYO SURROUNDING REGION-RELATED). Пептиды CLE имеют размер около 12-14 аминокислот, и образуются из более протяженных белков-предшественников, которые содержат высоко консервативный С-концевой СLE-домен, основная представляющий собой функциональную часть пептидов CLE, а также предполагаемый сигнал секреции на N-конце. Пептиды CLE участвуют в регуляции различных аспектов развития растений, в том числе ̶ в системном контроле развития симбиотических клубеньков. Клубеньки формируются на корнях бобовых растений при симбиозе с почвенными азотфиксирующими бактериями ризобиями. Наряду с механизмами локального контроля, активируемых в результате рецепции сигнальных молекул ризобий, Nod-факторов, в процесс формирования клубеньков у бобовых растений также вовлечены и системные механизмы - система авторегуляции клубенькообразования (Autoregulation Of Nodulation, AON). Ключевым компонентом AON является рецепторная CLAVATA 1 (CLV1) - подобная киназа, демонстрирующая высокий процент сходства по последовательности с киназой CLV1, ответственной за поддержание пула стволовых клеток в апикальной меристеме побега. Мутанты по генам, кодирующим CLV1-подобные рецепторы получены у ряда бобовых и все они образуют избыточное число клубеньков при симбиозе с ризобиями (суперклубенькообразующий фенотип). CLV1-подобные киназы, как было показано с помощью прививок, функционируют в побеге, и их лигандами являются регуляторные пептиды CLE, которые синтезируются в корнях и поступают по проводящей системе из корней – в побег. Однако, в работе системы авторегуляции клубенькообразования у бобовых растений еще многое остается неизученным. В частности, до конца неясно, какое число пептидов CLE и их рецепторов в ней задействовано.

В настоящей работе проводилось изучение полиморфизма *Medicago truncatula* по генам, кодирующим пептиды CLE, для которых показано или предполагается участие в системе авторегуляции клубенькообразования. В качестве объектов исследования были использованы 2 линии бобового растения *M. truncatula* – A17 и R108. В качестве генов для исследования были выбраны гены *MtCLE12* и *MtCLE13*, для которых ранее было показано участие в AON (Mortier et al., 2010), а также два родственных им гена, *MtCLE34* и *MtCLE35*. Как было недавно показано в нашей лаборатории, ген *MtCLE35* также задействован в AON и сверхэкспрессия этого гена подавляет развитие симбиотических клубеньков. Функция *MtCLE34* оставалась неизученной. В открытой рамке считывания гена *MtCLE34* у линии люцерны A17 содержится стоп-кодон, что должно приводить к синтезу укороченного нефункционального белка. Однако, у другой лабораторной линии *M. truncatula*, R108, согласно последовательностям, представленным в базе данных, стоп-кодон отсутствует, таким образом, у линии R108 ген *MtCLE34* может кодировать функциональный продукт. Изучение генетического полиморфизма генов *MtCLE* у бобового растения *M. truncatula* представляется весьма важным как для фундаментальной науки, так и, в будущем, для прикладных аспектов – селекционной работы и современной практики растениеводства.

Таким образом, целью работы является изучение полиморфизма генов *CLE*, регулирующих развитие клубеньков у *M. truncatula*.

В работе были поставлены следующие задачи:

1. Характеристика генов *CLE* *M. truncatula*: филогенетический анализ, оценка их расположения на хромосомах люцерны, оценка уровней экспрессии генов *CLE* с помощью доступных баз данных, поиск консервативных элементов в регуляторных последовательностях генов *CLE*;
2. Поиск вариабельных позиций в последовательностях генов *CLE* у линий *M. truncatula*, доступных в базе данных HapMap и оценка давления отбора на гены *CLE* с помощью методов молекулярной эволюции;
3. Изучение функции гена *MtCLE34* в регуляции развития симбиотических клубеньков у *M. truncatula*.
4. **ОБЗОР ЛИТЕРАТУРЫ**
   1. *Пептиды CLE и их роль в развитии растений*

Пептиды CLE являются семейством пептидных фитогормонов, регулирующих развитие растений. Пептиды CLE широко распространены у всех наземных растений, включая однодольные, двудольные, мхи и водоросли (Ганчева и др., 2019). В геноме *Arabidopsis thaliana* выявлено 32 гена (Strabala et al., 2006; Miyawaki et al., 2013), кодирующих пептиды этого класса, в геноме *Oryza sativa* – 47 *CLE* генов (Oelkers et al., 2008), в геноме *Medicago truncatula* – 52 *CLE* гена (Hastwell et al., 2017). Также 104 гена *CLE* было обнаружено у *Tricticum aestivum L.* (Li et al., 2019). Существует также группа CLE-подобных (CLEL, CLE-like) пептидов, которые имеют схожий функциональный домен, но проявляют иную биологическую активность (Hastwell et al., 2015).

Пептиды CLE играют важную роль в различных процессах развития растений и/или в опосредованных реакциях на внешние раздражители (Fletcher, 2020) (рис. 1).

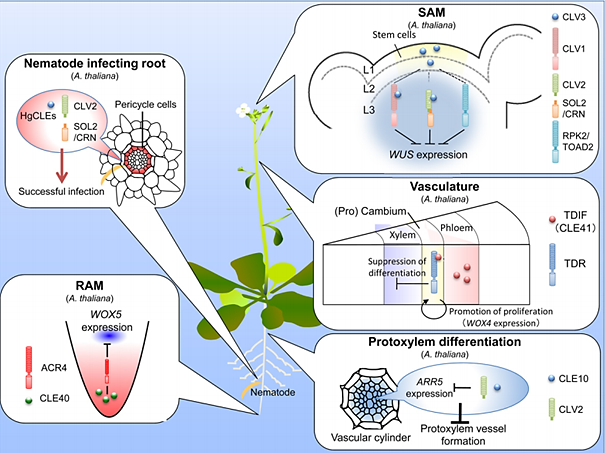


Рис. 1. Функции генов *CLE* в развитии растений (Betsuyaku et al., 2011).

В частности, пептиды CLE и их рецепторы задействованы в сигнальных путях, контролирующих поддержание пула стволовых клеток в апикальной меристеме побега и корня. Баланс стволовых клеток в апикальных меристемах контролируется с участием гомеобокс-содержащих генов *WUSCHEL (WUS)* и *WOX5,* экспрессирующихся в так называемых «организующих центрах» меристем, располагающихся рядом с пулами стволовых клеток (Sarkar et al., 2007).В меристеме побега, транскрипционный фактор (ТФ) WUS стимулирует экспрессию гена *CLAVATA3 (CLV3)* в стволовых клетках. Функциональным продуктом *CLV3* является додекапептид CLV3 - первый выявленный представитель семейства пептидов CLE. Пептид CLV3 активирует сигнальный каскад, приводящий к подавлению экспрессии гена *WUS* в организующем центре меристемы побега (Schoof et al., 2000)*.* При этом ТФ WUS активирует экспрессию гена *CLV3*, напрямую связываясь с его промотором (Yadav et al., 2011). В результате WUS-зависимой активации экспрессии *CLV3*, образуется петля негативной обратной связи в регуляции размера пулов стволовых клеток в меристеме побега (Лутова и др., 2015). В корневой апикальной меристеме пептид CLE40 выполняет функцию, схожую с функцией CLV3: он ограничивает зону экспрессии гена *WOX5 –* паралога гена *WUS* в меристеме корня (Stahl et al., 2009). CLV3 взаимодействует с целым рядом рецепторных комплексов, включающих белки CLV1, CLV2, CORYNE (CRN) и RECEPTOR-LIKE PROTEIN KINASE2 (RPK2). CLV1 и RPK2 - рецептор-подобные киназы с лейцин-богатыми повторами (LRR-RKs), а CLV2 - рецептор-подобный белок LRR (LRR-RP) без киназного домена, функционирующий в паре с белок CRN в передаче сигнала CLV3 (Kinoshita et al., 2010). Известно, что пептиды CLE являются лигандами LRR-подобных рецепторных киназ (Endo et al., 2014). В свою очередь, рецепторные белки LRR-RLK (LRR Receptor-Like Kinases) являются самой большой группой суперсемейства рецептор-подобных киназ (RLK), выполняющие разнообразные функции у растений, –в том числе - в процессах их развития (Liu, 2017). Регуляторные пептиды CLE и гены *WOX*, мишени активируемого ими сигнального каскада, работают в меристемах различных типов: в апикальной меристеме побега пептид CLE, CLV3, регулирует экспрессию гена *WUS*, в корне пептид CLE40 участвует в регуляции гена *WOX5*, в камбии пептид CLE41/44 (также известный как Tracheary Elements Differentiation Inhibitory Factor (TDIF)), контролирует экспрессию гена *WOX4*. Отмечается, что у *A. thaliana* пептид TDIF, кодируемый генами *CLE41/44*, участвует в подавлении дифференцировки прокамбальных клеток в ксилеме и способствует их пролиферации путем связывания с рецепторной киназой TDIF receptor/phloem intercalated with xylem (TDR/PXY), контролирующей экспрессию гена *WOX4* (Zhang et al., 2016). Потеря функции TDR/PXY приводит к развитию TDIF-нечувствительного фенотипа с редуцированным камбием и аномальным строением проводящих пучков (Hirakawa et al., 2010), а рецепторы РХС (PXY/TDR-CORRELATED) у *A. thaliana* участвуют в связывании пептидов TDIF и играют важную роль в дифференцировке сосудов (Wang et al., 2013). Кроме того, в рецепции пептидов CLE у *A. thaliana* также задействованы белки BARELY ANY MERISTEM 1, 2, 3 (BAM 1, 2, 3), CLV2, SUPPRESSOR OF LLP2/CORYNE (SOL2/CRN) и RPK2/TOAD2 наряду с другими комплексами, в частности, параллельно с CLV1, который вместе с BAM образует монофилетическую группу генов *A. thaliana*, кодирующих соответствующие рецепторные киназы (Somssich et al., 2016; Whitewoods, 2020).

Также пептиды CLE участвуют в передаче сигналов в развитии листьев. Установлено, что у *A. thaliana* регуляция транскрипции *CLE5* и *CLE6* осуществляется с помощью ТФ WOX1 и PRESSED FLOWER (PRS) у основания развивающихся листьев до получения окончательной формы листа, а потеря этой функции приводит к появлению более широких черешков (DiGennaro et al., 2018). Однако механизм того, как CLE опосредуют этот эффект пока точно неизвестен. Некоторые пептиды CLE вовлечены в регуляцию эмбриогенеза. В частности, одним из регуляторов ранних этапов эмбриогенеза и развития семян является пептид CLE8, мишень которого - ген *WOX8* (Yamaguchi et al., 2016). Таким образом, пептиды CLE играют важную роль в развитии растений, в частности, контролируют различные аспекты развития растений (поддержание пула стволовых клеток в апикальной меристеме побега и корня, передача сигналов в развитии листьев, регуляция эмбриогенеза), а также участвуют в системном контроле развития симбиотических клубеньков (см. раздел 1.2).

Помимо растений, способностью к синтезу пептиды CLE обладают некоторые паразитические нематоды, индуцирующие корневые галлы, а также грибы арбускулярной микоризы (Miyawaki et al., 2013; Chen et al., 2015). В частности, обработка пептидом RiCLE1, синтезированым Smartox Biotechnology (Saint-Egrève, Франция), проростков *M. truncatula* стимулировала развитие арбускулярной микоризы(Le Marquer et al., 2018). Также было показано, что сверхэкспрессия гена *MtCLE53* у люцерны негативно регулирует колонизацию корня грибами арбускулярной микоризы (Karlo et al., 2020). Однако до сих пор неясно, какие сигнальные компоненты используются CLE для обеспечения успешного симбиоза растений и грибов. Паразитические нематоды имеют собственные копии CLV3-подобных и TDIF-подобных генов, опосредующих образование синцития и инфекции. Установлено, что протеолитическая обработка инфицированным *A. thaliana* пептида GrCLE1 нематод *Globodera rostochiensis,* паразитирующих на корнях картофеля и томатов, изменяет развитие корня, связываясь с рецепторами CLV2, BAM1 и BAM2 (Guo et al., 2011; Guo et al., 2017).

*1.2. Пептиды CLE* *в регуляции развития клубеньков*

Пептиды CLE играют важную роль в развитии клубеньков. Наряду с механизмами локального контроля, активируемых при взаимодействии с ризобиями, в процесс формирования клубеньков у бобовых растений также вовлечены и системные механизмы - система авторегуляции клубенькообразования (Autoregulation Of Nodulation, AON). Предполагается, что система AON связана с поддержанием баланса между притоком азота и оттоком углеродных соединений (Ferguson et al., 2018). C помощью AON поддерживается оптимальное количество корневых клубеньков (Suzaki et al., 2019). У мутантов с нарушением AON образование избыточного количества клубеньков может приводить к замедлению роста растений (Nishida et al., 2018).

Известны пептиды CLE, являющиеся компонентами системы AON: они образуются в корнях при развитии симбиотических клубеньков, и перемещаются в побег, где, взаимодействуя со своими рецепторами, запускают каскад реакций, приводящий к подавлению развития новых клубеньков по механизму обратной связи. В частности, пептиды MtCLE12 и MtCLE13 у *M. truncatula*, CLE ROOT SIGNAL1 и 2 (CLE-RS1 и CLE-RS2) у *Lotus japonicus* подавляют избыточное образование клубеньков (Mortier et al., 2010; Yamaguchi et al., 2016). Рецепторами пептидов CLE в побеге являются рецепторные CLV1-подобные киназы MtSUNN у *M. truncatula*, LjHAR1 у *Lotus japonicus*, PsSYM29 у *Pisum sativum* и GmNARK у *Glycine max* (Mortier et al., 2011). Мутации в генах, кодирующих такие CLV1-подобные рецепторные киназы, приводят к развитию суперклубенькообразующего фенотипа у целого ряда бобовых растений.

Активация экспрессии генов *CLE*, вовлеченных в AON, происходит с участием как сигнального каскада, индуцируемого ризобиями (Wang et al., 2019), так и с участием нитрата (Okamoto et al., 2009). Активатором экспрессии *CLE* при взаимодействии с ризобиями является ТФ NODULE INCEPTION (NIN) - ключевой регулятор развития клубеньков (Laffont et al., 2020).

Наряду с пептидами CLE, пептиды другой группы – CEP (C-terminally encoded peptides), также вовлечены в авторегуляцию клубенькообразования, и способны транспортироваться из корней – в побег, где они, связываясь со своими рецептором, кодируемым у *M. truncatula* геном *MtCRA2* (*COMPACT ROOT ARCHITECTURE 2*), запускают реакции, стимулирующие образование симбиотических клубеньков (рис. 2). У *M. truncatula* экспрессия гена *MtCEP7* активируется также с участием ТФ NIN. Показано, что рецепторы пептидов CEP и CLE по-разному влияют на уровень микроРНК miR2111. Так, активация рецептора MtCRA2 с участием пептидов CEP приводит к накоплению miR2111 в побеге, тогда как активация рецептора CLE, MtSUNN, напротив, приводит к снижению уровня miR2111 (Gautrat et al., 2020). miR2111 является мобильной молекулой, транспортирующейся из побега – в корень и регулирующей экспрессию генов, контролирующих развитие симбиотических клубеньков (Gautrat et al., 2020). Мишенью miR2111 является мРНК гена *TOO MUCH LOVE (TML)* – негативного регулятора клубенькообразования, работающего в корне (Tsikou et al., 2018). Таким образом, miR2111 положительно регулирует симбиотическое клубенькообразование.

Наряду с ТФ NIN, активирующего экспрессию генов *CLE* при взаимодействии с ризобиями, экспрессия ряда генов *CLE* также регулируется с участием нитрата. Например, у *L. japonicus* экспрессия двух из трех генов *CLE*, индуцируемых при симбиозе с ризобиями, LjCLE-RS2 и LjCLE-RS3, индуцируется в ответ на обработку нитратом (Okamoto et al., 2009). У *Glycine max* были выявлены нитрат-индуцируемые CLE, относящиеся к группе NIC (NITRATE INDUCED CLE) - гены *GmNIC1* и *GmNIC2,* активирующиеся в ответ на присутствие нитрата в почве и ингибирующие клубенькообразование, а также индуцируемые ризобиями пептиды группы RIC - GmRIC1 и GmRIC2, задействованные в системном контроле клубенькообразования при участии CLV1-подобной киназы GmNARK (Reid et al., 2011). Согласно результатам исследований, пептиды GmNIC1 и GmNIC2 подавляют клубенькообразование локально через CLV1-подобную киназу GmNARK, функционирующую в корнях. Сверхэкспрессия генов *GmNIC1* и *GmNIC2* в трансгенных корнях, полученных в результате трансформации с помощью *Agrobacterium rhizogenes,* не приводила к подавлению клубенькообразования на нетрансгенных корнях у композитных растений (несущих как трансгенные корни, так и нетрансгенные)(Lim et al., 2014). Это, в свою очередь, может свидетельствовать о локальном подавлении клубенькообразования пептидами GmNIC1 и GmNIC2 при участии CLV1-подобной киназы GmNARK. Однако на сегодняшний день еще точно неизвестны механизмы, лежащие в основе работы корневой рецепции CLE с участием GmNARK.

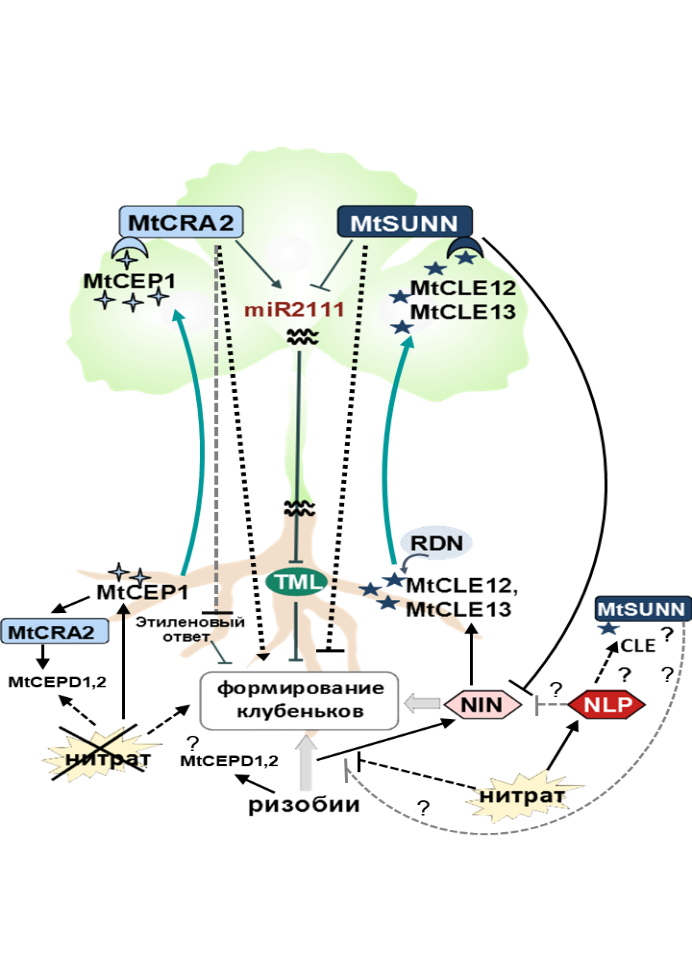


Рис. 2. Системный контроль развития клубеньков при участии пептидов CLE и CEP (Лебедева и др., 2020).

У *A. thaliana* также известны пептиды CLE, опосредующие реакции корней доступность питательных веществ. Так, экспрессия генов *CLE1, CLE3, CLE4* и *CLE7* индуцируется в корнях в условиях дефицита азота, а избыточная экспрессия *CLE3* в таких условиях ингибирует развитие боковых корней CLV1-зависимым образом (Araya et al., 2014). При этом CLV1 может ингибировать развитие бокового корня при низком содержании сульфатов в среде (Dong et al., 2019).

Можно заключить, что пептиды CLE и их рецепторы имеют важное значение в развитии растений, в частности - в развитии апикальной меристемы побегов и корней, меристем камбия, листьев, эмбриогенезе, в формировании устьиц, ветвлении корней и побегов, развитии пыльников, инфицировании паразитическими нематодами, а также в арбускулярном микоризном симбиозе и клубенькообразовании.

*1.3. Полиморфизм и естественный отбор*

В результате точечных мутаций может возникнуть явление, получившее название однонуклеотидный полиморфизм (Single Nucleotide Polymorphism, SNP). SNP представляет собой отличия последовательности ДНК размером в один нуклеотид в геноме между гомологичными участками гомологичных хромосом (рис. 3). На сегодняшний день SNP является одним из наиболее активно изучаемым видом полиморфизма. Частота возникновения однонуклеотидного полиморфизма достигает >1% в нормальной популяции (в отличие от других мутаций, которые происходят с частотой 1% или менее).

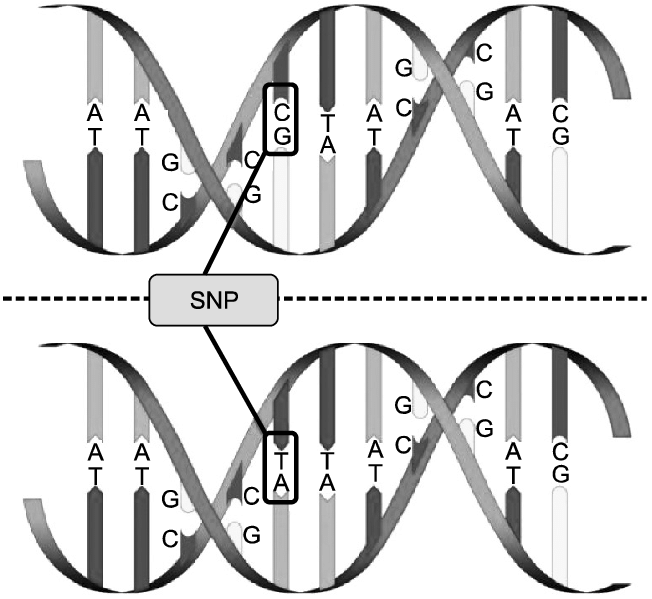


Рис. 3. Схема однонуклеотидного полиморфизма (SNP) (Doo et al., 2015).

SNP встречаются в кодирующих и в некодирующих последовательностях. Выделяют два основных типа замен: синонимичные и несинонимичные. Некоторые SNP в кодирующих участках могут быть не связаны с каким-либо изменением в аминокислотной последовательности и, таким образом, не иметь функциональных последствий, такие замены, оставляющие аминокислотную последовательность белка без изменения, называют синонимичные. Напротив, несинонимичные замены приводят к замене аминокислоты. SNP также подразделяют на три типа: сайленс, нонсенс и миссенс. При сайленс-мутациях, как правило, не имеющих функциональных последствий, кодон кодирует ту же аминокислоту из-за вырожденности генетического кода. Нонсенс-мутации вызывают образование стоп-кодонов, а при миссенс мутации кодон начинает кодировать другую аминокислоту. При этом, возможен вариант транзиции, т.е. мутации, обусловленной заменой одного пуринового основания на другое, либо аналогичной заменой пиримидиновых оснований, а также вариант трансверсии – мутации, при которой происходит замена пуринового основания на пиримидиновое, или наоборот.

Поиск и выявление полиморфизма может дать представление об эволюционных путях, ведущих к современному состоянию генов бобовых у родственных видов. Поэтому важно рассматривать полиморфизм в контексте естественного отбора. Известно, что естественный отбор представляет собой эволюционный процесс, приводящий к сохранению и увеличению в популяции особей с полезными признаками и уменьшению количества особей с менее полезными признаками (Futuyma et al., 2017). При этом естественный отбор, поддерживаемый генетические вариации, представляет основной интерес для изучения происхождения и сохранения генетического полиморфизма, который является результатом существования в популяции двух или более вариантов (аллелей) отдельных генов (Harrisson et al., 2014). В частности, поддержка естественным отбором аллельного полиморфизма, называется балансирующим полиморфизмом. В этом случае частоты селективно отличающихся аллелей поддерживаются в равновесии (Delph et al., 2013).

Согласно классификации, основанной на характере влияния форм отбора на изменчивость признака в популяции, выделяют три основные формы естественного отбора: движущий, стабилизирующий и дизруптивный (рис. 4).

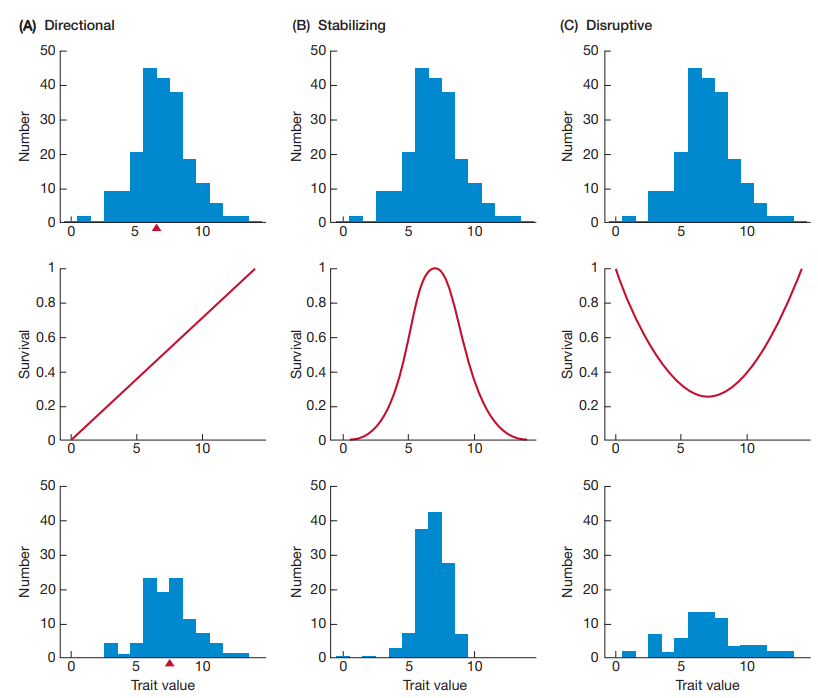


Рис. 4. Формы естественного отбора (Futuyma et al., 2017).

Движущий отбор способствует изменению признака в сторону большего или меньшего текущего среднего значения. При этом движущий отбор действует при направленном изменении условий окружающей среды и является основной движущей силой эволюционных изменений (Kingsolver et al., 2011). Стабилизирующий отбор направлен на сохранение средней величины значений признака, сужая крайние отклонения от средней нормы. В результате стабилизирующего отбора происходит уменьшение количества экстремальных признаков и увеличение средних (промежуточных) вариантов (Lieberman et al., 1996). Дизруптивный (разрывающий) отбор направлен в пользу сохранения двух крайних признаков, образованных из одной исходной. При этом окружающая среда является благоприятной только для крайних признаков, а для средних (промежуточных) является неблагоприятной. Важно, что дизруптивный отбор способствует появлению и поддержанию полиморфизма (Rueffler et al., 2006).

Отдельно отмечается балансирующий (дестабилизирующий) отбор, который поддерживает генетическую изменчивость в популяции (рис. 5). Предполагается, что балансирующий отбор может быть начальным этапом дизруптивного отбора. Одним из механизмов балансирующего отбора является избирательное преимущество гетерозигот с двумя различными аллелями перед гомозиготами с одним аллелем (Villanea et al., 2015).

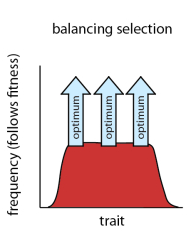


Рис. 5. Схема балансирующего (дестабилизирующего) отбора (Loewe, 2008).

Существует классификация естественного отбора по направленности действия: положительный (позитивный) и отрицательный (очищающий). Положительныйотбор направлен на распространение и фиксацию возникающего полиморфизма. Такой полиморфизм рассматривается как полезный. В ходе положительного отбора возможно увеличение частоты и фиксация полезного аллеля – такой процесс называется «выборочная зачистка» («selective sweep») (рис. 6). В свою очередь, под этим процессом также понимается ассоциированное уменьшение или устранение вариации последовательности ДНК на хромосоме. Выборочная зачистка возможна в случае увеличения частоты редкого аллеля в результате естественного отбора. Однако при увеличении распространенности этого редкого аллеля параллельно происходит увеличение распространенности других генетических вариантов, находящихся вблизи этого редкого аллеля. Такой процесс называется «эффект автостопа» («hitchhiking effect»). Следовательно выборочная зачистка может приводить к уменьшению генетического разнообразия не только по аллелям данного гена, но и в содержащей его области хромосомы.

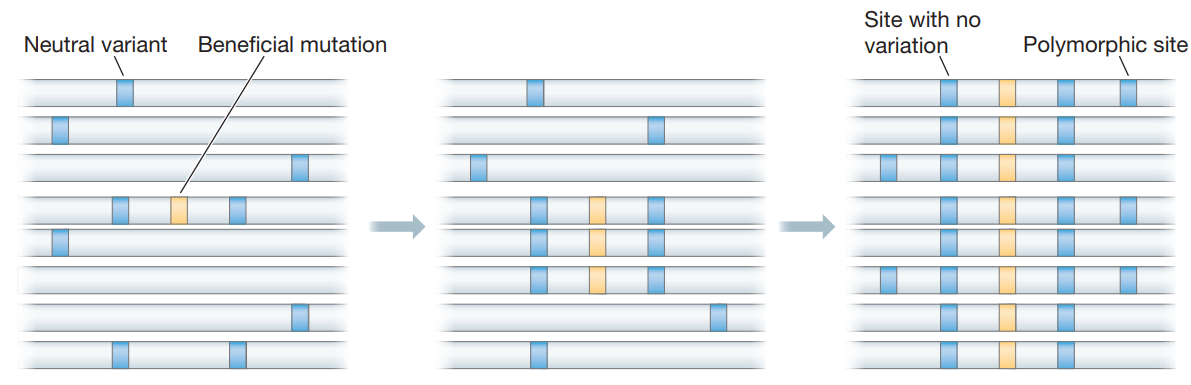


Рис. 6. Схема «выборочной зачистки» («selective sweep») в популяции (Futuyma et al., 2017).

Примером «выборочной зачистки» («selective sweep») у растений является одомашнивание теосинте (*Zea mays ssp. parviglumis*) и кукурузы (*Zea mays mays)* (рис. 7)*.* Произошла мутация в области хромосомы, которая регулирует экспрессию гена *tb1*, контролирующего ветвление побегов. По мере распространения мутации гетерозиготность рядом с хромосомой резко снижалась. Также существуют доказательства, указывающие на мутацию в некодирующей области около 60 т.п.н. выше от *tb1* (Studer et al., 2011).

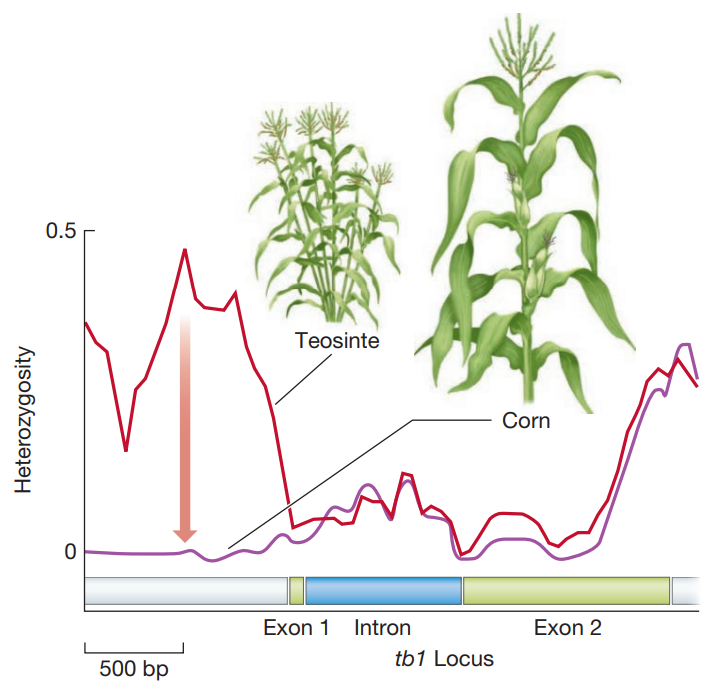


Рис. 7. Пример «выборочной зачистки» («selective sweep») у *Zea mays ssp. parviglumis* и *Zea mays mays* (Futuyma et al., 2017).

Отрицательный (очищающий) отбор направлен на удаление генетической изменчивости (рис. 8). При этом отбор в большей степени сказывается на функционально значимых геномных последовательностях и, как следствие, - наиболее значимые области генома являются высококонсервативными. Также отрицательный отбор является самым распространённым типом естественного отбора и эффективной формой движущего отбора.

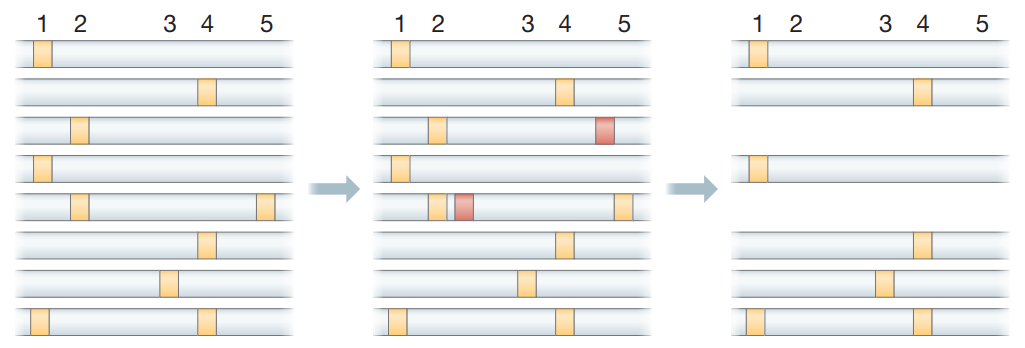


Рис. 8. Схема отрицательного (очищающего) отбора в популяции. Желтым цветом обозначен нейтральный аллель, красным – вредная мутация (Futuyma et al., 2017).

Важным фактором эволюции является изменение численности популяции. Примером такого изменения может быть прохождение популяции сквозь «бутылочное горлышко» («bottleneck») (рис. 9). Причиной возникновения эффекта «бутылочного горлышка» может быть ряд экологических событий (природные катастрофы, болезни) и деятельность человека (охота, геноцид), приводящих к гибели организмов.

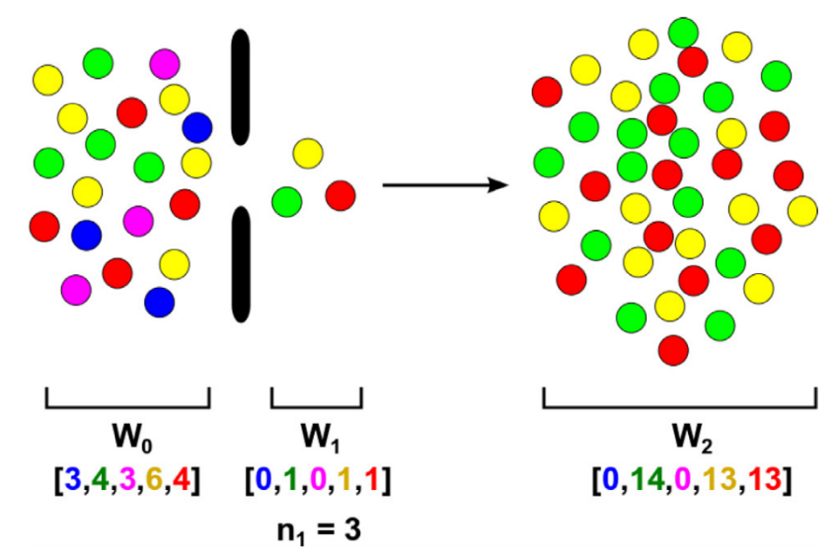


Рис. 9. Схема «бутылочного горлышка» («bottleneck») в популяции и его склонности к стохастической изменчивости с точки зрения размера (n1) и состава (w) генов (Dybowski et al., 2017).

Результатом такого прохождения является резкое уменьшение генетического разнообразия популяции, что может вызвать накопление вредных мутаций (Balakrishnan et al., 2008). В свою очередь, после прохождения сквозь «бутылочное горлышко» популяция с небольшим генетическим разнообразием сталкивается с высоким уровнем дрейфа генов и вероятностью потери аллелей, что может привести к дополнительному снижению генетического разнообразия. С одной стороны, такая популяция имеет меньшую способность к адаптации и выживанию в изменяющихся условиях окружающей среды, с другой стороны – возможно возникновение и эволюция новых видов.

Таким образом, естественный отбор может поддерживать полиморфизм и тем самым представлять основной интерес для изучения происхождения и сохранения редких аллелей, которые могут принимать участие в проявлении сельскохозяйственно-значимых признаков, а также дать представление об эволюционных путях, ведущих к современному состоянию генов у родственных видов.

*1.4. Методы оценки полиморфизма генов у люцерны*

Благодаря усовершенствованию технологий секвенирования, резко увеличился объем полученных данных о полиморфизме на уровне ДНК. Зачастую, скорость получения новых данных превышает скорость их обработки и понимания. Также активное развитие новых методов оценки полиморфизма, в частности - многофункциональных технических и программных средств, позволяет проводить многопрофильные исследования с большим количеством выборок в короткие сроки. С помощью полногеномного поиска ассоциаций (Genome-wide association study, GWAS) возможно проведение исследований, направленных на обнаружение взаимосвязей интересующих генов с фенотипическими признаками. В частности, с помощью GWAS и традиционного картирования локусов количественных признаков (Quantitative Trait Loci, QTL), были идентифицированы гены-кандидаты (SNP), лежащие в основе фенотипического изменения высоты *M. truncatula*, плотности трихом, времени цветения и клубенькообразования (Stanton-Geddes et al., 2013). Также благодаря GWAS на 174 модельных образцах *M. truncatula,* были идентифицированы локусы устойчивости к болезням, к различным изолятам возбудителя гнили корневых растений *Aphanomyces euteiches* (Bonhomme et al., 2019).

В настоящее время активно изучаются закономерности распределения аллелей SNP и гаплотипов (Salisbury et al., 2003). В частности, было идентифицировано несколько локусов устойчивости *M. sativa* к патогену *Verticillium wilt (VW)* (Yu et al., 2017)*.* Однако до сих пор остались неясны молекулярные механизмы этой устойчивости. Для интерпретации полученных результатов как правило требуется проведение дальнейших исследований. За последние несколько лет анализ гаплотипов получил широкое распространение в популяционно-генетических исследованиях. Картирование QTL *Pisum sativum* позволило получить потенциальные маркеры SNP, связанные с формой листа (Zheng et al., 2018). Одиночные аллельные SNP (Strand Displacement Amplification, SDA) используются для построения карт сцепления. Так были проанализированы ключевые признаки адаптации *M. sativa:* зимостойкость и устойчивость к полеганию (Adhikari et al., 2018). Оказалось, что эти признаки имеют независимое наследование и могут улучшаться по-отдельности.

С увеличением количества исследований, направленных на обнаружение SNP и определение сельскохозяйственно-значимых признаков, создаются базы данных, содержащие геномные последовательности различных живых организмов. Такие базы представляют собой общедоступный ресурс, в основе которого лежит описание структуры популяции и идентификации сегментов генома. В частности, база данных Hapmap Medicago (<http://www.medicagohapmap.org/>) содержит геномы 330 инбредных образцов *M. truncatula.* Hapmap Medicago позволяет изучать ассоциации всего генома, обнаруживать SNP, вставки и делеции (INDEL) с очень высоким разрешением. В настоящей работе мы также использовали базу данных Hapmap Medicago, с помощью которой был проведен поиск вариабельных позиций в последовательностях генов *MtCLE12*, *MtCLE13*, *MtCLE34* и *MtCLE35* у линий А17 и R108. На основе полученных сведений и количества доступной информации относительно геномов, возможно дальнейшее изучение полиморфизма генов *CLE*, регулирующих развитие клубеньков у *M. truncatula*.

Важное значение для оценки полиморфизма играют методы молекулярной эволюции. С их помощью возможно определять генетическую основу для эволюционно интересных признаков, а также выяснять причины мутаций и подтверждать их фенотипические последствия (Wray, 2013). Геномные технологии возродили интерес к фундаментальным вопросам эволюционной генетики и открыли ранее неизвестные эволюционные явления. В частности, ключевым моментом в молекулярной популяционной генетике является выявление типа естественного отбора, действующего в эволюции изучаемого гена, на основании встречающихся в популяции вариантов его нуклеотидной последовательности (Tajima, 1989). Известно, что полиморфизм в популяции может быть результатом не только мутаций и случайного дрейфа генов, но и естественного отбора. Для оценки значимости полиморфизма, а также выяснения его природы, необходимы критерии, отражающие наличие давления отбора. В частности, в качестве таких критериев используются тесты на нейтральность, позволяющие оценить отклонения выявленного полиморфизма в популяции от ожидаемых эволюционно нейтральных генетических изменений. Так, используется статистический тест на нейтральность Таджимы (Tajima, 1989). Тест основан на распределении частот аллелей изучаемых генов, а также представляет собой сравнение между средним числом замен нуклеотидов и полиморфных сайтов в образце. Тест используется для определения наличия давления отбора, действующего на изучаемые последовательности. В свою очередь, также используются оценки Уоттерсона (ThetaPerBP) и PiPerBP, которые являются основой теста Таджимы и позволяют сделать вывод об эволюционном режиме интересующих локусов (Watterson, 1975).

D-критерии теста на нейтральность Таджимы могут отражать отрицательное значение D-критерия (D<0), которое предполагает, что на изучаемую последовательность ДНК действовал положительный отбор, «выборочные зачистки» или в данный момент наблюдается расширение популяции после недавнего прохождения «бутылочного горлышка». Как правило, в таких популяциях наблюдаются редкие аллели, присутствующие с относительно высокой частотой (избыток редких аллелей). Также тест на нейтральность Таджимы может отражать положительное значение D-критерия (D>0), которое предполагает действие балансирующего отбора или внезапное сужение популяции. Однако в этом случае поддерживается, увеличивается или регулируется генетическая изменчивость без возникновения новых морфофизиологических адаптаций и жизненных форм, а также наблюдается отсутствие редких аллелей. Балансирующий отбор расширяет адаптивные возможности популяций без возникновения новых форм. Также тест на нейтральность Таджимы может показать нейтральное значение D-критерия (D=0), что предполагает нейтральную эволюцию ДНК и отсутствие доказательств отбора.

Также существует другой тест на нейтральность - Н-тест Фэя и Ву (Fay and Wu's H) (Fay, 2000), представляющий собой усовершенствованный тест Таджимы. С помощью теста Фэя и Ву возможно идентифицировать последовательности, подверженные «выборочным зачисткам» («selective sweep») в ходе эволюции. Важно, что отрицательное значение D-критерия (D<0) теста Таджимы не позволяет выяснить истинную причину неслучайной эволюции конкретной последовательности, а тест Фэя и Ву дает возможность точно установить тип естественного отбора и является ли наблюдаемый избыток редких аллелей следствием прохождения через «бутылочное горлышко». Также тест Фэя и Ву позволяет использовать аутгруппы, что поможет выяснить предковое состояние аллеля до разделения родов или видов. При этом комбинирование тестов Таджимы и Фэя и Ву поможет точно установить наличие положительного отбора, действовавшего на определённый локус в ходе эволюции, а также отличить эффект положительного отбора от эффекта прохождения «бутылочного горлышка».

Таким образом, для осуществления оценки полиморфизма у *M. truncatula* требуется применение ряда методов и средств, позволяющих эффективно проводить исследования с большим количеством выборок в короткие сроки. В частности, для достижения цели настоящей работы, мы провели оценку давления отбора на участках генов *CLE* с помощью методов молекулярной эволюции. Мы использовали D-критерий Таджимы для анализа последовательностей ДНК генов *MtCLE12*, *MtCLE13*, *MtCLE34* и *MtCLE35* у линий А17 и R108.

**2. МАТЕРИАЛЫ И МЕТОДЫ**

*2.1.* *Растительный материал*

В качестве объектов исследования в настоящей работе были использованы 2 линии бобового растения *Medicago truncatula* – A17 и R108, являющихся наиболее часто используемыми для исследований функциональной геномики *M. truncatula* (Li et al., 2014; Luo et al., 2016; Curtin et al., 2017).

Материал был представлен кафедрой генетики и биотехнологии СПбГУ.

*2.2. Выделение растительной ДНК*

Выделение ДНК из *M. truncatula* проводилось с использованием экстрагирующего буфера, содержащего цетилтриметиламмонийбромид (ЦТАБ). С помощью жидкого азота растирали растительный материал из молодых листьев *M. truncatula* и экстрагировали ДНК с помощью 2Х буфера с ЦТАБ - 65оC (40 мин). Чистка ДНК осуществлялась смесью хлороформ:изоамиловый спирт (24:1). Фазы разделяли центрифугированием. С помощью изопропанола осаждали ДНК, затем промывали осадок холодным этанолом 70%. Сушка осадка проводилась в ламинаре под током воздуха. В 30 мкл стерильной деионизованной воды растворяли ДНК.

*2.3. Амплификация фрагментов генов CLE*

Для уточнения последовательностей гена *MtCLE34* у двух линий *M. truncatula* нами были амплифицированы и секвенированы фрагменты этого гена. Для амплификации фрагментов генов использовали программу VectorNTI (Invitrogen, США). Было подобрано 4 праймера (таблица 1, рис. 10). Последовательность гена *MtCLE34* была взята в базе данных HapMap Medicago (<http://www.medicagohapmap.org/>).

Таблица 1. Праймеры, использованные для секвенирования гена *MtCLE34*.

|  |  |  |
| --- | --- | --- |
| **Ген** | **Название**  **праймера** | **Последовательность** |
| *MtCLE34* | MtCLE34\_flank\_FOR1 | TTTGTTGTGAAAATTGCCAAAACA |
| MtCLE34\_flank\_REV1 | TTCCTTTCCTCCAATTTCCCTAA |
| MtCLE34\_flank\_FOR2 | AGAAAGGGAACGGGGAGCATA |
| MtCLE34\_flank\_REV2 | TGAAGGTTGTCACAAATGAGTCG |

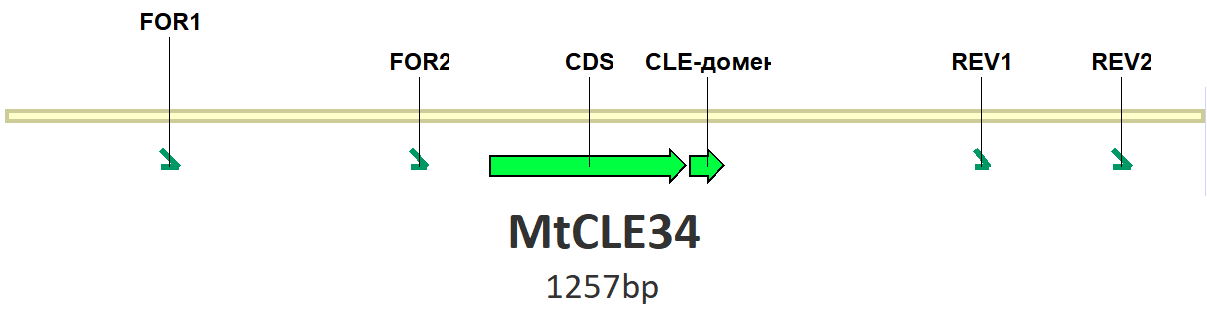
****

Рис. 10. Схема расположения праймеров.

Полимеразную цепную реакцию (ПЦР) проводили с помощью амплификатора CFX96 Real-Time PCR Detection System (Bio-Rad). Использовали необходимый набор реагентов в объеме 20 мкл, из которых: вода 12 мкл, полимераза Phusion High-Fidelity DNA Polymerase (ThermoFisher Scientific, США, кат.номер F530S) – 0,5 мкл (2 ед.акт./мкл, т.е. 1 ед. акт. на реакцию), буфер 5X (ThermoFisher Scientific, буфер прилагается к полимеразе) - 4 мкл, смесь дНТФ (Евроген, Россия) - 1 мкл (исходная концентрация 10 мМ каждого нуклеотида), праймеры (FOR1 и REV2) - 1 мкл каждого на реакцию (10 пкМ/мкл исходный раствор для каждого праймера). ПЦР проводили по следующей программе: денатурация – 95оС (20 секунд); отжиг праймеров – 52оC (30 сек); элонгация – 72оС (1 мин); температурный цикл амплификации повторяли 35 раз.

Электрофорез ПЦР-фрагментов осуществлялся в 1% агарозном геле, приготовленном на 1Х трис-ацетатном буфере (ТАЕ), в течение 20 мин при напряжении 120 В и мощности 80 Вт. Визуализация ПЦР-фрагментов в геле проводилась с применением интеркалирующего красителя бромистого этидия (0.1%) в ультрафиолетовом свете. Очищенные продукты ПЦР линий A17 и R108, соответствующие фрагментам гена *MtCLE34,* секвенировали в Ресурсном центре СПбГУ «Развитие молекулярных и клеточных технологий».

*2.4. Трансформация* *растений с помощью A. rhizogenes*

Растения *M. truncatula* трансформировали бактериями *A. rhizogenes,* несущими генетические конструкции *35S:MtCLE34* и *35S:GUS* (контроль), а также ген *GFP* под контролем промотора 35S для возможности отбора трансгенных корней. Такие конструкции были получены ранее в нашей лаборатории на основе вектора pB7WG2D (VIB, Гент, Бельгия). Трансформацию проростков люцерны осуществляли согласно протоколу (Limpens et al., 2004) следующим образом: проростки *M. truncatula* обрезали в области гипокотиля и на срез наносили суспензию агробактерий. Культивирование таких эксплантов проводили в течение 5 дней на твердой питательной среде Fahraeus (Fahraeus, 1957). Затем растения переносили на среду с добавлением антибиотика (цефотаксим, 300 мкг/мл) и культивировали до регенерации корней, после чего пересаживали в горшки, заполненные вермикулитом и инокулировали жидкой культурой бактерий *Sinorhizobium meliloti,* штамм 2011, для индукции клубенькообразования (1 мл культуры бактерий при оптической плотности OD600 на одно растение). Через 4 недели после инокуляции проводили подсчет клубеньков на корнях растений, при этом учитывали количество клубеньков как на трансгенных корнях (с флуоресценцией GFP), так и на нетрансгенных корнях (без флуоресценции GFP) с использованием флуоресцентного стереомикроскопа Leica M205FA (Leica, Германия) в ресурсном центре «Развитие молекулярных и клеточных технологий».

*2.5. Выделение РНК и обратная транскрипция*

Из трансгенных растений, содержащих конструкции 35S:MtCLE34 (в двух повторностях) и 35S:GUS (контроль) (в двух повторностях) была выделена РНК с помощью набора RNeasy Plant Mini Kit (QIAGEN, Германия) по приложенному протоколу Quick-Start Protocol RNeasy Plant Mini Kit, основанном на очистке РНК в несколько стадий с помощью колонок.

ДНКазную обработку РНК от примесей гДНК проводили с помощью фермента ДНКазы (1 мкл на реакцию, 1 ед./мкл) в течение 30 минут при 37оC с использованием RapidOut DNA Removal Kit (Thermo Fisher Scientific, США). Затем измерили концентрацию РНК на спектрофотометре Nanodrop (Thermo Scientific, США). Рассчитали необходимое количество РНК (300 нг) для постановки обратной транскрипции в соответствии с показаниями спектрофотометра.

Для обратной транскрипции использовали необходимый набор реагентов в объеме 20 мкл, из которых: РНК – 300 нг, буфер 5X - 4 мкл (ThermoFisher Scientific, США), дНТФ (Евроген, Россия) - 2 мкл (исходная концентрация 10 мМ каждого нуклеотида), праймеры OlygodT – 1 мкл (100 пкМ/мкл исходный раствор для каждого праймера), обратная транскриптаза RevertAid Transcriptase – 1 мкл (200 ед./мкл) (Thermo Scientific, США), ингибитор рибонуклеаз RNAzin – 1 мкл (20 ед./мкл) (Thermo Scientific, США) – 1 мкл. Обратную транскрипцию проводили при температуре 42°C в течение 60 минут, после чего реакцию останавливали при 70°C в течение 10 минут. Полученные образцы кДНК разводили водой для инъекций до объема 100 мкл и использовали для проведения ПЦР в реальном времени.

*2.6. Проведение ПЦР в реальном времени*

Для ПЦР в реальном времени было подготовлено 24 пробы, их них 12 проб – для амплификации с праймерами к референсному гену убиквитина *Ubi* (Ubi\_FOR 5’-ATGCAGATC/TTTTGTGAAGAC-3’, Ubi\_REV 5’-ACCACCACGG/AAGACGGAG-3’) и 12 проб – для амплификации с праймерами к гену *MtCLE34* (MtCLE34\_FOR AAGAAAATGGGCAACTTAAACAAA, MtCLE34\_REV TCAGTGGTGTCTAGGGTCAGG). В качестве матрицы использовали кДНК из контрольных растений со сверхэкспрессией бета глюкуронидазы (GUS) (GUS-1 и GUS-2), а также из растений со свекрхэкспрессией гена *MtCLE34* (CLE34-1 и CLE34-2), с каждым образцом проводили амплификацию в трех технических повторностях. Были подготовлены 2 стока: для Ubi и для CLE34. Использовали необходимый объем 20 мкл реагентов из набора ПЦР-Комплект (Синтол, Россия) на 1 пробу: 2 мкл буфер 10Х, 2 мкл дНТФ (исходная концентрация 10 mM каждого нуклеотида), 2 мкл MgCl2 (исходная концентрация 25mM), праймеры (FOR и REV для Ubi и отдельно FOR и REV для CLE34) - 1 мкл каждого праймера (10 пкМ/мкл исходный раствор для каждого праймера), 10,75 мкл вода, 0,25 мкл полимераза SynTaq (Синтол, Россия), 1 мкл ДНК.

ПЦР в реальном времени проводили с помощью системы CFX96 Real-Time PCR Detection System (Bio-Rad) при следующих тепловых циклах: 1 цикл - 94ºC (10 сек), 40 циклов - 94ºC (15 сек), 56ºC (30 сек), 72ºC (30 сек), затем повышение температуры каждые 5 секунд на 0.5ºС от 72ºС до 95ºС для оценки кривых плавления.

*2.7. Биоинформатический анализ*

Выравнивание фрагментов генов *MtCLE12, MtCLE13, MtCLE34, MtCLE35*, а также генов *CLE* других бобовых, их анализ, поиск вариабельных позиций в последовательностях генов *CLE* у линий *M. truncatula* и построение филогенетических деревьев осуществлялось в программах UniPro GENE (Okonechnikov et al., 2012), Vector NTI (Lu et al., 2004), MEGA-X (Kumar et al., 2018) и базе данных HapMap (<http://www.medicagohapmap.org/>) (Kang et al., 2015). В программах UniPro GENE и Vector NTI проводилось выравнивание нуклеотидных и аминокислотных последовательностей генов *CLE* по алгоритму Multiple sequence alignment (MSA) итерационным методом MUSCLE, а также размечались функционально значимые участки генов (в частности, последовательности, соответствующие доменам CLE). Применение метода MUSCLE объясняется использованием наиболее точных расстояний для оценки связи последовательностей. В программе MEGA-Х выполнялось построение филогенетических деревьев генов *CLE* *Arabidopsis thaliana*, *M. truncatula* и других бобовых, вовлеченных в AON. Также с использованием данных из базы Medicago truncatula Small Secreted Peptide Database (<http://mtsspdb.noble.org/database/>) (Boschiero et al., 2020) была проведена оценка экспрессии генов *CLE* в различных частях растения. С помощью программы MapInspect (http://mapinspect.apponic.com/) (Ma et al., 2013) было визуализировано распределение генов *MtCLE* по хромосомам.

Определение показателей генетической изменчивости, в том числе тест на нейтральность Таджимы (Tajima, 1989), проводилось с помощью программы Tassel (https://www.maizegenetics.net/tassel) (Bradbury et al, 2007). В свою очередь, также были проведены оценки Уоттерсона (ThetaPerBP) и PiPerBP. С помощью GWAS в программе Tassel с применением моделей Multidimensional Scaling (MDS), General Linear Model (GLM) и Mixed Minear Model (MLM) (Principal Component Analysis (PCA) + K), были построены графики (Q-Q Plot и Manhattan), отражающие вероятностное распределение участков кластеров генов *MtCLE* на хромосомах 2 и 4 с интересующими SNP: *MtCLE12* и *MtCLE13* вхромосоме 4, *MtCLE34* и *MtCLE35* в хромосоме 2, с фенотипическими характеристиками, в частности, с количеством образующихся клубеньков.

**3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ**

*3.1. Филогенетический анализ генов CLE*

Филогенетический анализ необходим не только для эволюционных и сравнительных исследований, ориентированных на выяснение эволюционных взаимоотношений, но и также для биохимических и молекулярных исследований, направленных на решение вопросов функций гена и белка. С помощью филогенетических деревьев можно визуализировать отношения между последовательностями.

Для построения филогенетического дерева со всеми генами семейства *CLE* (de Bang et al., 2017) у *M. truncatula* в программе MEGA-X (Kumar*,* 2018), использовался метод ближайшего соседа (Neighbour-Joining), основанный на расстояниях между последовательностями. Этот метод достаточно надежен для большинства наборов данных. Для удобства анализа был применен критерий загрузочной оценки >50% (Bootstrapped Neighbour-Joining Tree), отображающий только статистически надежные узлы.

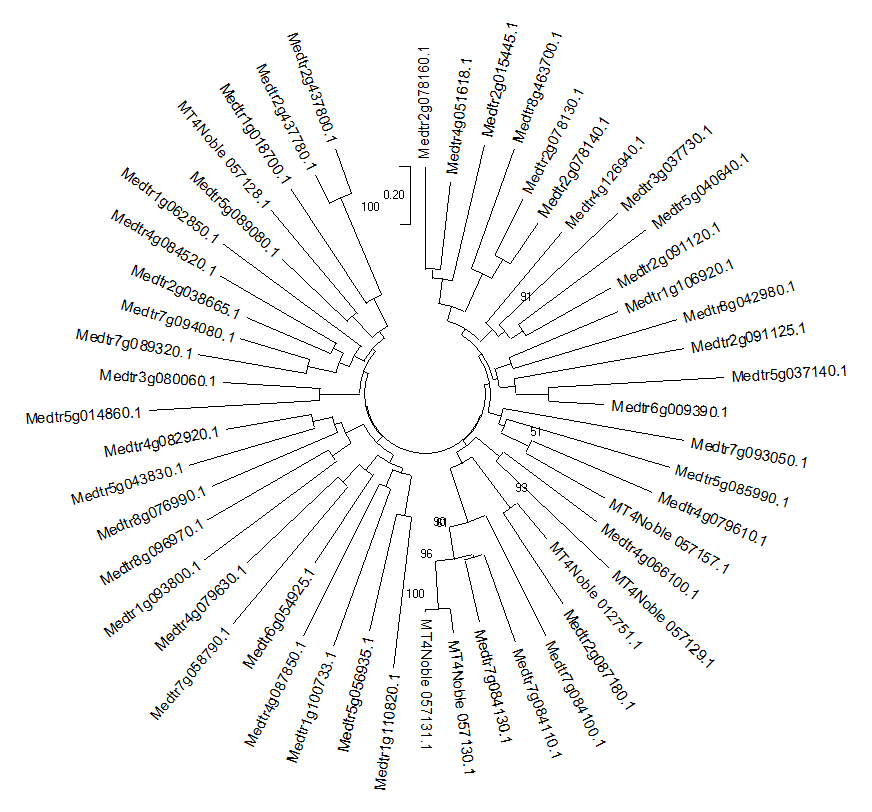


Рис. 11. Филогенетическое дерево со всеми генами семейства *CLE* у *M. truncatula*.

Полученное филогенетическое дерево приведено на рисунке 11. Процент реплицированных деревьев, в которых ассоциированные таксоны сгруппированы вместе в тесте bootstrap (500 реплик), показан рядом с ветвями. Эволюционные расстояния были вычислены с использованием метода максимального сложного правдоподобия (Tamura, 2004) и выражены в единицах числа базовых замен на участок. Этот анализ включал 52 нуклеотидные последовательности. Все неоднозначные позиции были удалены для каждой пары последовательностей (опция попарного удаления). В окончательном наборе данных насчитывалось в общей сложности 2007 позиций. После применения Bootstrapped Neighbour-Joining Tree отобразилось только 7 надежных узлов, среди которых: Medtr3g037730.1 (*MtCLE11*), Medtr5g085990.1 (*MtCLE17*), Medtr2g437780.1 (*MtCLE36*), Medtr2g437800.1 (*MtCLE37*), MT4Noble\_057129.1 (*MtCLE43*), MT4Noble\_057130.1 (*MtCLE47*), MT4Noble\_057131.1 (*MtCLE48*). Также дерево отражает разделение предковой последовательности генов *CLE* на Medtr2g091120.1 (*MtCLE34*) и Medtr5g040640.1 (*MtCLE24*). Интересно, что *MtCLE34* и *MtCLE24* расположены в разных хромосомах – 2 и 5. Medtr2g091125.1 (*MtCLE35*) располагается в отдельном узле, отдалившись от предковых последовательностей Medtr5g037140.1 (*MtCLE01/MtCLE25*) и Medtr6g009390.1 (*MtCLE02*). Medtr4g079630.1 (*MtCLE12*) и Medtr7g058790.1 (*MtCLE06*), также как Medtr4g079610.1 (*MtCLE13*) и MT4Noble\_057157.1 (*MtCLE19*) расположены в разных хромосомах – 7 и 4, и имеют общего предка.

Аналогичный метод (Neighbour-Joining) использовался для построения филогенетического дерева для аминокислотных последовательностей белков семейства CLE у *M. truncatula, A. thaliana* и других бобовых (рис. 12). Ветви, соответствующие разделам, воспроизведенным менее чем в 50% копий начальной загрузки, свернуты. Эволюционные расстояния были рассчитаны с использованием метода поправки Пуассона и выражены в единицах количества аминокислотных замен на сайт. Этот анализ включал 86 аминокислотных последовательностей. Все неоднозначные позиции были удалены для каждой пары последовательностей (опция парного удаления). Всего в итоговом наборе данных было 250 позиций. Построенное дерево на основе аминокислот позволит увидеть, какие CLE *M. truncatula* близки известным регуляторам клубенькообразования у других бобовых.

Согласно построенному дереву ближайшими гомологами CLE разных бобовых оказались: *MtCLE20* и *AtCLE43, MtCLE11* и *AtCLE45, MtCLE18* и *AtCLE12, GmRIC1* и *LjCLE-RS2, MtCLE13* и *LjCLE-RS1*. Гены *MtCLE34* и *GmNIC1* также являются ближайшими гомологами, расположенными тандемно на хромосоме 2 с *MtCLE35*. В свою очередь, ген *AtCLE1* также кластеризуется вместе с *MtCLE34* и *GmNIC1/2.*



Рис. 12. Филогенетическое дерево со всеми генами семейства CLE у *M. truncatula, Arabidopsis thaliana* и других бобовых.

*3.2. Распределение генов MtCLE по хромосомам*

С помощью программы MapInspect (http://mapinspect.apponic.com/) (Ma et al., 2013) было визуализировано распределение генов *MtCLE* по хромосомам (рис. 13). Оказалось, что некоторые гены *CLE* расположены тандемно в геноме (выделены красной рамкой). Как видно на рисунке 13, тандемное расположение имеют гены:

1. *MtCLE34* и *MtCLE35* находятся на расстоянии 9,6 Кб. Этигены *CLE* родственны генам *CLE*, ингибирующим клубенькообразование у других бобовых растений. Для *MtCLE35* недавно было показано участие в AON у люцерны (Lebedeva et al., 2020; Mens et al., 2021).
2. *MtCLE12* и *MtCLE13* - функционально характеризованные гены, для которых ранее было показано участие в AON (Mortier et al., 2010), находятся на расстоянии всего 6,3 Кб;
3. *MtCLE36* и *MtCLE37* с расстоянием в 6,7 Кб;
4. *MtCLE31* находится на расстоянии 6,7 Кб от *MtCLE32* и 15,3 Кб от *MtCLE33*;
5. *MtCLE09, MtCLE14* и *MtCLE46* находятся на расстоянии 9 Кб, а *MtCLE47* составляет 3,7 Кб от *MtCLE48.*

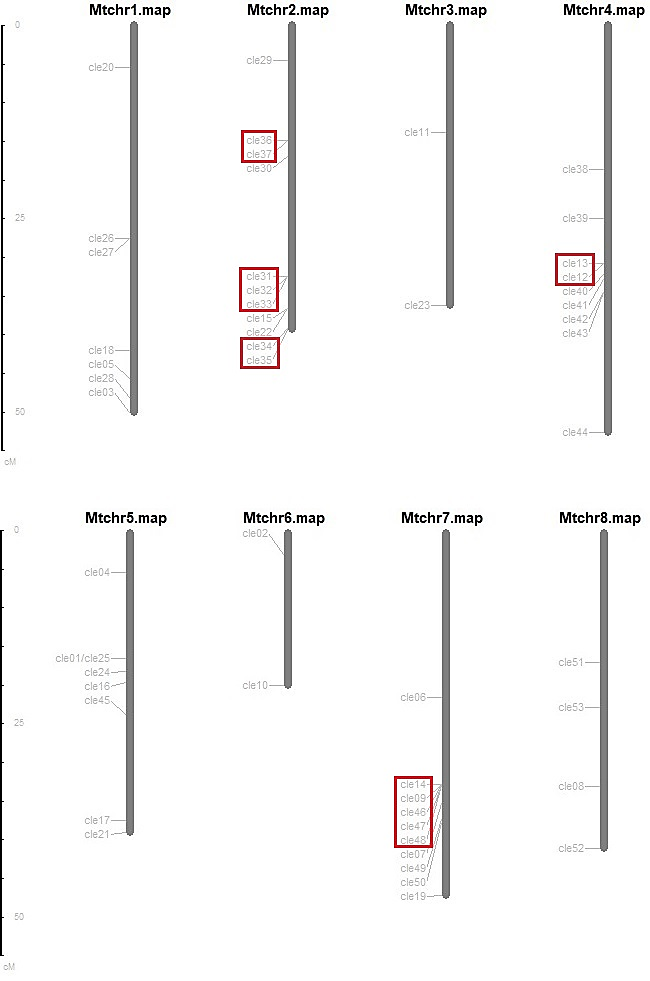


Рис. 13. Распределение генов *MtCLE* по хромосомам.

*3.3. Оценка экспрессии генов MtCLE с помощью доступных транскриптомных данных*

В базе данных *Medicago truncatula* Small Secreted Peptide Database (<http://mtsspdb.noble.org/database/>) была проведена оценка экспрессии генов *CLE* в клубеньках. В результате была составлена тепловая карта, содержащая сведения об экспрессии всех генов *CLE* *M. truncatula* в клубеньках (рис. 14)*.* Были выделены гены *CLE* с высоким уровнем экспрессии в клубеньках в разные периоды после инокуляции (красным, зеленым и черным цветом). Цвет соответствует относительным уровням экспрессии. На рисунке 14 приведены следующие условия эксперимента: 4wkNod – 28 дней, Nod0dpi – до инокуляции, Nod10dpi – 10 дней, Nod14dpi – 14 дней, Nod4dpi – 4 дня. Цвет отражает относительные уровни экспрессии генов (нормированные по медиане значения RPKM (reads per kilobase per million mapped reads) - количество ридов на килобазу на миллион картированных ридов).

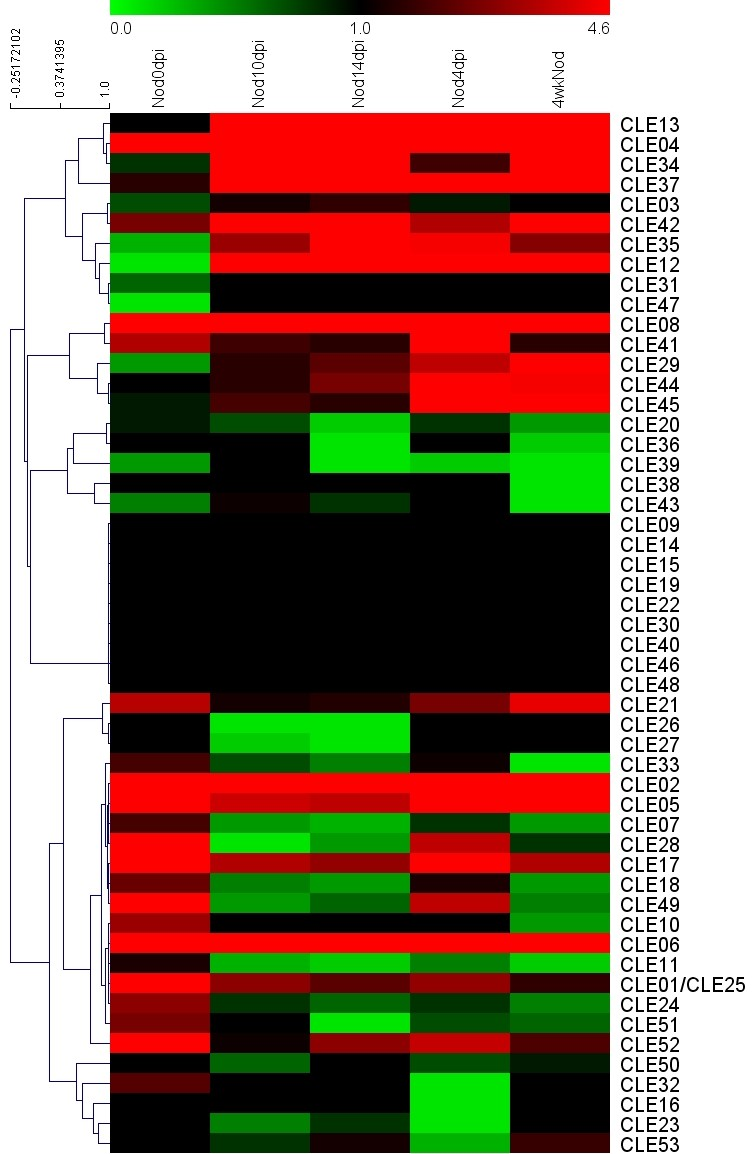


Рис. 14. Экспрессия генов *CLE* *M. truncatula* на разных сроках развития клубеньков.

После проведенного анализа было выделено 10 генов *CLE* с высокой экспрессией в клубеньках: *MtCLE02, MtCLE04, MtCLE06, MtCLE08, MtCLE12, MtCLE13, MtCLE34, MtCLE35, MtCLE37, MtCLE42.* При этом для генов *MtCLE02, MtCLE04, MtCLE06, MtCLE08* характерны относительно высокие уровни экспрессии как в контроле (Nod0dpi), так и при развитии клубеньков, тогда как для генов *MtCLE12, MtCLE13, MtCLE34, MtCLE35, MtCLE37, MtCLE42* наблюдаются более высокие уровни экспрессии в инокулированных корнях, по сравнению с контролем (Nod0dpi). Для наглядности экспрессия этих 10 генов *CLE* в клубеньках представлена в виде графика (рис. 15).

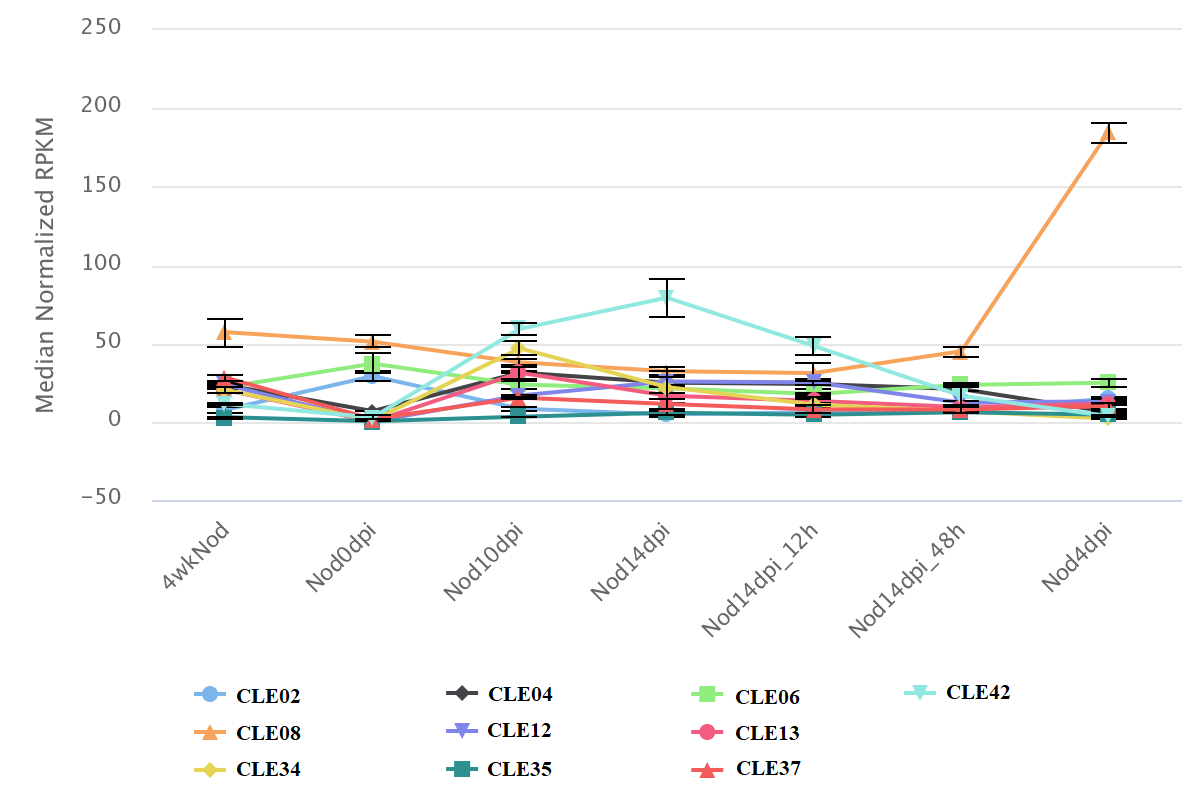


Рис. 15. Экспрессия 10 генов *CLE* в клубеньках. По оси абсцисс: дни после инокуляции (дпи, dpi); по оси ординат: относительные уровни экспрессии генов (нормированные по медиане значения RPKM).

Также мы оценили экспрессию генов *CLE* в других органах растения (рис. 16). Как видно на построенной тепловой карте, гены *MtCLE04, MtCLE34, MtCLE35, MtCLE42* – «корневые» CLE, экспрессирующиеся в корнях и в клубеньках. Гены *MtCLE12* и *MtCLE13* являются клубенек-специфичными, их экспрессия в других органах растений наблюдается на относительно низком уровне.

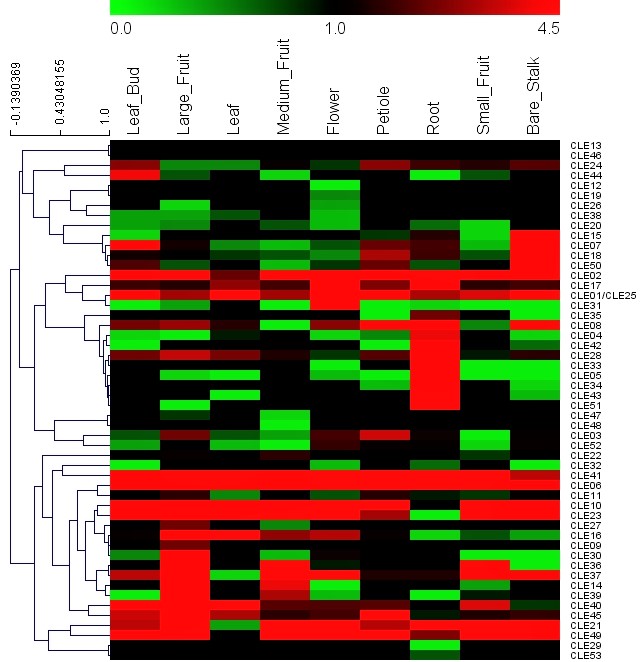


Рис. 16. Экспрессия генов *CLE* *M. truncatula* в разных частях растения.

Также был проведен анализ наличия схожего паттерна экспрессии генов *CLE* в разных органах у генов *CLE*, которые расположены кластерами на хромосомах. В результате среди тандемно расположенных на хромосоме генов *CLE*, были обнаружены схожие паттерны экспрессии. Экспрессия расположенных тандемно на хромосоме 2 генов *MtCLE12* и *MtCLE13* специфична для клубеньков. Гены *MtCLE34* и *MtCLE35* экспрессируются в клубеньках и в корнях. Похожие паттерны экспрессии наблюдаются у генов *MtCLE36* и *MtCLE37*, среди них *MtCLE37* экспрессируется в клубеньках, а у *MtCLE36* в целом выражен более низкий уровень экспрессии в разных частях растения относительно *MtCLE37.* У генов *MtCLE09, MtCLE14, MtCLE47*, *MtCLE46* и *MtCLE48*, расположенных в виде кластера на хромосоме 7, наблюдается относительно низкий уровень экспрессии во всех органах растения. Можно отметить относительную схожесть экспрессии генов *MtCLE09, MtCLE14, MtCLE47* в плоде. На основании проведенного анализа можно заключить, что среди генов *CLE*, расположенных кластерами на хромосомах, в частности, у генов *MtCLE34* и *MtCLE35,* выражена схожесть паттерна экспрессии генов *CLE* в разных органах растения. Более того, недавно было показано, что оба эти гена активируются у люцерны при добавлении нитрата (Lebedeva et al., 2020; Mens et al., 2021). Таким образом, для ряда близко расположенных генов *CLE* наблюдается тенденция к схожей регуляции экспрессии генов в разных органах растений и при различных условиях.

Также стоит отметить ряд близкородственных генов *CLE* с высоким уровнем экспрессии в клубеньках *M. truncatula*. Согласно построенному филогенетическому дереву, представленному ранее (рис. 12), близкородственными являются: *MtCLE02* - *AtCLE26* и *MtCLE01/MtCLE25*; *MtCLE04 – MtCLE16* и *MtCLE40*; *MtCLE06* - *AtCLE41* и *AtCLE44*; *MtCLE08* - *AtCLE8*; *MtCLE12* - *MtCLE13*, *LjCLE-RS1*; *MtCLE34* – *GmNIC1, GmNIC2* и *AtCLE1*; *MtCLE35 – MtCLE38*; *MtCLE37 - MtCLE36; MtCLE42 – MtCLE43*.

*AtCLE26* также экспрессируется в корневой системе, однако *MtCLE01/MtCLE25* в большей степени в цветке. *MtCLE16* и *MtCLE40* тоже отличаются местом экспрессии – в плоде. При этом *MtCLE40* вообще не экспрессируется в корне. Ближайший гомолог *AtCLE41* и *AtCLE44* также экспрессируются в корне, а *MtCLE06* не проявляет экспрессии в семенах и относительно своей интенсивности экспрессии в плоде, в меньшей степени экспрессируется в корне. Самый высокий уровень экспрессии в клубеньках отмечен у *MtCLE08*. Его ближайший гомолог *AtCLE8* участвует в пролиферации и дифференцировке клеток, и действует в эмбриогенезе. Гены *LjCLE-RS1, MtCLE12* и *MtCLE13* участвуют в регуляции клубенькообразования. Ген *MtCLE34*, который, как мы предполагаем, также может быть вовлечен в контроль клубенькообразования, демонстрирует наибольшее сходство с нитрат-индуцируемым геном сои *GmNIC1*, сверхэкспрессия которого у сои также локально подавляла развитие клубеньков на трансгенных корнях (Lim et al., 2014). Другой же гомолог - *AtCLE1* имеет специфические паттерны экспрессии в корне и участвует в контроле корневой архитектуры у *A. thaliana*.

Таким образом, обнаружены гены *CLE* *M. truncatula* с высоким уровнем экспрессии в клубеньках *MtCLE02, MtCLE04, MtCLE06, MtCLE08, MtCLE12, MtCLE13, MtCLE34, MtCLE35, MtCLE37, MtCLE42.* При этом гены *MtCLE34* и *MtCLE35* экспрессируются преимущественно в клубеньках, но также и в корнях без инокуляции, а экспрессия генов *MtCLE12* и *MtCLE13* является клубенек-специфичной. Известно, что гены *MtCLE12* и *MtCLE13* являются регуляторами AON (Mortier et al., 2010). Другие гены *CLE* также, возможно, могут быть задействованы в клубенькообразовании, но вероятно, помимо этого, имеют еще и другие функции у растений. Недавно в нашей лаборатории была показана ингибирующая роль гена *MtCLE35:* его сверхэкспрессия в трансгенных корнях люцерны системно подавляла образование симбиотических клубеньков (Lebedeva et al., 2020), что также было подтверждено и в исследованиях других авторов (Mens et al., 2021). Роль гена *MtCLE34* в клубенькообразовании к настоящему времени не изучена, но его сходство с нитрат-индуцируемым геном сои *GmNIC1*, ингибирующим развитие клубеньков, позволяет предположить, что он также может быть задействован в авторегуляции клубенькообразования у люцерны. Таким образом, для более детального изучения полиморфизма нами были выбраны четыре гена люцерны, для которых показано или предполагается участие в системном контроле клубенькообразования: *MtCLE12, MtCLE13, MtCLE34, MtCLE35*

*3.4. Оценка полиморфизма генов MtCLE люцерны*

В ходе проведенного поиска вариабельных позиций в последовательностях генов *CLE* у 262 линий *M. truncatula*, доступных в базе данных HapMap, были обнаружены существующие SNP и составлена таблица с SNP в генах *MtCLE12, MtCLE13, MtCLE34, MtCLE35* (таблица 2).

Таблица 2. Полиморфизм генов *MtCLE12, MtCLE13, MtCLE34, MtCLE35.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ген** | **SNP** | **Codon Change** | **Effect** | **Частота альтернативной аллели** |
| ***MtCLE12*** | 30800355  30800406  30800452  30800472 | (A/C) tcA/tcC  (T/C) tcT/tcC  (A/C) Aat/Cat  (C/T) ctC/ctT | Синонимичная  Синонимичная  Несинонимичная  (N37H)  Синонимичная | 0.0035  0.064  0.142  0.0035 |
| ***MtCLE13*** | 30793865  30793889 | (A/G) tcA/tcG  (C/T) cgC/cgT | Синонимичная  Синонимичная | 0.169  0.01 |
| ***MtCLE34*** | 39237030  39237043  39237092 | (T/A) ATG/AAG  (T/C) AAT/ AAC  (T/C) **TAG** /CAG | Несинонимичная  (M30N)  Синонимичная  **Stop-кодон** | 0.201  0.0071  0.983 |
| ***MtCLE35*** | 39246840  39246899  39246937  39246939  39246968 | (C/T) Cta/Tta  (C/T) caC/caT  (A/G) cAt/cGt  (G/A) Gtg/Atg  (A/G) ctA/ctG | Синонимичная  Синонимичная  Несинонимичная  (H43R)  Несинонимичная  (V44R)  Синонимичная | 0.042  0.023  0.01  0.092  0.08 |

Были обнаружены существующие SNP, из которых в гене *MtCLE13* - только синонимичные замены, что может свидетельствовать о консервативности этого гена и значимости его функций для жизни растения, а также о возможном существовании направленного отбора в эволюции. У генов *MtCLE12, MtCLE34* и *MtCLE35* обнаружены синонимичные и несинонимичные замены. В том числе в базе данных HapMap был обнаружен стоп-кодон в гене *MtCLE34* у линии А17 (линия HM101 согласно номенклатуре линий в базе HapMap) (рис*.* 17), а также еще у трех других линий (линии HM256, HM038 и HM058) из всех представленных в базе HapMap 262 линий.В свою очередь, предположительно, в базе данных ошибочно определены интрон-экзонные границы для гена *MtCLE34*: в гене *MtCLE34* аннотирован интрон размером более 5000 п.о. При этом на C-конце предполагаемого белкового продукта содержится CLE-домен. А в ORF (Open Reading Frame) соответствующей C-концевой части белка и части последовательности интрона присутствует стоп-кодон у линии A17, что должно приводить к синтезу укороченного белка без CLE-домена. У другой лабораторной линии *M. truncatula*, R108, согласно последовательностям, представленным в базе данных HapMap, стоп-кодон отсутствует, таким образом, у линии R108 ген *CLE34* может кодировать функциональный продукт. Для проверки были секвенированы фрагменты ПЦР на матрице ДНК, соответствующие гену *MtCLE34* у линий A17 и R108. Результаты секвенирования рассмотрены в следующем разделе.

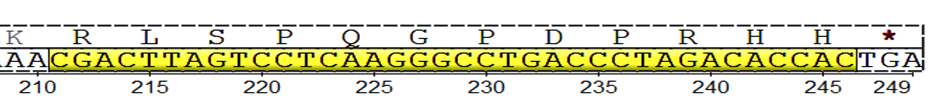


Рис. 17. Стоп-кодон у линии А17.

*3.5. Определение последовательностей фрагментов гена MtCLE34 с помощью секвенирования*

Для секвенирования фрагментов ПЦР на матрице ДНК, соответствующих гену *MtCLE34* у линий A17 и R108, были сделаны 4 пробы линий в двух повторностях (итого 8 проб). Пробы выглядели следующим образом:

1. 2 пробы R108 (F1+R1) – R108F1R1-1, R108F1R1-2;
2. 2 пробы R108 (F2+R2) - R108F2R2-1, R108F2R2-2;
3. 2 пробы A17 (F1+R1) – A17F1R1-1, A17F1R1-2;
4. 2 пробы A17 (F2+R2) - A17F2R2-1, A17F2R2-2.

Секвенирование проводили с прямых праймеров F1 и F2, подобранным к фланкирующему участку перед CDS. Из полученных ПЦР-смесей с ПЦР-фрагментами линий A17 и R108, пробу R108F1R1-2 и A17F1R1-2 передали в ресурсный центр СПбГУ на секвенирование. После получения сиквенсов, было выполнено выравнивание с последовательностью гена *MtCLE34*, взятойиз базы данных HapMap (рис. 18). Как можно видеть на приведенном выравнивании, последовательность нуклеотидов из пробы R108F1R1-2 соответствует последовательности фрагмента гена *MtCLE34* линии R108, взятой из базы данных, стоп-кодон в кодирующей области отсутствует. Интересно, что у линии R108 наблюдается также делеция в 3' фланкирующем участке после CDS (рис. 19).

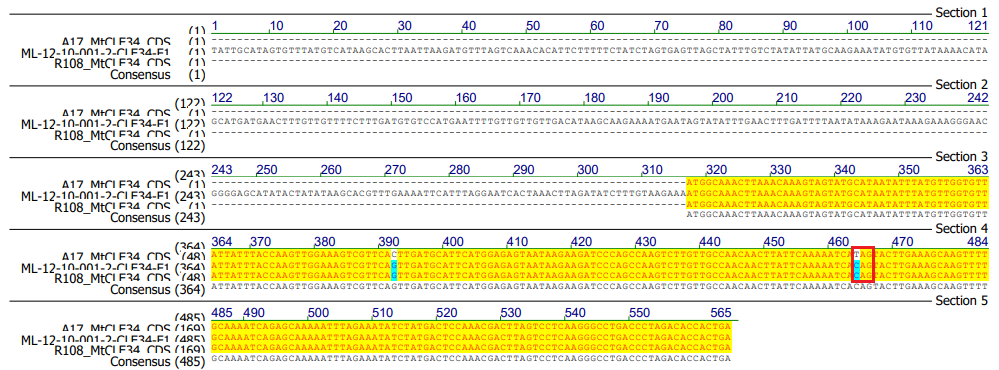


Рис. 18. Выравнивание пробы R108F1R1-2 с последовательностями R108 и А17 из баз.



Рис. 19. Делеция у линии R108 в 3' фланкирующем участке после CDS.

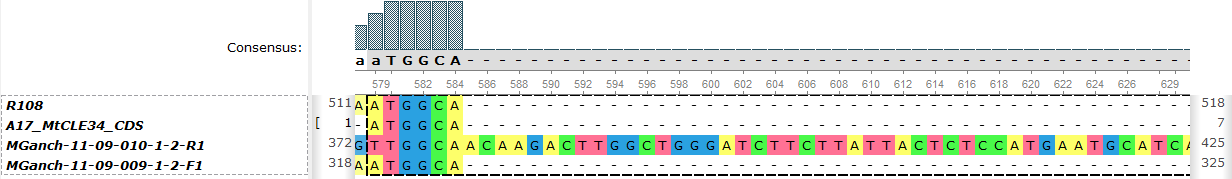


Рис. 20. Выравнивание повторной пробы R108F1R1-2 с последовательностями R108 и А17 из баз.

*3.6.* *Определение влияния отбора в участках генов MtCLE*

Для оценки значимости мутационных изменений было проведено определение показателей генетической изменчивости с помощью программы Tassel. Архив с данными по SNP для всех хромосом *M. truncatula* был взят из базы данных Hapmap Medicago. Была проведена фильтрация данных - выбраны участки кластеров генов *CLE* на хромосомах 2 и 4 с интересующими SNP: для *MtCLE12* и *MtCLE13* промежуток 30790000…30804500 в хромосоме 4, а для *MtCLE34* и *MtCLE35* промежуток 39229000…39253183 в хромосоме 2. В результате проведенной фильтрации данных, была получена 61 позиция, которые были пронумерованы в порядке возрастания (таблица 3).

Таблица 3. Позиции кластеров генов *CLE* с SNP на хромосомах.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **№ п/п** | **Хромосома** | **Позиция на хромосоме** | **Гены** | **SNP** |
| 1 | 2 | 39 229 002…39 229 622 | *Intergenic region (IGR)* | 39229002  39229052  39229268  39229506 |
| 2 | 2 | 39 229 631…39 230 416 | *Intergenic region (IGR)* | 39229634  39229775  39230011  39230285 |
| 3 | 2 | 39 230 428…39 230 888 | *Intergenic region (IGR)* | 39230436  39230581  39230878 |
| 4 | 2 | 39 230 892…39 231 282 | *Intergenic region (IGR)* | 39230911  39230941  39231028  39231082  39231278 |
| 5 | 2 | 39 231 285…39 231 889 | *Intergenic region (IGR)* | 39231371  39231407  39231716  39231874 |
| 6 | 2 | 39 231 890…39 232 848 | *Medtr2g091120.1*  *hypothetical protein*  *(CDS)* | 39231925  39231985  39232116  39232505  39232697  39232760 |
| 7 | 2 | 39 232 851…39 233 113 | *Medtr2g091120.1*  *hypothetical protein*  *(Intron)* | 39232856  39232883  39232969  39233078  39233105 |
| 8 | 2 | 39 233 116…39 233 509 | *Medtr2g091120.1*  *Intergenic region (IGR)* | 39233116  39233186  39233247  39233472 |
| 9 | 2 | 39 233 513…39 234 139 | *Medtr2g091120.1*  *hypothetical protein*  *Intergenic region (IGR)* | 39233515  39233592  39233812  39233879  39234139 |
| 10 | 2 | 39 234 144…39 234 775 | *Medtr2g091120.1*  *hypothetical protein*  *Intergenic region (IGR)* | 39234144  39234222  39234463 |
| 11 | 2 | 39 234 782…39 235 783 | *Medtr2g091120.1*  *hypothetical protein*  *Intergenic region (IGR)* | 39234789  39235217  39235280  39235745 |
| 12 | 2 | 39 236 010...39 237 260 | *MtCLE34*  *(CDS)* | 39237030  39237043  39237092 |
| 13 | 2 | 39 237 263...39 237 963 | *Intergenic region (IGR)* | 39237263  39237294  39237449  39237500  39237757  39237963 |
| 14 | 2 | 39 237 964...39 238 426 | *Intergenic region (IGR)* | 39237964  39238023  39238182  39238274  39238392 |
| 15 | 2 | 39 238 440...39 238 963 | *Intergenic region (IGR)* | 39238452  39238496  39238722  39238938  39238961 |
| 16 | 2 | 39 238 965...39 239 360 | *Intergenic region (IGR)* | 39238987  39239007  39239135  39239254  39239292 |
| 17 | 2 | 39 239 361...39 240 113 | *Intergenic region (IGR)* | 39239383  39239440  39239486  39239588  39239928  39240099 |
| 18 | 2 | 39 240 119...39 241 175 | *Intergenic region (IGR)* | 39240119  39240160  39240270  39240423  39240842  39241028  39241150 |
| 19 | 2 | 39 241 185...39 242 425 | *Intergenic region (IGR)* | 39241185  39241208  39241262  39241333  39241409  39241476  39242159  39242419 |
| 20 | 2 | 39 242 436...39 243 012 | *Intergenic region (IGR)* | 39242436  39242488  39242647  39242753  39242909 |
| 21 | 2 | 39 243 016...39 243 436 | *Intergenic region (IGR)* | 39243027  39243039  39243146  39243350  39243411 |
| 22 | 2 | 39 243 438...39 243 919 | *Intergenic region (IGR)* | 39243817  39243826  39243833  39243838  39243898 |
| 23 | 2 | 39 243 920...39 244 483 | *Intergenic region (IGR)* | 39243920  39243965  39244354  39244381  39244483 |
| 24 | 2 | 39 244 486...39 244 856 | *Intergenic region (IGR)* | 39244486  39244510  39244711  39244763  39244802 |
| 25 | 2 | 39 244 889...39 245 305 | *Intergenic region (IGR)* | 39244914  39244941  39244970  39245038  39245122  39245138  39245227 |
| 26 | 2 | 39 245 310...39 245 817 | *Intergenic region (IGR)* | 39245376  39245432  39245461  39245600  39245671  39245710  39245750  39245808 |
| 27 | 2 | 39 245 820...39 246 159 | *Intergenic region (IGR)* | 39245820  39245834  39245867  39245878  39245926  39245958  39245999  39246118  39246142 |
| 28 | 2 | 39 246 167...39 246 509 | *Intergenic region (IGR)* | 39246168  39246198  39246236  39246279  39246384  39246428  39246447  39246506 |
| 29 | 2 | 39 246 512...39 247 439 | *MtCLE35*  *(CDS)* | 39246840 39246899 39246937 39246939 39246968 |
| 30 | 2 | 39 247 446…39 247 852 | *Intergenic region (IGR)* | 39247468  39247622  39247796 |
| 31 | 2 | 39 247 855…39 248 143 | *Intergenic region (IGR)* | 39247855  39247924  39248011  39248122 |
| 32 | 2 | 39 248 145…39 248 513 | *Intergenic region (IGR)* | 39248150  39248196  39248366  39248502 |
| 33 | 2 | 39 248 514…39 249 194 | *Medtr2g091130.1*  *transcription factor GTE4-like protein (CDS)* | 39248587  39248755  39248938  39249194 |
| 34 | 2 | 39 249 196…39 249 630 | *Medtr2g091130.1*  *transcription factor GTE4-like protein (CDS)* | 39249196  39249276  39249389  39249540 |
| 35 | 2 | 39 249 631…39 250 382 | *Medtr2g091130.1*  *transcription factor GTE4-like protein (CDS)* | 39249645  39249765  39250128  39250359 |
| 36 | 2 | 39 250 407…39 250 968 | *Medtr2g091130.1*  *transcription factor GTE4-like protein (CDS)* | 39250664  39250755  39250872 |
| 37 | 2 | 39 251 009…39 251 475 | *Medtr2g091130.1*  *transcription factor GTE4-like protein (CDS)* | 39251148 |
| 38 | 2 | 39 251 487…39 252 587 | *Medtr2g091130.1*  *transcription factor GTE4-like protein (Intron)* | 39251487  39251681  39252164  39252578 |
| 39 | 4 | 30 790 003…30 790 455 | *Intergenic region (IGR)* | 30790003  30790160  30790275  30790319 |
| 40 | 4 | 30 790 457…30 790 889 | *Intergenic region (IGR)* | 30790484  30790557  30790791 |
| 41 | 4 | 30 790 890…30 791 160 | *Intergenic region (IGR)* | 30790944  30791028 |
| 42 | 4 | 30 791 181…30 792 117 | *Intergenic region (IGR)* | 30791231  30791443  30792012 |
| 43 | 4 | 30 792 136…30 792 535 | *Intergenic region (IGR)* | 30792188  30792457 |
| 44 | 4 | 30 792 537…30 793 500 | *Intergenic region (IGR)* | 30793264  30793393  30793489 |
| 45 | 4 | 30 793 508...30 794 412 | *MtCLE13*  *(CDS)* | 30793865 30793889 |
| 46 | 4 | 30 794 413...30 794 895 | *Intergenic region (IGR)* | 30794424  30794502  30794673  30794811  30794823 |
| 47 | 4 | 30 794 896...30 795 417 | *Intergenic region (IGR)* | 30794896  30795016  30795093  30795280 |
| 48 | 4 | 30 795 438...30 795 895 | *Intergenic region (IGR)* | 30795438  30795480  30795642  30795835 |
| 49 | 4 | 30 795 914...30 796 519 | *Intergenic region (IGR)* | 30795914  30796012  30796229  30796378  30796519 |
| 50 | 4 | 30 796 522...30 796 960 | *Medtr4g079620*  *hypothetical protein*  *(CDS)* | 30796599  30796619  30796749  30796796  30796904 |
| 51 | 4 | 30 796 961...30 797 519 | *Intergenic region (IGR)* | 30796981  30796991  30797098  30797289  30797350  30797453 |
| 52 | 4 | 30 797 524...30 797 926 | *Intergenic region (IGR)* | 30797524  30797543  30797683  30797862  30797926 |
| 53 | 4 | 30 797 944...30 798 912 | *Intergenic region (IGR)* | 30797976  30798047  30798647  30798681  30798810 |
| 54 | 4 | 30 798 914...30 799 450 | *Intergenic region (IGR)* | 30798946  30798985  30799067  30799136  30799264  30799388 |
| 55 | 4 | 30 799 455...30 799 983 | *Intergenic region (IGR)* | 30799470  30799524  30799604  30799649  30799872 |
| 56 | 4 | 30 799 990...30 800 791 | *MtCLE12*  *(CDS)* | 30800355 30800406 30800452 30800472 |
| 57 | 4 | 30 800 837…30 801 786 | *Intergenic region (IGR)* | 30800941  30801234  30801754 |
| 58 | 4 | 30 801 800…30 802 531 | *Intergenic region (IGR)* | 30801803  30802226  30802387  30802523 |
| 59 | 4 | 30 802 532…30 802 968 | *Intergenic region (IGR)* | 30802542  30802657  30802718  30802836  30802956 |
| 60 | 4 | 30 802 969…30 803 690 | *Intergenic region (IGR)* | 30802972  30803300  30803437  30803659 |
| 61 | 4 | 30 803 708…30 804 324 | *Intergenic region (IGR)* | 30803708  30803979  30804306  30804324 |

Затем был проведен тест на нейтральность Таджимы (Tajima, 1989) и оценки Уоттерсона (ThetaPerBP) и PiPerBP, которые являются основой теста Таджимы и позволяют сделать вывод об эволюционном режиме интересующих локусов. Выбор этих тестов обусловлен их наибольшей распространенностью.

На основании проведенного теста были построены графики, отражающие основные значения показателей TajimaD, ThetaPerBP и PiPerBP для хромосомы 2 (рис. 21) и 4 (рис. 22), а также изображена схема расположения генов. Отрицательное значение D-критерия (D<0) предполагает, что на изучаемую последовательность ДНК действовал положительный отбор, «выборочные зачистки» или в данный момент наблюдается расширение популяции после недавнего прохождения «бутылочного горлышка». Как правило, в таких популяциях наблюдаются редкие аллели, присутствующие с относительно высокой частотой (избыток редких аллелей). В случае положительного значения D-критерия (D>0), предполагается действие балансирующего отбора или внезапное сужение популяции. Однако в этом случае поддерживается, увеличивается или регулируется генетическая изменчивость без возникновения новых морфофизиологических адаптаций и жизненных форм, а также наблюдается отсутствие редких аллелей. Балансирующий отбор расширяет адаптивные возможности популяций без возникновения новых форм. Нейтральное значение D-критерия (D=0) предполагает нейтральную эволюцию ДНК и отсутствие доказательств отбора.

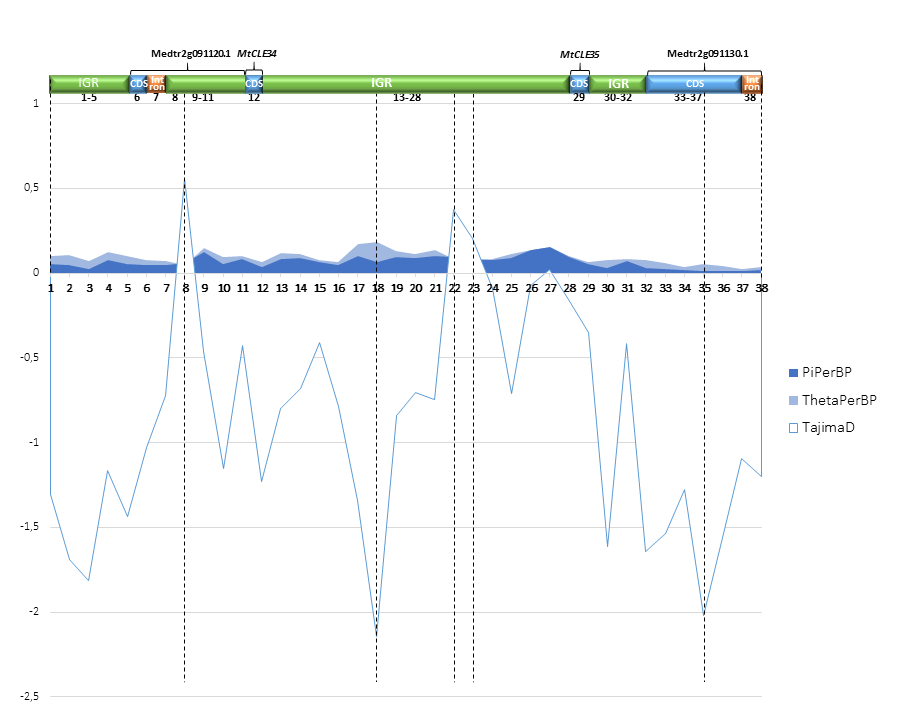


Рис. 21. График показателей TajimaD, ThetaPerBP и PiPerBP кластеров генов *CLE* с SNP и схема расположения этих генов на хромосоме 2.

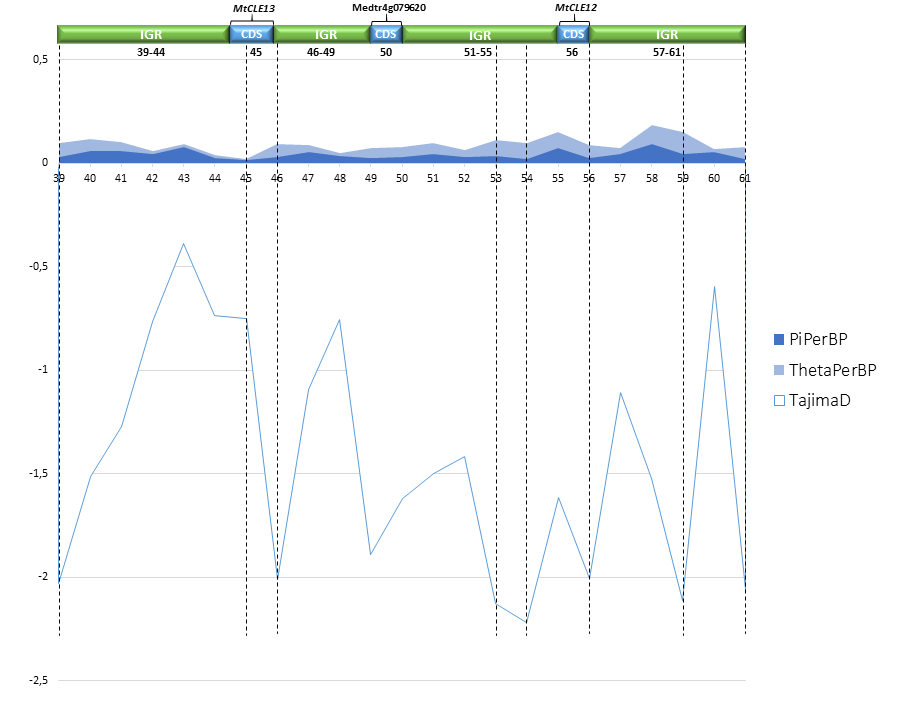


Рис. 22. График показателей TajimaD, ThetaPerBP и PiPerBP кластеров генов *CLE* с SNP и схема расположения этих генов на хромосоме 4.

Согласно полученным результатам, на подавляющее большинство участков хромосом 2 и 4 под номерами 1-7, 9-21, 24-26, 28-61 действует положительный отбор (D<0). В частности, на такие участки кластеров генов *CLE* на хромосомах 2 и 4 с интересующими SNP генов, как *MtCLE34* (номер 12), *MtCLE35* (номер 29), *MtCLE12* (номер 56) и *MtCLE13* (номер 45). У *MtCLE35* (номер 29) наблюдается меньший показатель относительно других генов - *MtCLE34* (номер 12), *MtCLE12* (номер 56) и *MtCLE13* (номер 45). Важно, что SNP в этих генах расположены в кодирующих последовательностях (CDS).

Таблица 4. Показатели D-критериев и SNP генов *MtCLE12, MtCLE13, MtCLE34, MtCLE35.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Название гена | PiPerBP | ThetaPerBP | TajimaD | Число SNP | Отношение числа несинонимичных замен к синонимичным в пределах CDS (nsy/syn) |
| *MtCLE12* | 0,02415 | 0,0846 | -2,00485 | 4 | 1 nsy/3 syn |
| *MtCLE13* | 0,01263 | 0,02017 | -0,75137 | 2 | 2 syn |
| *MtCLE34* | 0,0355 | 0,06478 | -1,23074 | 3 | 2 nsy/1 syn |
| *MtCLE35* | 0,05526 | 0,06376 | -0,35079 | 5 | 2 nsy/3 syn |

Отталкиваясь от показателей D-критериев, возможно дать оценку изменчивости интересующим генам и SNP в этих генах (таблица 4). В частности, гены *MtCLE12* (номер 56) и *MtCLE34* (номер 12) по результатам теста Таджимы, имеют самые низкие показатели D-критериев, что может свидетельствовать о их меньшей вариабельности и консервативности. Также согласно отношению числа несинонимичных замен к синонимичным (nsy/syn) видно, что у гена *MtCLE12* (номер 56) выражено преобладание syn над nsy, что может быть обусловлено важностью выполняемых этим геном функций. Так, известно, что ген *MtCLE12* (номер 56) строго специфично экспрессируется в клубеньках и незначительно в корнях и, по-видимому, действие строгого отбора обусловлено сохранением этого гена для выполнения важных функций в растении. Это согласуется с полученным низким показателем D-критериев. Однако для гена *MtCLE34* (номер 12) анализ SNP показал преобладание nsy над syn. Такой результат может означать как избыток редких аллелей, так и недавно произошедшую «выборочную зачистку» (selective sweep). При этом «выборочные зачистки» в случае положительного отбора могут привести к избытку низкочастотных вариантов.

В свою очередь, гены *MtCLE13* (номер 45) и *MtCLE35* (номер 29) согласно показателям D-критериев, являются более вариабельными генами. При этом у этих генов наблюдается преобладание syn над nsy, что может означать важность функций генов для растения. Так, *MtCLE13* (номер 45) наряду с *MtCLE12* (номер 56) строго специфично экспрессируется в клубеньках, но при этом не экспрессируется в корнях, а гены *MtCLE35* (номер 29) и *MtCLE34* (номер 12) являются корневыми CLE, экспрессирующимися в корнях и в клубеньках. Отсутствие строгого отбора и преобладание syn над nsy, может указывать на недавнее прохождение «бутылочного горлышка» и/или «выборочной зачистки» (selective sweep). Однако, при этом в гене *MtCLE13* были выявлены только две синонимичные замены, и согласно этому показателю – она является наименее вариабельным среди изученных генов *CLE*. В гене *MtCLE35* было выявлено наибольшее количество SNP, и значения D-критерия для него составляет -0,35, самое приближенное к нулю значение по сравнению с остальными генами, а также 2 несинонимичных на 3 синонимичные замены. Получается, что среди проанализированных «клубеньковых»генов люцерны, ген *MtCLE35* по всем показателям более вариабельный по сравнению с другими изученными генами *CLE*.

Однако среди проанализированных участков оказались позиции с выраженным положительным значением критериев (D>0). Так, на участки под номерами 8, 22 и 23, расположенных в межгенной области (Intergenic region (IGR)), действует балансирующий отбор, из которых участок под номером 8 имеет SNP с эффектом MODIFIER, что предполагает изменения вне областей кодирования. Известно, что IGR содержат функционально важные элементы, такие как промоторы и энхансеры. В частности, энхансерные последовательности ДНК могут активировать экспрессию дискретных наборов генов на большом расстоянии, а изменения в белках, связанных с энхансерами, могут влиять на клеточный фенотип. В контексте полученных результатов можно предположить, что генетические изменения в участках под номерами 8, 22 и 23 могут привести к расширению адаптивных возможностей растения без возникновения новых форм. С другой стороны, также можно предположить и внезапное сужение популяции растения – генетические изменения приводят к уменьшению редких аллелей и отсутствию новых адаптивных возможностей.

Отдельно отмечаются участки под номерами 24, 26-28 в IGR, значение D-критерия у которых нейтрально (D=0) или приближается к нулю. В таких участках наблюдается отсутствие доказательств отбора.

Таким образом, проведенный анализ показателей генетической изменчивости в локусах, охватывающих кластеры генов *CLE* *M. truncatula*, позволил установить влияние отбора на основании критериев TajimaD, ThetaPerBP и PiPerBP. В целом можно говорить о преобладании положительного отбора (D<0) над балансирующим отбором (D>0) и нейтральной эволюцией (D=0). Это может объясняться расширением популяции *M. truncatula* после недавнего прохождения «бутылочного горлышка», а также генетическими изменениями – SNP, которые могут влиять на функционирование генома растения.

*3.7.* *Поиск ассоциаций полиморфизма генов MtCLE с количеством образующихся клубеньков*

В рамках данной работы мы также предприняли попытку оценить, существуют ли возможные ассоциации SNP в геномных участках, охватывающих гены *MtCLE12, MtCLE13, MtCLE34* и *MtCLE35*, с количеством образующихся клубеньков.Поскольку известно, что продукты генов *MtCLE12, MtCLE13, MtCLE35* оказывают ингибирующий эффект на развитие симбиотических клубеньков (Mortier et al., 2010; Lebedeva et al., 2020; Mens et al., 2021), и ген *MtCLE34* также является потенциальным регулятором развития симбиотических клубеньков, мы предположили, что SNP, приводящие к изменению уровня активности данных генов, могут оказывать влияние на количество образующихся симбиотических клубеньков. Данные о количестве образующихся клубеньков у разных линий *M. truncatula* были взяты из работы Stanton-Geddes и соавторов (Stanton-Geddes et al., 2013), которые отдельно учитывали количество клубеньков в верхней части корневой системы (признак nodule\_above), что отражает более ранние события инокуляции, и количество клубеньков в нижней части корневой системы (nodule\_below), отражающее более поздние события инокуляции.

Прежде всего, нами была построена точечная диаграмма (XYScatter), отражающая распределение 262 линий *M. truncatula* из базы HapMap в соответствии с количеством у них клубеньков в верхней и нижней части корневой системы (рис. 23). На этой диаграмме мы отметили синим цветом линии, у которых имеется стоп-кодон в гене *MtCLE34* (SNP 39237092) (это линии HM038, HM058, HM101 (линия A17), HM256), чтобы первоначально оценить, насколько в целом отличается количество клубеньков у данных конкретных линий от количества клубеньков у других линий.

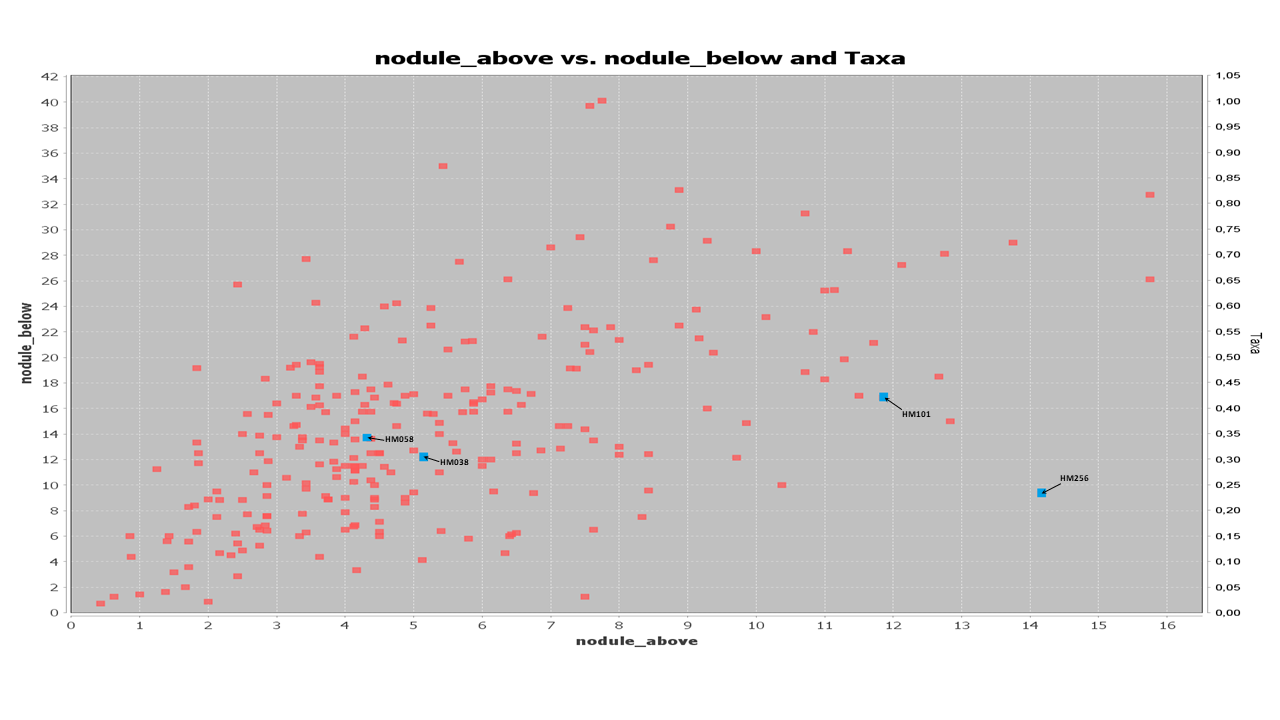


Рис. 23. Точечная диаграмма (XYScatter), отражающая распределение линий *M. truncatula* в соответствии с количеством клубеньков в верхней (nodule\_above) и нижней части (nodule\_below) корневой системы. Синим цветом отмечены линии, имеющие SNP 39237092 – стоп-кодон в гене *MtCLE34*.

Как видно из данного распределения, у двух линий, HM038 и HM058, количество клубеньков в верхней и нижней части корня находится в пределах значений признака, характерных для большинства других линий, тогда как точки, соответствующие двум другим линиям, HM101 (линия А17) и HM256, отклоняются от общего распределения – для этих двух линий характерны более высокие значения признака «nodule\_above». Таким образом, на основании данного распределения, невозможно однозначно определить, связана ли мутация в гене *MtCLE34,* приводящая к потере функции *MtCLE34* вследствие наличия преждевременного стоп-кодона, к изменению количества клубеньков.

Для выявления генов, ответственных за встречающиеся в природе фенотипические вариации, часто используют полногеномный поиск ассоциаций (Genome-wide association study, GWAS). Этот подход позволяет выявить участки генома, полиморфизм в которых демонстрирует ассоциации с проявлением изучаемого признака. В нашей работе с помощью GWAS в программе Tassel мы провели поиск ассоциаций наличия SNP в участках генома, включающих кластеры генов *CLE* (для *MtCLE12* и *MtCLE13* - в хромосоме 4, *MtCLE34* и *MtCLE35* в хромосоме 2), с количеством образующихся клубеньков на корнях разных линий *M. truncatula*. Данные по SNP были взяты из базы данных HapMap Medicago (http://www.medicagohapmap.org/). Нами были использованы модели для поиска ассоциаций, доступные в программе Tassel: Multidimensional Scaling (MDS), General Linear Model (GLM) и Mixed Minear Model (MLM) (Principal Component Analysis (PCA) + K).

В результате проведенного анализа были построены графики Manhattan, отражающие вероятностное распределение участков кластеров генов *CLE* с SNP на хромосомах 2 и 4, и их связь с фенотипическими характеристиками – количество клубеньков в верхней части корневой системы (nodule\_above) и в нижней части корневой системы (nodule\_below). График Manhattan представляет собой точечный график, который суммирует результаты GWAS (приложение 1-2), где по оси X указана геномная позиция SNP, а по оси Y – отрицательный логарифм значения *P* (вероятность наличия ассоциации SNP в данной позиции с изучаемым признаком). Таким образом, более высокие пики на графике, указывают на то, что SNP в данном конкретном положении имеет более высокую ассоциацию с изучаемым признаком (в нашем случае, количество образующихся клубеньков в нижней или верхней части корневой системы). В частности, в позиции 39236000 и соседних точках на хромосоме 2 можно отметить наличие более высоких пиков, что позволяет предположить, что SNP в данном участке генома имеют более сильную связь с количеством образующихся клубеньков в верхней части растения (приложение 2). Эта область генома расположена перед кодирующей последовательностью гена *MtCLE34*, и возможно, включает в себя регуляторные элементы, влияющие на его уровень экспрессии.

Гены *MtCLE12* и *MtCLE13* на графиках Manhattan представлены позициями 30793500 и 30799500-30800500, и в этих участках не было выявлено SNP, демонстрирующих высокие показатели степени ассоциации с количеством образующихся клубеньков (приложение 3-4).

Таким образом, с помощью подхода GWAS нами был выявлен участок, соответствующий позиции 39236000 и соседним точкам в хромосоме 2, который расположен перед кодирующей последовательностью гена *MtCLE34* и в пределах которого наблюдаются SNP, демонстрирующие более высокую степень ассоциации с количеством симбиотических клубеньков, по сравнению с соседними участками генома. Для SNP, расположенных в области других проанализированных генов *MtCLE*, таких ассоциаций выявлено не было.

*3.8. Изучение функциональной значимости гена MtCLE34 в регуляции развития симбиотических клубеньков у M. truncatula*

Для изучения роли гена *MtCLE34* в контроле развития клубеньков у *M. truncatula* нами была использована конструкция, включающая функциональную кодирующую последовательность гена *MtCLE34*, полученную от линии R108, под контролем промотора мозаики цветной капусты 35S, которая была получена ранее в нашей лаборатории. С использованием данной конструкции мы провели трансформацию проростков *M. truncatula* линии A17 с помощью *A. rhizogenes* с целью оценки эффекта сверхэкспрессии *MtCLE34* на развитие симбиотических клубеньков. В качестве контроля была использована конструкция для сверхэкспрессии гена GUS (бета-глюкуронидазы) 35S:GUS.

Трансгенные корни отбирали по флуоресценции репортерного белка GFP. Мы смотрели как трансгенные (GFP+), так и не трансгенные корни (GFP-). Это проводилось для того, чтобы оценить возможный системный эффект. Например, для *MtCLE35* и для других «клубеньковых» *CLE* показан системный эффект, т.е. наблюдается подавление развития клубеньков как на трансгенных корнях трансформированных композитных растений со сверхэкспрессией гена *CLE*, так и на нетрансгенных корнях композитных растений, несущих также и трансгенные корни (Mortier et al., 2010; Lebedeva et al., 2020).

Для подтверждения наличия сверхэкспрессии гена *MtCLE34* мы выделили РНК из трансгенных корней, трансформированных конструкцией 35S:MtCLE34 (в двух повторностях: CLE34-1 и CLE34-2) и контрольных корней, трансформированных конструкцией 35S:GUS (в двух повторностях: GUS-1 и GUS-2), провели обратную транскрипцию и получили пробы кДНК для проведения ПЦР в реальном времени с целью оценки наличия сверхэкспрессии гена *MtCLE34* в трансгенных корнях, трансформированных конструкцией для сверхэкспрессии гена *MtCLE34*. По данным ПЦР в реальном времени, уровень экспрессии *MtCLE34* в трансгенных корнях, несущих конструкцию 35S:MtCLE34, примерно в 200 раз превышает уровень экспрессии *MtCLE34* в контрольных корнях, несущих конструкцию 35S:GUS (рис. 24), что подтверждает наличие сверхэкспрессии этого гена.

Относительные уровни экспрессии гена *MtCLE34*

Рис. 24. График уровней экспрессии гена *MtCLE34* в образцах GUS-1 (контроль) и CLE34-1 (сверхэкспрессия гена *MtCLE34*).

На трансгенных корнях со сверхэкспрессией гена *MtCLE34* развивались симбиотические клубеньки (рис. 25, А-Г).

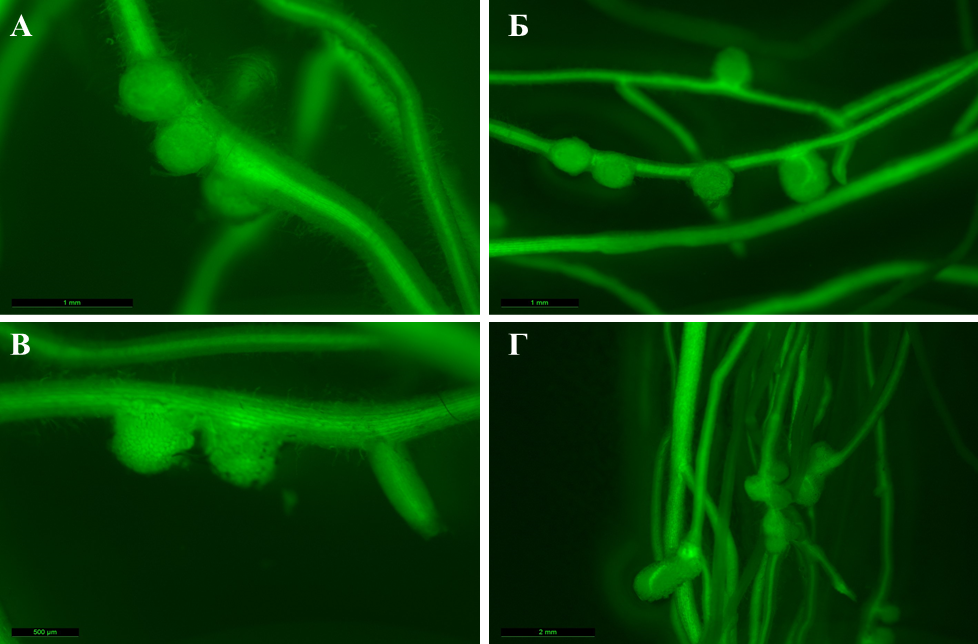


Рис. 25. Фотографии клубеньков на трансгенных корнях со сверхэкспрессией гена *MtCLE34*, отобранных по флуоресценции репортерного белка GFP.

Мы не выявили статистически значимых отличий в количестве клубеньков на трансгенных корнях со сверхэкспрессией *MtCLE34* и на контрольных корнях со сверхэкспрессией GUS (рис. 26). На основании этого мы можем сделать вывод о том, что сверхэкспрессия гена *MtCLE34* не оказывает существенного влияния на количество закладывающихся симбиотических клубеньков у *M. truncatula*.

Стоит отметить, что при сверхэкспрессии гена *MtCLE34* в некоторых случаях на трансгенных корнях наблюдались близко расположенные друг к другу кластеры клубеньков (рис. 27, В, Г), тогда как у контрольных растений со сверхэкспрессией *GUS* клубеньки в целом были более равномерно распределены вдоль корня (рис. 27, А, Б). Однако детальной оценки пространственного распределения клубеньков вдоль корня при сверхэкспрессии *MtCLE34* мы не проводили.

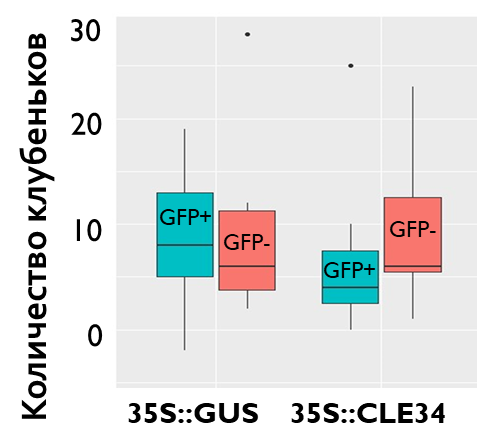


Рис. 26. Количество клубеньков на трансгенных корнях при сверхэкспрессии гена *MtCLE34* (CLE34\_OE) и в контроле (GUS\_OE).

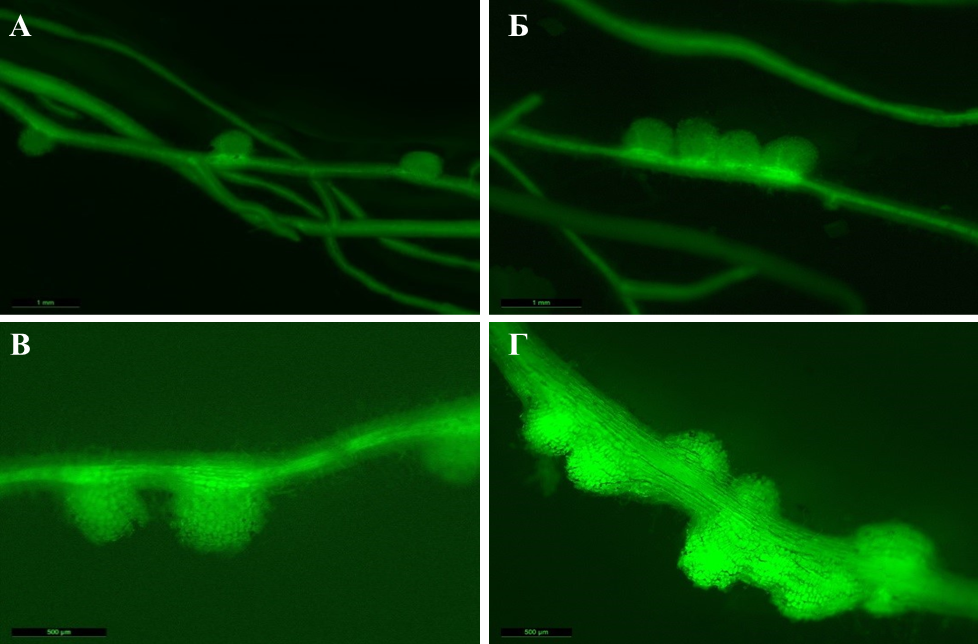


Рис. 27. Фотографии клубеньков на контрольных трансгенных корнях со сверхэкспрессией гена GUS (A, Б) и на трансгенных корнях со сверхэкспрессией гена *MtCLE34* (В, Г).

Таким образом, в отличие от его близких гомологов, генов *MtCLE35, MtCLE12* и *MtCLE13,* ген *MtCLE34* не является ингибитором клубенькообразования у *M. truncatula*, и, по-видимому, не является компонентом AON.

**ЗАКЛЮЧЕНИЕ**

Таким образом, в настоящей работе проводилось изучение полиморфизма *Medicago truncatula* по генам, кодирующим пептиды CLE – охарактеризованные и потенциальные участники авторегуляции клубенькообразования. В качестве объектов исследования были использованы 2 линии бобового растения *M. truncatula* – A17 и R108. Для исследования полиморфизма были выбраны следующие гены: гомологи *MtCLE12* и *MtCLE13* - функционально характеризованные гены, для которых ранее было показано участие в AON; *MtCLE35,* функция в клубенькообразовании для которого была показана совсем недавно и ген *MtCLE34,* – роль которого в клубенькообразовании оставалась неизученной. Гены *MtCLE12* и *MtCLE13* расположены рядом на хромосоме4, тогда как гены *MtCLE34* и *MtCLE35* занимают соседние позиции на хромосоме 2. При этом расположенные рядом пары генов *CLE* демонстрируют схожий паттерн экспрессии: экспрессия генов *MtCLE12* и *MtCLE13* активируется при инокуляции ризобиями и в целом специфична для клубеньков, тогда как расположенные рядом гены *MtCLE34* и *MtCLE35* активируются при клубенькообразовании, при действии нитрата, а также их экспрессия наблюдается в корневой системе неинокулированных растений. Характер экспрессии генов *MtCLE34* и *MtCLE35* позволяет предположить, что их роль в растениях может быть более широкой и может не ограничиваться регуляцией клубенькообразования.

В результате проведенного поиска вариабельных позиций в последовательностях генов *CLE* у линий *M. truncatula*, доступных в базе данных HapMap, были обнаружены существующие SNP, из которых в гене *MtCLE13* – только две синонимичные замены, что может свидетельствовать о консервативности этого гена и значимости его функций для жизни растения, а также о возможном существовании направленного отбора в эволюции. У генов *MtCLE12, MtCLE34* и *MtCLE35* - синонимичные и несинонимичные замены. При этом в гене *MtCLE35* было выявлено наибольшее количество SNP и значения D-критерия для него составляет -0,35, самое приближенное к нулю значение, по сравнению с остальными генами, а также 2 несинонимичных на 3 синонимичные замены. Среди проанализированных «клубеньковых» генов люцерны, ген *MtCLE35* по всем показателям более вариабельный по сравнению с другими изученными генами *CLE.*

Выявление стоп-кодона в гене *MtCLE34* у лабораторной линии A17, а также у трех других линий из 262 линий, данные o SNP для которых представлены в базе HapMap, свидетельствует в пользу того, что функция этого гена не является жизненно важной для растений. Проведенная трансформация линии A17 растения *M. truncatula* с целью оценки эффекта сверхэкспрессии функционального гена *MtCLE34*, полученного от линии R108, на клубенькообразование показала, что сверхэкспрессия гена *MtCLE34* не оказывает эффекта на количество образующихся клубеньков, что отличает этот ген от родственных ему «клубеньковых» генов *MtCLE13, MtCLE12* и *MtCLE35 –* ингибиторов клубенькообразования. Таким образом, пептид MtCLE34*,* во-видимому, не является участником AON. Представляет особый интерес осуществить с помощью экспериментальных подходов поиск возможных причин того, почему в ходе эволюции этот близкий гомолог индуцируемых нитратом генов *CLE* сои утратил способность подавлять развитие симбиотических клубеньков у люцерны.

**Выводы**

1. Занимающие соседние позиции на хромосомах пары проанализированных генов *CLE* люцерны *Medicago truncatula* демонстрируют схожий паттерн экспрессии: экспрессия генов *MtCLE12* и *MtCLE13*, расположенных рядом на хромосоме № 2, является специфичной для инокулированных ризобиями корней и симбиотических клубеньков, тогда как экспрессия расположенных рядом на хромосоме № 4 генов *MtCLE34* и *MtCLE35* наблюдается в клубеньках, в корневой системе растений в отсутствие ризобий, а также активируется при обработке нитратом.
2. Среди генов *MtCLE12, MtCLE13, MtCLE34*, *MtCLE35*, наиболее вариабельным по всем проанализированным критериям является ген *MtCLE35*.
3. В гене *MtCLE34* у четырех из 262 линий люцерны из коллекции HapMap, в том числе у широко используемой лабораторной линии A17, присутствует стоп-кодон, наличие которого должно приводить к синтезу нефункционального белкового продукта.
4. Сверхэкспрессия функциональной копии гена *MtCLE34,* полученной от линии R108, не содержащей стоп-кодона, не оказывает эффекта на численность образующихся симбиотических клубеньков.

**СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ**

1. Белозерова М.Ю., Ткаченко А.А., Додуева И.Е., Жуков В.А. Анализ полиморфизма генов, кодирующих CLE-пептиды и их рецепторы, у гороха посевного (Pisum sativum L.). Школа-конференция молодых ученых «Молекулярно-генетические и клеточные аспекты растительно-микробных взаимодействий»: материалы. Санкт-Петербург-Пушкин. 1-3 декабря 2017. С. 11-12.
2. Ганчева М. С., Маловичко Ю. В., Полюшкевич Л. О., Додуева И. Е., Лутова Л. А. Пептидные гормоны растений. Журнал «Физиология растений». 2019. С. 83-103.
3. Лебедева М. А., Яшенкова Я. С., Додуева И. Е., Лутова Л. А. Молекулярный диалог корня и побега с участием регуляторных пептидов и его роль в системном контроле развития растений. Журнал «ФИЗИОЛОГИЯ РАСТЕНИЙ». 2020. 67(6). С. 1–21. doi:10.31857/S0015330320060111
4. Лутова Л. А., Додуева И. Е., Лебедева М. А., Творогова В. Е. Транскрипционные факторы в генетике развития и эволюции высших растений. Журнал «ГЕНЕТИКА». 2015. 51(5). С. 539-557. doi:10.7868/S001667581503008X
5. Adhikari L., Lindstrom O. M., Markham J., Missaoui A. M. Dissecting Key Adaptation Traits in the Polyploid Perennial Medicago sativa Using GBS-SNP Mapping. Frontiers in Plant Science. 2018. 9. P. 934. doi:10.3389/fpls.2018.00934
6. Araya T., Miyamoto M., Wibowo J., Suzuki A., Kojima S., Tsuchiya Y.N., … Takahashi H. CLE-CLAVATA1 peptide-receptor signaling module regulates the expansion of plant root systems in a nitrogen-dependent manner. Proceedings of the National Academy of Sciences. 2014. 111(5). P. 2029–2034. doi:10.1073/pnas.1319953111
7. Balakrishnan C.N., Edwards S.V. Nucleotide Variation, Linkage Disequilibrium and Founder-Facilitated Speciation in Wild Populations of the Zebra Finch (Taeniopygia guttata). Genetics. 2008. 181(2). P. 645–660. doi:10.1534/genetics.108.094250
8. Betsuyaku S., Sawa S., Yamada M. The Function of the CLE Peptides in Plant Development and Plant-Microbe Interactions. The Arabidopsis Book. 2011. 9. e0149. doi:10.1199/tab.0149
9. Bonhomme M., Fariello M. I., Navier H., Hajri A., Badis Y., Miteul H., … Pilet-Nayel, M.-L. A local score approach improves GWAS resolution and detects minor QTL: application to Medicago truncatula quantitative disease resistance to multiple Aphanomyces euteiches isolates. Heredity. 2019. 123(4). P. 517-531. doi:10.1038/s41437-019-0235-x
10. Boschiero C., Dai X., Lundquist PK., Roy S., de Bang T., Zhang S., Zhuang Z., Torres-Jerez I., Udvardi MK., Scheible WR., Zhao PX. MtSSPdb: the Medicago truncatula Small Secreted Peptide Database. Plant Physiol. 2020. 183(1). P. 399-413. doi:10.1104/pp.19.01088
11. Bradbury P.J., Zhang Z., Kroon D.E., Casstevens T.M., Ramdoss Y., Buckler E.S. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics. 2007. 23(19). P. 2633-2635. doi:10.1093/bioinformatics/btm308
12. Chen S., Lang P., Chronis D., Zhang S., De Jong W.S., Mitchum M.G., Wang X. In planta processing and glycosylation of a nematode CLAVATA3/ENDOSPERM SURROUNDING REGION-like effector and its interaction with a host CLAVATA2-like receptor to promote parasitism. Plant Physiol. 2015. P. 262-72. doi:10.1104/pp.114.251637
13. Curtin S.J., Xiong Y., Michno J.-M., Campbell B.W., Stec A.O., Čermák T., … Stupar R.M. CRISPR/Cas9 and TALENs generate heritable mutations for genes involved in small RNA processing of Glycine max and Medicago truncatula. Plant Biotechnology Journal. 2017. 16(6). P. 1125–1137. doi:10.1111/pbi.12857
14. de Bang T.C., Lay K.S., Scheible W.-R., Takahashi H. Small peptide signaling pathways modulating macronutrient utilization in plants. Current Opinion in Plant Biology. 2017. 39. P. 31–39. doi:10.1016/j.pbi.2017.05.005
15. Delph L.F., Kelly J.K. On the importance of balancing selection in plants. New Phytologist. 2013. 201(1). P. 45–56. doi:10.1111/nph.12441
16. DiGennaro P., Grienenberger E., Dao T.Q., Jun J.H., Fletcher J.C. Peptide signaling molecules CLE5 and CLE6 affect Arabidopsis leaf shape downstream of leaf patterning transcription factors and auxin. Plant Direct. 2018. 2(12). e00103. doi:10.1002/pld3.103
17. Dong W., Wang Y., Takahashi H. CLE-CLAVATA1 Signaling Pathway Modulates Lateral Root Development under Sulfur Deficiency. Plants. 2019. 8(4). P. 103. doi:10.3390/plants8040103
18. Doo M., Kim Y. Obesity: Interactions of Genome and Nutrients Intake. Preventive Nutrition and Food Science. 2015. 20(1). P. 1–7. doi:10.3746/pnf.2015.20.1.1
19. Dybowski R., Restif O., Price D.J., Mastroeni P. Inferring within-host bottleneck size: A Bayesian approach. Journal of Theoretical Biology. 2017. 435. P. 218–228. doi:10.1016/j.jtbi.2017.09.011
20. Endo S., Betsuyaku S., Fukuda H. Endogenous peptide ligand–receptor systems for diverse signaling networks in plants. Curr. Opin. Plant Biol. 2014. 21. P. 140–146. doi:10.1016/j.pbi.2014.07.011
21. Fahraeus G. The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. J. Gen. Microbiol. 1957. 16. P. 374–381. doi:10.1099/00221287-16-2-374
22. Fay J.C., Wu C.I. Hitchhiking under positive Darwinian selection. Genetics. 2000. 155(3). P. 1405-1413.
23. Ferguson B.J., Mens C., Hastwell A.H., Zhang M., Su H., Jones C.H., … Gresshoff P.M. Legume nodulation: The host controls the party. Plant, Cell & Environment. 2018. 42(1). P. 41-51. doi:10.1111/pce.13348
24. Fletcher J.C. Recent Advances in Arabidopsis CLE Peptide Signaling. Trends in Plant Science. 2020. S1360-1385(20). P. 30129-1. doi:10.1016/j.tplants.2020.04.014
25. Fu Y.X., Li W.H. Statistical Testsof Neutrality of Mutations. Centerfor Demographic and Population Genetics, University of Texas, Houston, Texas. 1992. P. 693-705.
26. Futuyma D.J., Kirkpatrick M. Evolution. 4th edition. Oxford: Sinauer Associates. 2017.
27. Gautrat P., Laffont C., Frugier F. Compact Root Architecture 2 Promotes Root Competence for Nodulation through the miR2111 Systemic Effector. Current Biology. 2020. 30. P. 1–7. doi:10.1016/j.cub.2020.01.084
28. Guo X., Wang J., Gardner M., Fukuda H., Kondo Y., Etchells J.P., Wang X., Mitchum M.G. Identification of cyst nematode B-type CLE peptides and modulation of the vascular stem cell pathway for feeding cell formation, PLoS Pathog. 2017. 13. P. 1–19. doi:10.1504/IJICT.2017.085458
29. Guo Y., Ni, J., Denver, R., Wang, X., & Clark, S. E. Mechanisms of Molecular Mimicry of Plant CLE Peptide Ligands by the Parasitic Nematode Globodera rostochiensis. Plant Physiology. 2011. 157(1). P. 476–484. doi:10.1104/pp.111.180554
30. Harrisson K.A., Pavlova A., Telonis-Scott M., Sunnucks. Using genomics to characterize evolutionary potential for conservation of wild populations. Evol Appl. 2014. P. 1008-1025. doi:10.1111/eva.12149
31. Hastwell A.H., de Bang T.C., Gresshoff P.M., Ferguson B.J. CLE peptide-encoding gene families in Medicago truncatula and Lotus japonicus, compared with those of soybean, common bean and Arabidopsis. Scientific Reports. 2017. 7(1). P. 9384. doi:10.1038/s41598-017-09296-w
32. Hastwell A.H., Gresshoff P.M., Ferguson B.J. Genome-wide annotation and characterization of clavata/esr (cle) peptide hormones of soybean (Glycine max) and common bean (Phaseolus vulgaris), and their orthologues of Arabidopsis thaliana. J. Exp. Bot. 2015. 66. P. 5271-5287. doi:10.1093/jxb/erv351
33. Hirakawa Y., Kondo Y., Fukuda H. TDIF peptide signaling regulates vascular stem cell proliferation via the WOX4 homeobox gene in Arabidopsis. Plant cell. 2010. P. 2618-2629. doi:10.1105/tpc.110.076083
34. Kang Y., Sakuroglu M., Krom N., Stanton-Geddes J., Wang M., Lee Y-C., Young N.D., Udvardi M. Genome-wide assocation of drought-related and biomass traits with HapMap SNPs in Medicago truncatula. Plant Cell and Environment. 2015. 38. P. 1997-2011. doi: 10.1111/pce.12520
35. Karlo M., Boschiero C., Landerslev K.G., et al. The CLE53-SUNN genetic pathway negatively regulates arbuscular mycorrhiza root colonization in Medicago truncatula. J Exp Bot. 2020. 71(16). P. 4972-4984. doi:10.1093/jxb/eraa193
36. Kingsolver J.G., Diamond S.E. Phenotypic selection in natural populations: What limits directional selection? Am. Nat. 2011. 177. P. 346–357. doi:10.1086/658341
37. Kinoshita A., Betsuyaku S., Osakabe Y., Mizuno S., Nagawa S., Stahl Y., Simon R., Yamaguchi-Shinozaki K., Fukuda H., Sawa S. Rpk2 is an essential receptor-like kinase that transmits the clv3 signal in Arabidopsis. Development. 2010. 137. P. 3911–3920. doi:10.1242/dev.048199
38. Kumar S., Stecher G., Li M., Knyaz C., Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution. 2018. 35. P. 1547-1549. doi:10.1093/molbev/msy096
39. Laffont C., Ivanovici A., Gautrat P. et al. The NIN transcription factor coordinates CEP and CLE signaling peptides that regulate nodulation antagonistically. Nat Commun. 2020. 11. 3167. P. 1-13. doi:10.1038/s41467-020-16968-1
40. Le Marquer M., Bécard G., Frei dit Frey N. Arbuscular mycorrhizal fungi possess a CLAVATA 3/Endosperm surrounding region‐related gene that positively regulates symbiosis. New Phytologist. 2018. 222(2). P. 1030-1042. doi:10.1111/nph.15643
41. Lebedeva M., Azarakhsh M., Yashenkova Y., Lutova L. Nitrate-Induced CLE Peptide Systemically Inhibits Nodulation in Medicago truncatula. Plants (Basel). 2020. 28;9(11). P. 1456. doi:10.3390/plants9111456
42. Li G., Wang, B., Tian, Q., Wang, T., & Zhang, W.-H. Medicago truncatula ecotypes A17 and R108 differed in their response to iron deficiency. Journal of Plant Physiology. 2014. 171(8). P. 639–647. doi:10.1016/j.jplph.2013.12.018
43. Li Z., Liu D., Xia Y., Li Z., Niu N., Ma S., Wang J., Song Y., Zhang G. Identification and Functional Analysis of the CLAVATA3/EMBRYO SURROUNDING REGION (CLE) Gene Family in Wheat. Int. J. Mol. Sci. 2019. 20. P. 4319. doi:10.3390/ijms20174319
44. Lieberman B.S., Dudgeon S. An evaluation of stabilizing selection as a mechanism for stasis. Palaeogeog. Lapaeoclimat. Palaeoecol. 1996. 127. P. 229–238.
45. Lim C.W., Lee Y.W., Lee S.C., Hwang C.H. Nitrate inhibits soybean nodulation by regulating expression of CLE genes // Plant Science. 2014. 229. P. 1. doi:10.1016/j.plantsci.2014.08.014
46. Limpens E. RNA interference in Agrobacterium rhizogenes-transformed roots of Arabidopsis and Medicago truncatula. Journal of Experimental Botany. 2004. 55(399). P. 983–992. doi:10.1093/jxb/erh122
47. Liu P. Pricing Strategies of a Three-Stage Supply Chain: A New Research in the Big Data Era. Discrete Dynamics in Nature and Society. 2017. P. 1–16. doi:10.1155/2017/9024712
48. Loewe L. Negative selection. Nature Education. 2008. 1(1). P. 59.
49. Lu G., Moriyama E. Vector NTI, a balanced all-in-one sequence analysis suite. Briefings in Bioinformatics. 2004. 5(4). P. 378–388. doi:10.1093/bib/5.4.378
50. Luo S., Sun Y., Zhou X., Zhu T., Zhu L., Arfan M., … Lin H. Medicago truncatula genotypes Jemalong A17 and R108 show contrasting variations under drought stress. Plant Physiology and Biochemistry. 2016. 109. P. 190–198. doi:10.1016/j.plaphy.2016.09.019
51. Ma L.C., Wang Y.R., Liu W.X., Liu Z.P. Expression analysis of seed-specific genes in four angiosperm species with an emphasis on the unconserved expression patterns of homologous genes. Seed Sci. Res. 2013. 23. P. 223–231. doi:10.1017/S0960258513000305
52. Mens C., Hastwell A.H., Su H., Gresshoff P.M., Mathesius U., Ferguson B.J. Characterisation of Medicago truncatula CLE34 and CLE35 in nitrate and rhizobia regulation of nodulation. New Phytol. 2021. 229. P. 2525-2534. doi:10.1111/nph.17010
53. Miyawaki K., Tabata R., Sawa S. Curr Opin Evolutionarily conserved CLE peptide signaling in plant development, symbiosis, and parasitism. Plant Biol. 2013. P. 598-606. doi:10.1016/j.pbi.2013.08.008
54. Mortier V., Den Herder G., Whitford R., Van de Velde W., Rombauts S., D'Haeseleer K., Holsters M., Goormachtig S. CLE peptides control Medicago truncatula nodulation locally and systemically. Plant Physiol. 2010. P. 222-37. doi:10.1104/pp.110.153718
55. Mortier V., De Wever E., Vuylsteke M., Holsters M., Goormachtig S. Nodule numbers are governed by interaction between CLE peptides and cytokinin signaling. The Plant Journal. 2012. 70(3). P. 367–376. doi:10.1111/j.1365-313x.2011.04881.x
56. Mortier V., Fenta B. A., Kunert K., Holsters M, Goormachtig S. Identification of putative CLE peptide receptors involved in determinate nodulation on soybean. Plant Signaling and Behavior. 2011. 6(7). P. 1019-1023. doi:10.4161/psb.6.7.15575
57. Nishida H., Ito M., Miura K., Kawaguchi M., Suzaki, T. Autoregulation of nodulation pathway is dispensable for nitrate-induced control of rhizobial infection. Plant Signaling & Behavior. 2020. 26. P. 1733814. doi:10.1080/15592324.2020.1733814
58. Nishida H., Suzaki T. Two Negative Regulatory Systems of Root Nodule Symbiosis: How Are Symbiotic Benefits and Costs Balanced? Plant and Cell Physiology. 2018. 59(9). P. 1733–1738. doi:10.1093/pcp/pcy102
59. Nishida H., Tanaka S., Handa Y., Ito M., Sakamoto Y., Matsunaga S., … Suzaki T. A NIN-LIKE PROTEIN mediates nitrate-induced control of root nodule symbiosis in Lotus japonicus. Nature Communications. 2018. 9(1). P. 499. doi:10.1038/s41467-018-02831-x
60. Oelkers K., Goffard N., Weiller G. F., Gresshoff P. M., Mathesius U., Frickey T. Bioinformatic analysis of the CLE signaling peptide family. BMC Plant Biology. 2008. 8(1). P. 1. doi:10.1186/1471-2229-8-1
61. Okamoto S., Ohnishi E., Sato S., Takahashi H., Nakazono M., Tabata S., Kawaguchi M. Nod Factor/Nitrate-Induced CLE Genes that Drive HAR1-Mediated Systemic Regulation of Nodulation. Plant and Cell Physiology. 2009. 50(1). P. 67–77. doi:10.1093/pcp/pcn194
62. Okonechnikov K., Golosova O., Fursov M., et al. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics. 2012. 28. P. 1166-1167. doi:10.1093/bioinformatics/bts091
63. Reid D.E., Ferguson B.J., Gresshoff P.M. Inoculationand nitrate-induced CLE peptides of soybean control NARK-dependent nodule formation // Mol. Plant Microbe Interact. 2011. 24. P. 606. doi:10.1094/MPMI-09-10-0207
64. Rueffler C., Van Dooren Tom J.M., Olof Leimar, Abrams P. A. Disruptive selection and then what? Trends in Ecology and evolution. 2006. P. 238-245. doi:10.1016/j.tree.2006.03.003
65. Salisbury B.A., Pungliya M., Choi J.Y., Jiang R., Sun X.J., Stephens J.C. SNP and haplotype variation in the human genome. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 2003. 526(1-2). P. 53–61. doi:10.1016/s0027-5107(03)00014-9
66. Sarkar A.K., Luijten M., Miyashima S., Lenhard M., Hashimoto T., Nakajima K., Scheres B., Heidstra R., Laux T. Conserved factors regulate signalling in Arabidopsis thaliana shoot and root stem cell organizers. Nature. 2007. 446. P. 811–814. doi:10.1038/nature05703
67. Schoof H., Lenhard M., Haecker A., Mayer K.F., Jürgens G., Laux T. The stem cell population of Arabidopsis shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. Cell. 2000. V. 100. P. 635–644. doi:10.1016/s0092-8674(00)80700-x
68. Somssich M., Je Il B., Simon R., Jackson D. CLAVATA-WUSCHEL signaling in the shoot meristem. Development 143. 2016. P. 3238–3248. doi:10.1242/dev.133645
69. Stahl Y., Wink R. H., Ingram G. C., Simon R. A Signaling Module Controlling the Stem Cell Niche in Arabidopsis Root Meristems. Current Biology. 2009. 19(11). P. 909–914. doi:10.1016/j.cub.2009.03.060
70. Stanton-Geddes J., Paape T., Epstein B., Briskine R., Yoder J., Mudge J., … Tiffin P. Candidate Genes and Genetic Architecture of Symbiotic and Agronomic Traits Revealed by Whole-Genome, Sequence-Based Association Genetics in Medicago truncatula. PLoS ONE. 2013. 8(5). e65688. doi:10.1371/journal.pone.0065688
71. Strabala T.J., O’Donnell P.J., Anne-Marie S., Charles A.D., E Jane M., Natalie N., Nieuwenhuizen N.J., Quinn B.D., Foote H.C.C., Hudson K.R. Gain-of-function phenotypes of many clavata3/esr genes, including four new family members, correlate with tandem variations in the conserved clavata3/esr domain. Plant Physiol. 2006. 140. P. 1331–1344. doi:10.1104/pp.105.075515
72. Studer A., Zhao Q., Ross-Ibarra J., Doebley J. Identification of a functional transposon insertion in the maize domestication gene tb1. Nat. Genet. 2011. 43. P. 1160–1163. doi:10.1038/ng.942
73. Suzaki T., Nishida H. Autoregulation of Legume Nodulation by Sophisticated Transcriptional Regulatory Networks. Molecular Plant. 2019. 12(9). P. 1179-1181. doi:10.1016/j.molp.2019.07.008
74. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989. P. 585–595.
75. Tamura K., Nei M., Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proceedings of the National Academy of Sciences. 2004. 101(30). P. 11030–11035. doi:10.1073/pnas.0404206101
76. Tsikou D., Yan Z., Holt D. B., Abel N. B., Reid D. E., Madsen L. H., … Markmann K. Systemic control of legume susceptibility to rhizobial infection by a mobile microRNA. Science. 2018. 362(6411). P. 233-236. doi:10.1126/science.aat6907
77. Villanea Fernando A., Safi Kristin N., Busch Jeremiah W. A general model of negative frequency dependent selection explains global patterns of human ABO Polymorphism. Plos one. 2015. 10(5). e0125003. doi:10.1371/journal.pone.0125003
78. Wang J., Kucukoglu M., Zhang L., Chen P., Decker D., Nilsson O., Jones B., Sandberg G., Zheng B. The Arabidopsis LRR-RLK, PXC, is a regulator of secondary wall formation correlated with the TDIF-PXY/TDR-WOX4 signaling pathway. Plant biology. 2013. P. 13. doi:10.1186/1471-2229-13-94
79. Wang L., Sun Z., Su C., Wang Y., Yan Q., Chen J., … Li X. A GmNINa-miR172c-NNC1 Regulatory Network Coordinates the Nodulation and Autoregulation of Nodulation Pathways in Soybean. Molecular Plant. 2019. 12(9). P. 1211-1226. doi:10.1016/j.molp.2019.06.002
80. Watterson G.A. On the number of segregating sites in genetical models without recombination. Theoretical Population Biology. 1975. 7(2). P. 256–276. doi:10.1016/0040-5809(75)90020-9
81. Whitewoods C.D. Evolution of CLE peptide signalling. Semin Cell Dev Biol. 2020. 109. P. 12-19. doi:10.1016/j.semcdb.2020.04.022
82. Wray G.A. Genomics and the Evolution of Phenotypic Traits. Annual Review of Ecology, Evolution, and Systematics. 2013. 44(1). P. 51–72. doi:10.1146/annurev-ecolsys-110512-135828
83. Yadav R.K., Perales M., Gruel J. et al. WUSCHEL protein movement mediates stem cell homeostasis in the Arabidopsis shoot apex. Genes and Development. 2011. 25. P. 2025–2030. doi:10.7868/S001667581503008X
84. Yamaguchi Y.L., Ishida T., Sawa S. CLE peptides and their signaling pathways in plant development. J. Exp. Bot. 2016. 67. P. 4813–4826. doi:10.1093/jxb/erw208
85. Yu L.-X., Zheng P., Bhamidimarri S., Liu X.-P., Main D. The Impact of Genotyping-by-Sequencing Pipelines on SNP Discovery and Identification of Markers Associated with Verticillium Wilt Resistance in Autotetraploid Alfalfa (Medicago sativa L.). Frontiers in Plant Science. 2017. 8. P. 89. doi:10.3389/fpls.2017.00089
86. Zhang H., Lin X., Han Z., Qu L.-J., Chai J. Crystal structure of PXY-TDIF complex reveals a conserved recognition mechanism among CLE peptide-receptor pairs. Cell Research. 2016. 26(5). P. 543–555. doi:10.1038/cr.2016.45
87. Zheng Y., Xu F., Li Q., Wang G., Liu N., Gong Y., Li L., Chen ZH., Xu S. QTL Mapping Combined With Bulked Segregant Analysis Identify SNP Markers Linked to Leaf Shape Traits in Pisum sativum Using SLAF Sequencing. Front Genet. 2018. 5(9). P. 615. doi:10.3389/fgene.2018.00615

**ПРИЛОЖЕНИЯ**

**Приложение 1**

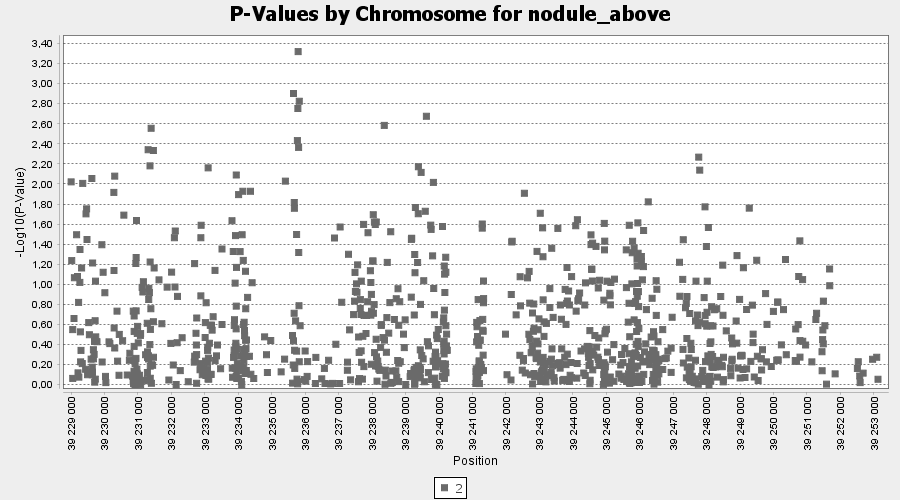


Рисунок. График Manhattan для *MtCLE34* и *MtCLE35* в хромосоме 2 по nodule\_above.

**Приложение 2**

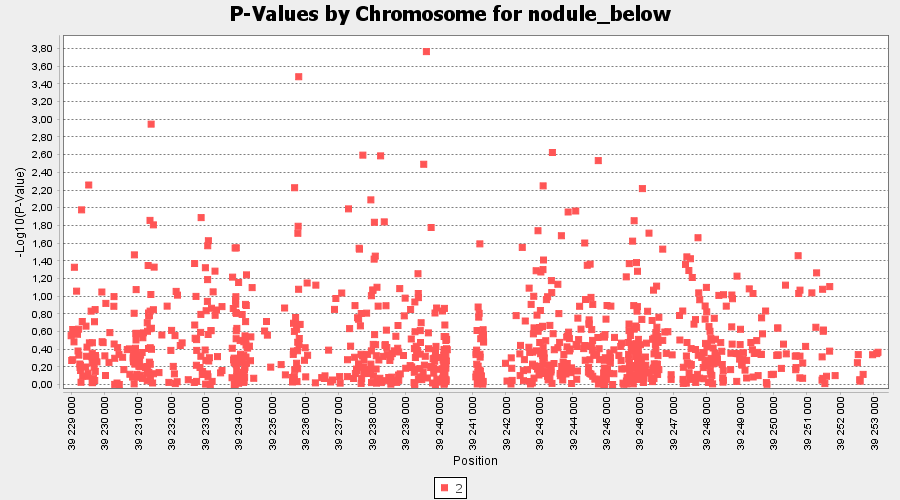


Рисунок. График Manhattan для участка хромосомы 2, включающего гены *MtCLE34* и *MtCLE35,* в хромосоме 2 по nodule\_below.

**Приложение 3**



Рисунок. График Manhattan для участка хромосомы 4, включающего гены *MtCLE12* и *MtCLE13* по признаку «количество клубеньков в верхней части корневой системы».

**Приложение 4**

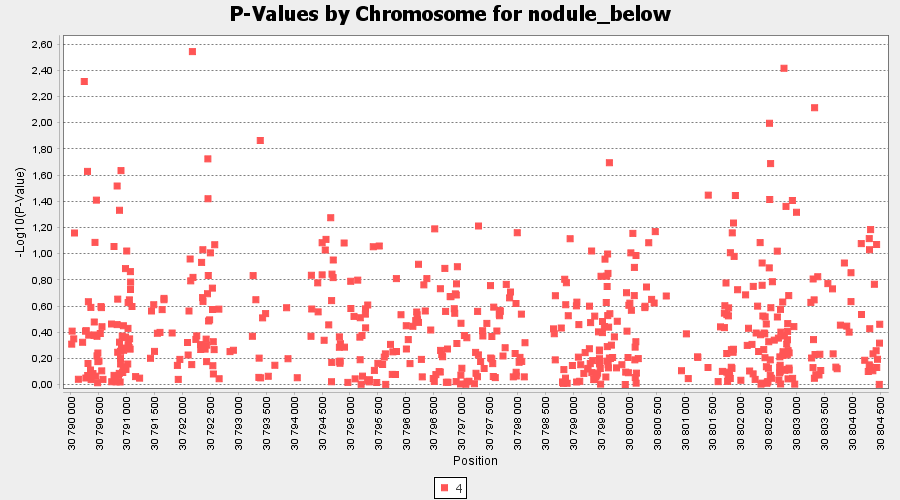


Рисунок. График Manhattan для участка хромосомы 4, включающего гены *MtCLE12* и *MtCLE13* по признаку «количество клубеньков в нижней части корневой системы».

**БЛАГОДАРНОСТИ**

Выражаю особую благодарность и признательность своему научному руководителю Лебедевой Марии Александровне за высокий профессионализм и помощь в проведении экспериментальных работ, а также за всестороннюю поддержку на протяжении всего периода обучения в магистратуре.

Работа выполнена при поддержке гранта СПбГУ, №60235510 «Изучение механизмов нитрат-опосредованной регуляции развития симбиотических клубеньков у гороха посевного (*Pisum sativum* L.) и люцерны слабоусеченной (*Medicago truncatula* Gaertn.) с использованием геномных технологий».