

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ НАУЧНОЕ
УЧРЕЖДЕНИЕ «ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
ИНСТИТУТ ЦИТОЛОГИИ И ГЕНЕТИКИ СИБИРСКОГО ОТДЕЛЕНИЯ
РОССИЙСКОЙ АКАДЕМИИ НАУК»

На правах рукописи

ДАРЬЯ ДМИТРИЕВНА НОВИКОВА

**Поиск новых чувствительных к ауксину регуляторных
элементов в промоторах генов *Arabidopsis thaliana* L.**

Генетика – 03.02.07

Математическая биология, биоинформатика – 03.01.09

Диссертация на соискание ученой степени

кандидата биологических наук

Научные руководители:

д.б.н., Алексей Владимирович Кочетов

к.б.н., Виктория Владимировна Миронова

Новосибирск 2019

ОГЛАВЛЕНИЕ

ОГЛАВЛЕНИЕ	2
ВВЕДЕНИЕ	5
ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ	11
1.1. Гормональная регуляция роста и развития растений	11
1.1.1. Основные фитогормоны растений и их функции	11
1.1.2. Сигнальный путь ауксина	13
1.1.3. Цис-регуляторные элементы в передаче сигнала от ауксина	19
1.1.4. Методы изучения регуляции активности генов	21
Заключение по обзору литературы	27
ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ	29
2.1. Используемые транскриптомные данные	29
2.2. Поиск и аннотация цис-регуляторных элементов в промоторах генов	31
2.3. Растительный материал	38
2.4. Подготовка РНК-образцов	38
2.5. Определение уровня экспрессии генов методом ОТ-количественной ПЦР	39
2.6. Метод безлигазного клонирования	40
2.7. Мутагенез методом ПЦР с использованием перекрывающихся праймеров	41
2.8. Трансформация растений с помощью агробактерии	43

2.9. Выявление взаимодействий белок-ДНК в дрожжевой одногобридной системе	44
2.10. Конфокальная микроскопия	47
ГЛАВА 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ	48
3.1. Биоинформатический анализ данных	48
3.1.1. Метод транскриптомного анализа ассоциаций (ТАА)	48
3.1.2. Поиск и функциональная аннотация простых ауксин-чувствительных элементов в промоторах генов <i>A.thaliana</i>	51
3.1.3. Обогащение предсказанных одиночных AuxRE в районах связывания ТФ ARF	56
3.1.4. Поиск и функциональная аннотация парных AuxRE в промоторах генов <i>A.thaliana</i>	59
3.1.5. Поиск регуляторных модулей, ассоциированных с ответом на ауксин	66
3.2. Экспериментальная верификация результатов биоинформатического анализа и биологической функциональности отобранных парных AuxRE	71
3.2.1. Верификация обнаруженных парных AuxRE методом ОТ-количественной ПЦР	71
3.2.2. Проверка функциональности предсказанных AuxRE методом направленного мутагенеза	78
3.2.3. Паттерны экспрессии белка GFP в репортерных линиях при стандартных условиях	80

3.2.4. Изменение паттернов экспрессии белка GFP в репортерных линиях <i>Arabidopsis</i> при обработке растений ауксином	85
3.2.5. Выявление транскрипционных факторов, связывающихся с идентифицированными парными AuxRE в промоторе гена <i>IAA30</i>	91
3.2.6. ARF5-зависимое подавление экспрессии гена <i>WOX4</i>	95
ЗАКЛЮЧЕНИЕ	97
ВЫВОДЫ	100
СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ	101
СПИСОК ЛИТЕРАТУРЫ	103
ПРИЛОЖЕНИЯ	123

ВВЕДЕНИЕ

Актуальность темы исследования

Фитогормон ауксин играет ключевую роль в регуляции развития и жизнедеятельности растений. Функции ауксина многосторонни и широко используются в разных технологиях, начиная от агротехники, где ауксин применяют и как гербицид для уничтожения сорных растений, и как стимулятор роста, до культуры клеток растений, где он служит индуктором соматического эмбриогенеза. Молекулярные основы реализации таких множественных сценариев клеточного ответа на ауксин, являющийся химически простой молекулой, остаются одним из фундаментальных и центральных вопросов генетики растений.

В основе регулируемых ауксином процессов лежит активация или подавление экспрессии тысяч генов (Weijers and Wagner, 2016). Транскрипционные факторы (ТФ) ARF (Auxin Response Factor) связываются с ауксин-чувствительными элементами (Auxin Responsive Elements; AuxRE) в регуляторных областях генов-мишеней. Геномы растений включают большие семейства ТФ ARF, которые могут регулировать уникальные наборы генов путем связывания с разными AuxRE (Blenau et al., 2011). Несмотря на прилагаемые существенные усилия, из-за функциональной избыточности ТФ ARF, найти мишени конкретных ARF и исследовать пути передачи сигнала ауксина, функционирующие при различных условиях, оказалось технически сложно. В данной работе мы предлагаем решить обратную задачу – найти и описать многообразие и структуру AuxRE и охарактеризовать ТФ, которые с ними связываются. ARF-связывающие AuxRE, выявляемые в

регуляторных районах многих ауксин-чувствительных генов, содержат последовательности TGTCGG, TGTCTC или TGTCCC (Zemlyanskaya et al., 2016). Однако эти элементы присутствуют и во многих нечувствительных к ауксину генах, из чего можно предположить существование дополнительных, еще неизученных особенностей регуляции транскрипционной активности генов ауксином.

Недавние исследования показали, что ТФ ARF связываются с ДНК в виде гомодимеров (Boer et al., 2014) и гетеродимеров (обзор в Cherenkov, Novikova et al., 2018). Биоинформатический анализ районов связывания ТФ ARF5 и ARF2 выявил в них обогащение прямых и инвертированных повторов AuxRE с различной длиной спейсера (Stigliani et al., 2018). Кроме того, в анализе промоторов ауксин-чувствительных генов было выявлено обогащение не ARF-связывающими цис-регуляторными элементами (Berendzen et al., 2012), в том числе в непосредственной близости от последовательностей TGTCNN (Mironova et al., 2014). Ключевым вопросом остается, как много таких цис-регуляторных элементов и какие еще ТФ, помимо ARF, модулируют ответ на ауксин?

Экспериментальный анализ районов связывания ТФ ARF имеет существенные ограничения – на данный момент удалось исследовать взаимодействие с ДНК лишь N-терминального ДНК-связывающего домена ТФ ARF, а не целого белка. Так как формирование гетеродимеров ARF с ТФ-партнерами, как правило, происходит через С-терминальный домен, то выявить такие взаимодействия по данным ChIP(DAP)-seq анализа не представляется возможным. Кроме того, если подразумевается наличие ТФ, которые модулируют ответ на ауксин независимо от ARF, то для решения обратной задачи надо использовать другие данные. В настоящее время в базах данных накоплено большое число

транскриптомных данных по ответу на ауксин, что позволяет провести их анализ и выявить новые AuxRE, обогащенные в промоторах дифференциально экспрессирующихся генов (ДЭГ), и, таким образом, дополнить теоретические знания о механизме регуляции генов ауксином.

Цели и задачи работы

Целью данной работы является идентификация ауксин-чувствительных цис-регуляторных элементов (AuxRE) в промоторах генов *Arabidopsis thaliana*.

Задачи:

1. Разработка биоинформатического метода по выявлению потенциальных цис-регуляторных элементов, ассоциированных с дифференциальной экспрессией генов.
2. Апробация биоинформатического метода в поиске потенциальных AuxRE.
3. Анализ функциональности кандидатных AuxRE в ответе на ауксин *in vivo* методом направленного мутагенеза промоторов генов *Arabidopsis thaliana*.
4. Экспериментальная проверка связывания кандидатных AuxRE с транскрипционными факторами *in vitro*.

Научная новизна

Новизна решения заявленной проблемы заключается в применении системно-биологического подхода к поиску молекулярно-генетических механизмов ответа на ауксин через анализ структуры и функциональности одиночных и парных AuxRE. Для решения этой задачи был разработан новый биоинформатический метод ТАА (Транскриптомного Анализа Ассоциаций), который по серии транскриптомных данных описывает разнообразие цис-элементов,

систематически обогащенных в промоторах ДЭГ. В экспериментальной части проекта были использованы современные методики, направленные на проверку функциональности предсказанных цис-элементов в ответе а ауксин.

Теоретическая и практическая значимость работы

Выявление ауксин-зависимых механизмов регуляции активности генов является актуальным и фундаментальным направлением исследований в биологии растений. Изучение молекулярных основ генерации множественных ответов на ауксин в развитии растения на транскрипционном уровне, выявление новых ТФ и цис-регуляторных элементов, вовлеченных в регуляцию этих процессов ауксином, открывает новые возможности для биотехнологии и сайт-направленной селекции растений. Предсказанные цис-регуляторные элементы с подтвержденной *in vivo* функциональностью могут быть использованы для создания ауксин-чувствительных репортерных конструкций, аналогов существующего DR5 (DIRECT REPEAT 5), но с более высокой и специфической чувствительностью.

Методология и методы диссертационного исследования

В данной работе разработан метод транскриптомного анализа ассоциаций (ТАА) для поиска связанных с транскрипционным ответом цис-регуляторных элементов (Cherenkov, Novikova et al., 2018). ТАА основан на анализе данных, полученных из множества транскриптомных экспериментов, что повышает достоверность предсказаний. Использование в анализе одновременно большого количества транскриптомных данных позволяет максимально сократить количество ложно-положительных результатов и найти наиболее полный список цис-элементов, вовлеченных в регуляцию транскрипции генов ауксином.

Для проверки функциональности предсказанных при помощи метода ТАА цис-регуляторных элементов в промоторах ауксин-чувствительных генов *A.thaliana* мы использовали методику мутагенеза с перекрывающимися праймерами. С помощью высокоэффективного безлигазного клонирования (Wendrich et al., 2015) мы получили генетические конструкции на основе плазмидного вектора pPLV104v2 (De Rybel et al., 2011). Трансформация растений *A.thaliana* осуществлялась при помощи *Agrobacterium tumefaciens*. Экспрессия паттернов активности белка GFP изучалась на корнях проростков *A.thaliana* на конфокальных микроскопах LSM 510 Meta и Leica SP5. Для поиска ТФ, вовлеченных в регуляцию транскрипции исследуемых генов, была использована дрожжевая одногибридная система.

Положения, выносимые на защиту:

- Разработан метод Транскриптного Анализа Ассоциаций (ТАА), позволяющий выявить многообразие ассоциированных с изучаемым феноменом регуляторных элементов в промоторах дифференциально экспрессирующихся генов.
- Промоторы генов первичного ответа на ауксин *Arabidopsis thaliana* обладают сложной структурой расположения ауксин-чувствительных регуляторных элементов и обогащены потенциальными сайтами связывания семейств транскрипционных факторов bZIP, bHLH, TCP и A/T-богатыми последовательностями.
- Транскрипционный фактор ARF5 может выполнять функцию активатора и репрессора транскрипции генов в ответ на фитогормон ауксин в зависимости от структуры цис-регуляторных элементов AuxRE.

Структура работы

Работа состоит из введения, обзора литературы, результатов и обсуждений работы, заключения, выводов, списка публикаций по теме диссертации, списка литературы (176 наименований). Материал изложен на 138 страницах, содержит 21 рисунок, 10 таблиц и 7 приложений.

Личный вклад автора

Основные результаты, изложенные в диссертации, получены и проанализированы автором лично.

Апробация результатов

Материалы настоящей работы вошли в отчеты по грантам Российского научного фонда и Российского фонда фундаментальных исследований. Основные результаты работы были представлены на научных конференциях и школах молодых ученых в виде устных и стендовых докладов: Международная конференция по биоинформатике, структуре и регуляции генома (BGRS\SB'2014 и BGRS\SB'2018, г. Новосибирск, Россия), Международная конференция по системной биологии (ICSB'16, г. Барселона, Испания), Всероссийская конференция с международным участием "Высокопроизводительное секвенирование в геномике" (HGS2017, г. Новосибирск, Россия), Международная конференция по экспериментальной биологии растений (EPS2018 и EPS2019, г. Люнтерен, Нидерланды), Международная конференция "Ауксины и цитокинины в развитии растений" (ACPD2018, г. Прага, Чехия), Международная конференция по ауксинам (Auxin Workshop 2018, г. Лидс, Великобритания), Международная конференция по биологически активным веществам растений (ICPGS2019, г. Париж, Франция).

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Гормональная регуляция роста и развития растений

1.1.1. Основные фитогормоны растений и их функции

Впервые идея о гормональной регуляции роста и развития растений была высказана Чарльзом Дарвином в 1880 году (Darwin, 1880). В дальнейшем это учение было развито в трудах отечественных и зарубежных физиологов растений (Кулаева, 1995; Кравец и др., 2008; Дерфлинг, 1985; Guern, 1987; Klee and Estelle, 1991).

Фитогормоны оказывают различные физиологические эффекты на растения, обеспечивают регуляцию морфогенеза и ростовых процессов растений, в том числе реакций на внешние воздействия и стрессы, половое размножение, переход в стадию покоя, старение и др. (Медведев, 2004). К классическим фитогормонам относят ауксины, цитокинины, гиббереллины, абсцизовую кислоту, брассиностероиды, этилен, кроме того функции фитогормонов могут выполнять оксипирины, салициловая кислота и ряд пептидов. Фитогормоны проявляют регуляторные функции в низких концентрациях, синтезируются в одних частях растительного организма и транспортируются в места функционирования, или же действуют непосредственно в том месте, где образуются (Ежова и др., 2003).

Фитогормоны образуют сложные сети взаимодействия при регуляции роста и развития растений. Действие различных фитогормонов может быть взаимосвязано на уровне их синтеза, транспорта, катаболизма, рецепции и передачи сигнала. На молекулярном уровне

гормональная регуляция активности генов может осуществляться на транскрипционном (инициация и элонгация транскрипции) и посттранскрипционном (функционирование рибосом, регуляция посредством микроРНК, деградация белка, модификации белка) уровнях. Регуляция транскрипции осуществляется фитогормонами через посредников. Посредниками могут выступать трансмембранные рецепторы, сигнал от которых проходит по многокомпонентной каскадной реакции в ядро, например, в случае действия brassinosteroidов, gibberellinов, cytokininов и этилена (Медведев, 2004). Также гормоны могут действовать на определенные процессы косвенно, через активацию других физиологических процессов, например, ауксин изменяет работу натрий-калиевой аденозинтрифосфотазы (H^+/K^+ -АТФазы) и аквапоринов, что приводит к изменению тургора и гидростатического давления, показатели которых важны для роста и деления клеток (Péret et al., 2012).

Ауксин — это производное триптофана, играющее роль морфогена во всех растительных организмах, в том числе и водорослях, регулирует рост и деление растительной клетки, развитие партенокарпических плодов, стимулирует образование боковых корней у стеблевых черенков, а также индуцирует образование этилена. Цитокинины вместе с ауксином контролируют процессы деления клеток и апикальное доминирование, а также участвуют в регуляции роста плодов, морфогенеза недифференцированных тканей, оказывают аттрагирующий эффект (Кулаева, 1995). Другие циклические биологически активные соединения — gibberellinы, вызывают удлинение стебля, активируя деление и растяжение клеток, индуцируют цветение растений, влияют на детерминацию пола, угнетают развитие семян в плодах. Немаловажную

роль в развитии растения играет абсцизовая кислота, она ингибирует прорастание семян и рост почек, выступает антагонистом ауксина, цитокинов и гиббереллинов (Ракитина и др., 1994). Единственным известным газообразным гормоном растений является этилен. Этилен участвует в ответе на стрессовые воздействия, контролирует созревание плодов, старение тканей, опадение плодов и листьев, развитие цветков, эпинастию, образование корневых волосков и ускорение прорастания семян (Кулаева, 1998). Брассиностероиды, вещества стероидной природы, обладают ростостимулирующим эффектом, являются агонистами ауксина (Ковганко, 1991). Жасмонаты, относящиеся к оксипиринам, участвуют в ответных реакциях при повреждении растений патогенами и насекомыми (Медведев, 2004).

1.1.2. Сигнальный путь ауксина

Механизм реализации клеточного ответа на ауксин осуществляется в основном через изменение транскрипции генов. Тысячи генов изменяют свою экспрессию в ответ на ауксин (Paponov et al., 2008). Путь передачи сигнала от ауксина был рассмотрен во многих работах (Chapman and Estelle, 2009; Salehin et al., 2015; Leyser 2018). Ауксин обеспечивает взаимодействие белков TRANSPORT INHIBITOR RESPONSE1/AUXIN SIGNALING F-BOX (TIR1/AFB) семейства F-бокс с белками-репрессорами ответа на ауксин Aux/IAA (Tan et al., 2007). F-бокс белки являются субъединицей узнавания субстрата в комплексах убиквитинлигаз SCF, также включающих субъединицу Skp1 и субъединицу Cullin (Smalle and Vierstra, 2004). Убиквитинируемый белок доставляется в комплекс SCF благодаря его взаимодействию с С-терминальным доменом белка семейства F-бокс TIR1/AFB, состоящего

из лейцин-богатого повтора и центра связывания ауксина (Tan et al., 2007). Связывание ауксина стабилизируется белками Aux/IAA на входе в центр связывания за счет короткого аминокислотного мотива белка Aux/IAA, известного как домен II (Tan et al., 2007). Таким образом, ауксин-опосредованное связывание белков Aux/IAA с TIR1/AFB обеспечивает их транспортировку к комплексу SCF для убиквитинирования (Ub) и деградации Aux/IAA (Gray et al., 2001; Maraschin et al., 2009) (Рисунок 1).

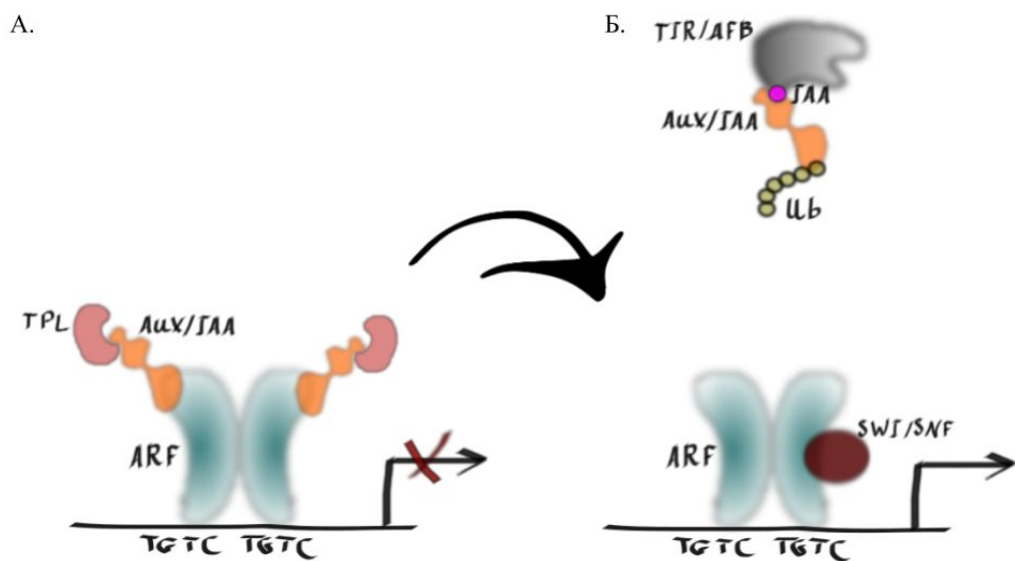


Рисунок 1. Схема механизма регуляции транскрипции генов ауксином через ТФ семейства ARF. А. В отсутствии ауксина транскрипция ингибируется. Б. Активация транскрипции в присутствии ауксина. ARF — ауксин-чувствительные ТФ, Aux/IAA — белки-ингибиторы ТФ ARF, TPL — белки-корепрессоры, Ub — убиквитин, TIR/AFB - белки семейства F-бокс, IAA — ауксин, SWI/SNF — белки-модуляторы упаковки хроматина, TGTC — коровая последовательность AuxRE.

В геноме *Arabidopsis thaliana* присутствуют 29 генов, кодирующих белки Aux/IAA (Paponov et al., 2008). Период полураспада этих белков и степень влияния на него экзогенного ауксина сильно варьирует (Dreher et al., 2006). Также в геноме *Arabidopsis* имеются гены шести белков TIR/AFB (Dharmasiri et al., 2005; Prigge et al., 2016), взаимодействующих с разными белками Aux/IAA с различной аффинностью (Villalobos et al., 2012). Белки Aux/IAA являются транскрипционными репрессорами (Ulmasov et al., 1997) за счет наличия консервативного домена EAR, связывающегося с белками-корепрессорами семейства TOPLESS (TPL) (Tiwari et al., 2004; Szemenyei et al., 2008) (Рисунок 1). В свою очередь, белки TPL подавляют транскрипцию, задействуя моделирующие хроматин белки SWI/SNF (Szemenyei et al., 2008). Белки Aux/IAA не связываются с ДНК самостоятельно, но могут димеризоваться с ТФ семейства AUXIN RESPONSE FACTOR (ARF). Димеризация осуществляется через С-терминальный домен (PROTEIN BINDING 1; PB1), присутствующий в белках обоих семейств (Guilfoyle, 2015).

Белки ARF также могут гомодимеризоваться через их N-терминальные домены В3 и в виде гомо- и гетеродимеров связываться с ДНК (Boer et al., 2014). Белки ARF связываются с AuxRE, представленными консенсусной последовательностью TGTCNN в регуляторных районах ауксин-чувствительных генов (Mironova et al., 2014). Группа ТФ ARF с Q-богатым средним доменом, расположенным между доменами В3 и PB1, обладает активационной способностью, осуществляемой через белки SWI/SNF, изменяющие структуру хроматина (Ulmasov et al., 1999; Wu et al., 2015). Связывание ТФ ARF с белками Aux/IAA препятствует активации транскрипции, а ауксин, в свою

очередь, модулирует экспрессию генов с сайтами связывания ТФ ARF за счет регуляции деградации белков Aux/IAA.

Одним из механизмов, обеспечивающих специфичность ответа на ауксин, является тканеспецифичная экспрессия белков семейства ARF, так как от их наличия полностью зависит регуляторная способность белков Aux/IAA (Rademacher et al., 2012; Bargmann et al., 2013). Для *Arabidopsis* было описано 23 разных ТФ ARF, отличающихся по степени аффинности к белкам-репрессорам Aux/IAA (Vernoux et al., 2011). Были продемонстрированы отличия по степени аффинности связывания димеров ТФ ARF с разными вариантами повторов TGTCNN в зависимости от расстояния (спейсера) между ними (Boer et al., 2014). Белки ARF, не обладающие Q-богатым средним доменом, могут выступать в роли репрессоров транскрипции (Ulmasov et al., 1999). Белки-репрессоры ARF конкурируют с белками-активаторами ARF за одни и те же сайты связывания на промоторах генов (Vert et al., 2008). Наличие белков-репрессоров ARF в клетке обеспечивает дополнительный механизм ответа на ауксин. Возможно, что конкуренция между ARF белками на промоторах ауксин-чувствительных генов подвержена влиянию других белков, не взаимодействующих непосредственно с регуляторным районом.

Описанный механизм обеспечивает быстрые изменения транскрипции в ответ на ауксин, которые детектируются уже после 3-5 минут обработки растений этим гормоном (Abel и Theologis, 1996). Период полураспада многих белков Aux/IAA очень мал, даже в отсутствии экзогенной обработки ауксином (Abel et al., 1994), и гены, кодирующие эти белки, быстро повышают свою экспрессию в ответ на ауксин (Abel и Theologis, 1996). Таким образом, клетки непрерывно

синтезируют и подвергают деградации белки Aux/IAA, этот непрерывный поток регулируется ауксином и перестраивается в зависимости от его концентрации. Передача сигнала от ауксина очень чувствительна к нарушениям в основных частях механизма деградации и синтеза белка (del Pozo et al., 2002; Gray et al., 2003; Hellmann et al., 2003; Chuang et al., 2004; Stirnberg et al., 2012).

Существует гипотеза, что способность ауксина регулировать транскрипцию могла быть добавлена к системе регуляции транскрипции, возникшей раньше в эволюции (Leyser, 2018). Полностью ауксин-независимая система, состоящая только из конкурирующих за один промотор активаторов и репрессоров ARF, могла обеспечивать модуляцию транскрипции многих генов, например, изменение роста в ответ на факторы окружающей среды. При добавлении в такую систему Aux/IAA-зависимой репрессии транскрипции и ауксин-зависимой деградации белков Aux/IAA, гены стали ауксин-чувствительными в присутствии белков-активаторов ARF. Свидетельства описанного порядка эволюционных событий были продемонстрированы в последних работах на низших растениях (Flores-Sandoval et al., 2015; Kato et al., 2015) — печеночном мхе *Marchantia polymorpha* и настоящем мхе *Physcomitrella patens*. В геноме *Marchantia* присутствует только три ТФ ARF, один белок Aux/IAA и один рецептор TIR1/AFB. Согласно последним исследованиям система передачи сигнала от ауксина у упомянутых низших растений функционирует так же, как и у высших (Flores-Sandoval et al., 2015; Kato et al., 2015). Белки ARF можно разделить на три клады, каждая из которых представлена в геномах низших и высших растений (Kato et al., 2015; Mutte et al., 2018). При анализе нокдаун мутантов и мутантов с редуцированной функцией

основных компонентов передачи сигнала от ауксина у низших растений были выявлены разные аспекты регуляции роста и развития ауксином (Flores-Sandoval et al., 2015; Kato et al., 2015).

У *P. patens* присутствуют 16 белков ARF, представители всех трех клад, и три белка Aux/IAA (Prigge et al., 2010). Растения-нокдауны по трем генам *Aux/IAA* демонстрируют полное отсутствие транскрипционного ответа на ауксин и фенотипически похожи на растения, обработанные высокой концентрацией ауксина (Lavy et al., 2016). Так повышенная экспрессия белка-репрессора ARF в нокдаун мутанте по всем генам *Aux/IAA* подавляла фенотип, свойственный растениям, обработанным высокой концентрацией ауксина, и придавала фенотип, сходный с фенотипом, обусловленным экспрессией стабилизированных ауксин-устойчивых вариантов белка Aux/IAA (Lavy et al., 2016). Эти результаты подтверждают идею, что механизм передачи сигнала от ауксина возник как усовершенствование системы, основанной на контроле соотношения белков-активаторов ARF и белков-репрессоров ARF. Особенный интерес представляет недавно открытый неканонический механизм передачи сигнала от ауксина через ТФ ETTIN/ARF3. ТФ ETTIN не обладает доменом PB1, взаимодействующим с белками Aux/IAA, но тем не менее обеспечивает ответ на ауксин при развитии гинецея цветка (Sessions et al., 1997; Nemhauser et al., 2000). В этом процессе ТФ ETTIN ауксин-зависимо димеризуется с ТФ INDEHISCENT (IND) для регуляции транскрипции ключевых генов (Simonini et al., 2016).

1.1.3. Цис-регуляторные элементы в передаче сигнала от ауксина

Последовательность TGTCTC является первым наиболее коротким цис-регуляторным элементом, идентифицированным как ауксин-чувствительный (AuxRE) (Ulmasov et al., 1997). Он может функционировать в составе парных AuxRE с элементом-партнером, сайтом-связывания другого ТФ, или в прямых и палиндромных повторах последовательностей TGTCNN. Белок ARF связывается с палиндромными повторами одиночных AuxRE со спейсером в 7-9 п.н. в виде гомо- или гетеродимера (Boer et al., 2014). Синтетические конструкции с множественными повторами последовательностей TGTCNN в 5-10 раз более чувствительны к ауксину, чем натуральные (Ulmasov et al., 1999), что было использовано при создании сенсорной конструкции DR5 (DIRECT REPEAT 5), которая используется для мониторинга клеточного или тканевого ответа на ауксин (Péret et al., 2012).

Ранее была показана взаимосвязь ряда известных сайтов связывания ТФ с передачей сигнала от ауксина (Trenner et al., 2016). Более того, промоторные районы ауксин-чувствительных генов обогащены не только TGTCNN-содержащими цис-регуляторными элементами, связывающими ТФ ARF, но также сайтами связывания других ТФ (Berendzen et al., 2012). Возможно, что специфика и комплексность в регуляции процессов ауксином в значительной степени обеспечивается за счет присутствия элементов-партнеров, сайтов связывания других ТФ, рядом с TGTCNN и взаимодействия этих ТФ с белками ARF (Vernoux et al., 2011; Boer et al., 2014). Было показано взаимодействие между белками ARF7 и MYB77 (Shin et al., 2007); белком ARF8 и ТФ bHLH (bPER) (Varaud et al., 2011);

ТФ ARF6 и bHLH (PIF4) (Oh et al., 2014); ТФ ARF6/8 с MADS (FUL) (Ripoll et al., 2015); ТФ ARF3 с семействами ТФ bHLH (IND), Homeobox (RPL, KNAT1 и KNAT3), AP2 (BBM и PLT5) и TCP (TCP4 и TCP18) (Simonini et al., 2016).

Чувствительный к абсцизовой кислоте цис-регуляторный элемент (Abscisic Acid Response Element; ABRE) ACGTG(G/T)C (Choi et al., 2000) был впервые описан как элемент-партнер к последовательности TGTCTC в промоторе гена *GH3* у сои, также было подтверждено его связывание с ТФ bZIP (Ulmasov et al., 1995; Liu et al., 1997). До настоящего времени не было показано физического взаимодействия между белками bZIP и ARF, но ТФ bZIP11 и его гомологи у *A.thaliana* регулируют ответ на ауксин посредством взаимодействия с белком-модулятором хроматина ADA2b (Weiste and Dröge-Laser, 2014). Таким образом, сайты связывания белков bZIP могут быть элементами-партнерами для последовательностей TGTCNN в парных AuxRE, в этом случае они вместе с последовательностью TGTCNN опосредуют изменение транскрипции с промотора гена *GH3* в ответ на ауксин (Ulmasov et al., 1995; Weiste and Dröge-Laser, 2014).

Наряду с чувствительными к абсцизовой кислоте цис-регуляторными элементами, ТФ bZIP связываются с другими ACGT-содержащими сайтами, которые делятся на А-боксы (TACGTA), С-боксы (GACGTC) и G-боксы (CACGTG) (Izawa et al., 1993; Foster et al., 1994; Jakoby et al., 2002). ТФ семейства bHLH (Martínez-García et al., 2000; Oh et al., 2012; Kim et al., 2016; Dombrecht et al., 2007; Fernandez-Calvo et al., 2011), белки AP2/ERF, ABI4 (Zhang et al., 2013) и белок BZR1/BES1 (Yu et al., 2011) также могут связываться с последовательностью G-бокс.

Ранее показали, что мотив CACATG, с которым связываются ТФ bHLH и BZR1/BES1 (Walcher and Nemhauser, 2012; Oh et al., 2014), цис-элементы MRE (AACC) и MYB (CNGTTR) (Shin et al., 2007), последовательности CArG-боксы (CC[A/T]6GG) (Ripoll et al., 2015), с которыми взаимодействуют ТФ семейства MADS, часто располагаются рядом с последовательностью TGTCNN и, возможно, являются частью функциональных парных AuxRE. К тому же, регуляторные районы ауксин-чувствительных генов, наряду с последовательностью TGTCNN, обогащены большим числом других цис-регуляторных элементов (Pufky et al., 2003; Doi et al., 2008; Berendzen et al., 2012; Mironova et al., 2014), которые также могут быть элементами-партнерами в парных AuxRE. Однако, за исключением очень небольшого числа случаев (Boer et al., 2014, Mironova et al., 2014), системный анализ сопредставленности обогащенных цис-элементов и последовательности TGTCNN ранее не проводился. Ключевым вопросом остается, какую роль такие парные AuxRE играют в регуляции экспрессии ауксин-чувствительных генов. Возможно, что с ними связываются другие белки ARF, еще не изученные детально, или димеры ТФ ARF и других ТФ. Таким образом, в то время как основные механизмы ауксин-зависимой регуляции генов были исследованы (обзор в Weijers and Wagner, 2016), многие специфические особенности этого процесса до сих пор остаются неизвестными. Неизвестны также цис-регуляторные элементы, ассоциированные с подавлением экспрессии генов ауксином.

1.1.4. Методы изучения регуляции активности генов

Методы исследования регуляторных элементов можно разделить на экспериментальные и теоретические. Из экспериментальных методов

наиболее часто используются методы, основанные на иммунопреципитации (ChIP, ChIP-chip, ChIP-seq, DAP-seq), методы делеционного анализа регуляторных районов, методы направленного мутагенеза регуляторных районов, метод измерения электрофоретической подвижности в геле (electrophoretic mobility shift assay; EMSA) и метод с применением дрожжевой одногибридной системы (yeast one hybrid; Y1H). Из теоретических методов исследования цис-регуляторных элементов основными являются предсказания с помощью консенсуса, позиционно-весовых матриц, филогенетического футпринтинга и методы машинного обучения.

Наиболее простым теоретическим подходом к поиску регуляторных элементов является метод консенсуса. Он заключается в использовании консенсусной символьной последовательности. Данный метод позволяет проводить быстрое сравнение идентичных последовательностей, но в нем не учитываются численные характеристики: в вырожденном коде не принимаются во внимание вклады каждого из невырожденных нуклеотидов, находящихся в одной позиции. Использование расширенного алфавита нуклеотидов IUPAC позволяет составлять консенсусы, которые распознают больше реальных сайтов, но в тоже время они чаще предсказывают ложные сайты (Wasserman and Sandelin, 2004).

Более точным методом распознавания цис-регуляторных элементов является метод позиционно-весовых матриц. Частотная матрица содержит информацию о встречаемости каждого нуклеотида в определенной позиции. Каждый столбец матрицы содержит частоты для определенной позиции изучаемой последовательности. После нормировки такой матрицы на общее число последовательностей в

выравнивании получается таблица вероятностей встречаемости для каждого нуклеотида в каждой позиции. С помощью вероятностной матрицы вычисляется весовая матрица (Berg and von Hippel, 1987). Матричный метод позволяет дать численную оценку любой последовательности ДНК для идентификации потенциального сайта связывания ТФ (Stormo, 2000).

Существуют также методы предсказания цис-регуляторных элементов путем машинного обучения, основанные на статистическом анализе, моделях Маркова и нейронных сетях (Vityaev et al., 2001; Vityaev et al., 2002).

Метод филогенетического футпринтинга, применяемый для предсказания сайтов связывания ТФ, основывается на сходстве последовательностей ортологичных генов. Главное предположение этого метода заключается в том, что мутации в функциональных районах генов накапливаются значительно медленнее, чем в участках без контекстно-специфичной функции. Сравнение последовательностей в 5'-областях ортологичных генов может указать на некодирующие последовательности, задействованные в процессах регуляции транскрипции. Метод филогенетического футпринтинга позволяет обнаруживать регулирующие экспрессию районы генома с небольшой вероятностью пропуска функционально-значимых участков. Но сам исследователь должен принимать решение, на основании изучения эволюционных расстояний, верно ли это предположение в определенном случае. Методы филогенетического футпринтинга, основанные на выравнивании регуляторных районов генов, уместны только для ортологов (Lenhard, 2006).

Делеционный анализ является базовым экспериментальным методом исследования регуляторных районов *in vivo*. Его используют, например, для определения длины промотора и расположения регуляторных элементов в нем (Liu, 1994). Для этого создают набор генетических конструкций, несущих промотор интересующего гена различной длины. “Нарезка” промоторного района гена производится так, что 3'-конец регуляторного района фиксирован, а 5'-конец варьируется или наоборот. Ниже этих вариантов промотора располагается кодирующая последовательность репортерного гена. Различия в паттернах экспрессии репортера в линиях с разными вариантами промотора изучают экспериментально. Анализ этих данных позволяет делать выводы о локализации интересующего регуляторного элемента в исследуемом промоторе.

При наличии потенциальных функциональных сайтов связывания ТФ в промоторе используется более прицельный подход — метод направленного мутагенеза, который позволяет мутировать одиночные нуклеотиды или последовательности нуклеотидов в интересующих исследователя позициях. Мутагенез может осуществляться при помощи разных методик в зависимости от конечной цели: CRISPR (Li et al., 2018), ПЦР с использованием перекрывающихся праймеров (Ho et al., 1989) и др.

При изучении сайтов связывания ТФ важным этапом является идентификация взаимодействий ТФ-сайт связывания, последняя может быть определена несколькими методами, описанными ниже. Эти методы предназначены для изучения взаимодействия белка с ДНК или РНК *in vitro*. Дрожжевая одногибридная система основана на том, что ТФ состоят, как минимум, из двух частей (доменов) – ДНК-связывающего и

активирующего. Эти домены можно искусственно разделить и “слить” с другими исследуемыми белками. Гаплоидные дрожжи несут репортерную систему с встроенными регуляторными последовательностями и генами, кодирующими ТФ, взаимодействие которых необходимо протестировать. После скрещивания дрожжей, при наличии взаимодействия между ТФ и ДНК, происходит связывание активирующего домена с репортерной системой и активация синтеза белка с гена-репортера (Gaudinier et al., 2011).

Метод измерения электрофоретической подвижности в геле позволяет определить возможность связывания определенных белков, в том числе и ТФ, с нуклеотидными последовательностями, потенциальными сайтами связывания этих белков. Метод заключается в электрофоретическом разделении смесей, состоящих из белков и последовательностей нуклеотидов (ДНК или РНК), в полиакриламидном или агарозном геле (Orchard and May, 1993). Для визуализации результатов данного эксперимента фрагменты нуклеотидных последовательностей метят радиоактивными, флуоресцентными или биотиновыми метками.

Для анализа цис-регуляторных последовательностей также используют метод иммунопреципитации хроматина (chromatin immunoprecipitation; ChIP) (Liu et al., 2002). Он служит для изучения взаимодействия ТФ с последовательностями ДНК, для определения участков ДНК, ассоциированных с различными модификациями гистонов, а также для анализа взаимодействия ТФ с белками, напрямую не контактирующими с последовательностью ДНК. На первом этапе поиска сайтов связывания ТФ производят “сшивание” белков с хроматином. На следующем этапе хроматин обрабатывают при помощи

ультразвука или ферментативного расщепления. Иммунопреципитацию “сшитого” хроматина осуществляют, используя антитела, которые распознают специфичные белки, например, ТФ. В результате эксперимента можно получить коллекцию всех сайтов связывания интересующего ТФ в масштабах генома (Valouev et al., 2008), также для подобного анализа используют биочипы (ChIP-chip) или секвенирование (ChIP-seq). В ChIP-chip эксперименте иммунопреципитированный и контрольный образцы метят флуоресцентными метками и совместно или отдельно гибридизуют на ДНК-микрочип. Сайты связывания идентифицируют по интенсивности сигнала иммунопреципитированного образца относительно контрольного образца. ChIP-seq – это полногеномный анализ сайтов связывания изучаемого ТФ, при котором производится секвенирование всех фрагментов ДНК, связанных с белком. Технология ChIP-seq используется для определения мишеней ТФ и других белков, ассоциированных с хроматином.

Помимо указанных выше методов, поиск и анализ цис-регуляторных последовательностей может быть основан на транскриптомных методах определения экспрессии генов. В настоящее время накоплено большое число транскриптомных данных, полученных методами иммунопреципитации (DNA-chip) и секвенирования транскриптомов (RNA-seq). При таком анализе после определения дифференциально экспрессирующихся генов возможно исследование их регуляторных районов для поиска обогащенных последовательностей, потенциальных сайтов связывания ТФ. Например, при анализе девяти транскриптомных экспериментов были выявлены сайты связывания ТФ, ассоциированные с раком груди (He et al., 2016). Также несколько транскриптомных

экспериментов могут быть использованы для поиска ТФ, контролируемых одни и те же гены (Kim and Kim, 2006).

Заключение по обзору литературы

Гормональная регуляция играет ключевую роль в росте и развитии растений, реализации защитных механизмов и реакции на изменения окружающей среды. Среди фитогормонов растений важнейшую роль играет ауксин, большинство ключевых посредников сигнального пути ауксина хорошо изучены. Помимо ТФ ARF и их сайтов связывания TGTCNN, в регуляцию могут быть вовлечены другие ТФ за счет непосредственной близости их сайтов связывания. Участие ТФ, не относящихся к семейству белков ARF, в регуляции ответа на ауксин является одним из возможных механизмов, обеспечивающих специфическую регуляцию ауксином транскрипционной активности генов, вовлеченных в большое количество процессов. Однако в литературе присутствуют лишь разрозненные данные по участию тех или иных ТФ в ответе на ауксин.

Существует ряд теоретических и экспериментальных подходов к изучению регуляции экспрессии генов. К первым относится предсказание цис-регуляторных элементов в промоторах генов методом консенсуса, позиционно-весовых матриц, филогенетического футпринтинга и др. Среди экспериментальных методов по выявлению взаимодействия между белком и цис-регуляторной последовательностью ДНК исторически устоявшимися являются метод измерения электрофоретической подвижности в геле и метод дрожжевой одногибридной системы, выявление биологической функциональности цис-элементов *in vivo* зачастую осуществляется при помощи мутации цис-элемента в репортерной конструкции (раздел 1.1.4).

В настоящее время в базах данных накоплены данные большого числа транскриптомных экспериментов по ответу на ауксин, что позволяет провести мета-анализ этих данных и предсказать разнообразие ТФ, участвующих в регуляции экспрессии генов ауксином. Поэтому целью данной работы стал поиск связанных с ответом на ауксин цис-регуляторных элементов в промоторах чувствительных к ауксину генов *A.thaliana* и их экспериментальная верификация с помощью перечисленных выше методов.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ

2.1. Использованные транскриптомные данные

Для биоинформатического анализа были использованы данные всех транскриптомных экспериментов *A.thaliana*, имеющих в свободном доступе, в которых мы выявили как минимум десять дифференциально-экспрессирующихся генов (ДЭГ) в ответ на обработку растений ауксином (Таблица 1) (Omelyanchuk et al., 2017). Данные транскриптомных экспериментов были взяты из базы данных GEO (<https://www.ncbi.nlm.nih.gov>; Barrett T. et al., 2005). Активация транскрипции гена учитывалась при изменении экспрессии $> 1,5$ и средней доле ложных отклонений гипотез (false discovery rate; FDR) $< 0,05$, ингибирование – при изменении экспрессии $< 1/1,5$ и FDR $< 0,05$. Доля ложных отклонений гипотез рассчитывалась по методу Бенджамини-Якутели (Benjamini-Yekutieli) (Benjamini and Yekutieli, 2001).

Таблица 1. Используемые в работе транскриптомные данные.

Идентификатор данных	Ткань	Стадия развития (сутки)	Время обработки (ч)	Активно е вещество	Концентрация (мкМ)
GDS151	корень	3	6	НУК	10
GDS151	корень	3	2	НУК	10
GSE35580	корень	7	3	ИУК	5
GSE35580	корень	7	3	ИУК	5
GSE35580	корень	7	3	ИУК	5
GSE35580	корень	7	3	ИУК	5
GSE35580	корень	7	3	ИУК	5
GSE35580	корень	7	3	ИУК	5
GSE42007	корень	6	4	ИУК	1
GSE42007	корень	6	2	ИУК	1
GSE627	корень	7	2	ИУК	5
GSE42007	корень	6	12	ИУК	1
GSE42007	корень	6	8	ИУК	1
GSE42007	корень	6	24	ИУК	1
GSE42007	корень	6	1	ИУК	1
GDS1044	проросток	7	1	ИУК	10
GDS672	проросток	10-12	1	ИУК	1
GDS744	проросток	5	2	ИУК	10
GDS672	проросток	10-12	3	ИУК	0,1
GDS672	проросток	10-12	1	ИУК	0,1
GSE42007	корень	6	0,5	ИУК	1
GDS672	проросток	10-12	3	ИУК	1
GDS3505	корень	3	4	ИУК	1
GSE39384	проросток	7	1	ИУК	1
GSE39384	проросток	7	3	ИУК	1
GSE42896	корень	3	2	НУК	10
GSE42896	корень	3	6	ИУК	10
GSE17479	лист	35	1	ИУК	10
GSE17479	лист	35	2	ИУК	10
GSE6349	проросток	5	2	ИУК	10
GSE6349	проросток	5	6	НУК	10
GSE18975	проросток	7	0,5	ИУК	1
GSE18975	проросток	7	1	ИУК	1
GSE18975	проросток	7	3	ИУК	1
GSE59426	корень	3	6	ИБК	10
GSE52966	корень	7	4	ИУК	5

Примечание: НУК - 1-нафтилуксусная кислота, ИУК - индолил-3-уксусная кислота, ИБК - индолбутировая кислота.

2.2. Поиск и аннотация цис-регуляторных элементов в промоторах генов

Модуль пакета программ metaRE

На основе разработанного мною метода ТАА (транскриптомного анализа ассоциаций, см. раздел 3.1.1) для поиска цис-регуляторных элементов, перепредставленных в промоторах генов, изменивших свою экспрессию в ответ на внешнее воздействие, Черенковым П.А. в рамках его дипломной работы был разработан пакет программ metaRE (Рисунок 2) (<https://github.com/cheburechko/metaRE>). Особенностью этого пакета программ является возможность искать ассоциированные с изучаемым феноменом цис-регуляторные последовательности ДНК на основании анализа большого количества транскриптомных данных, тестирующих одну и ту же гипотезу (например, одна и та же обработка, одно и то же заболевание и т.д.). Работа пакета metaRE была апробирована в данном исследовании на поиске потенциальных AuxRE в промоторах чувствительных к ауксину генов *A.thaliana*.

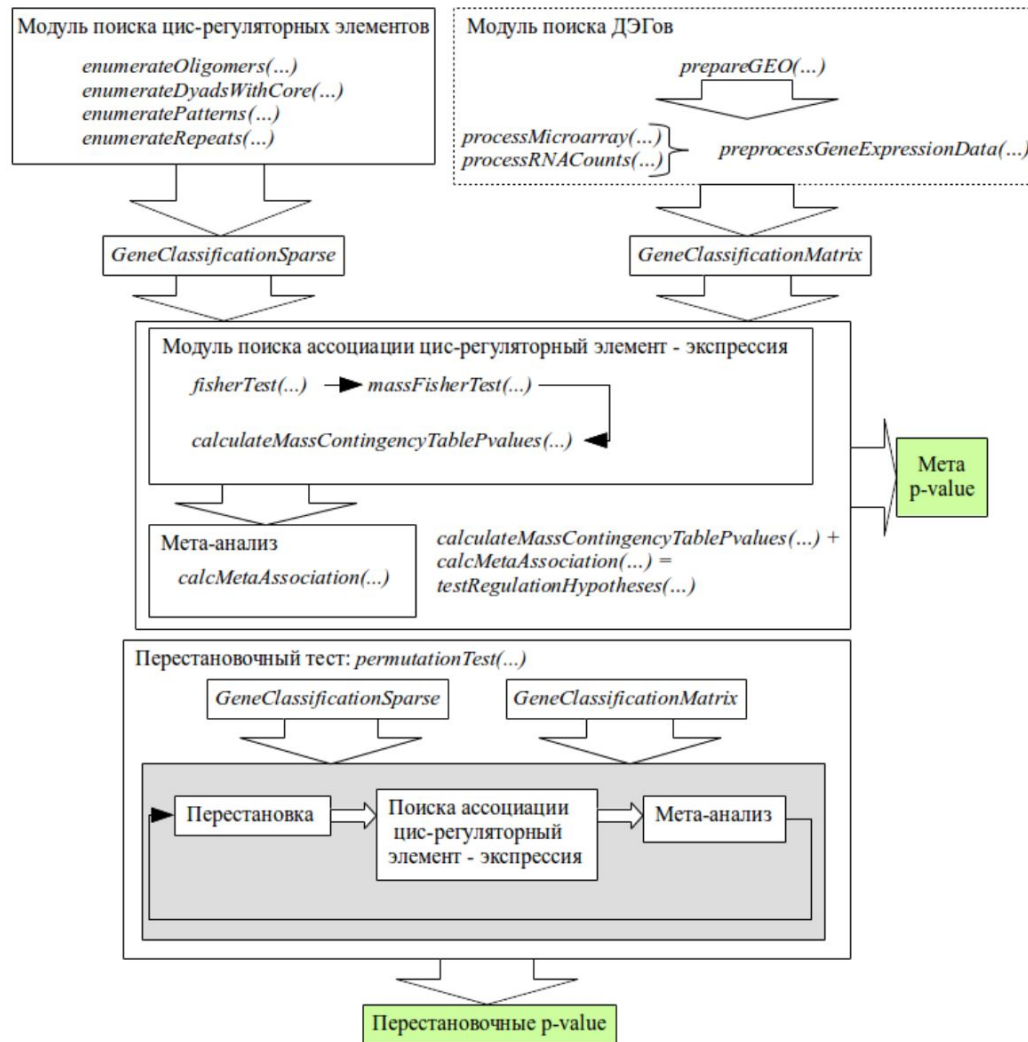


Рисунок 2. Схема работы модулей пакета программ metaRE по поиску обогащения цис-регуляторными последовательностями регуляторных районов ауксин-чувствительных генов. Тонкая и широкие стрелки отображают последовательность выполнения функций и блоков функций (заключены в прямоугольники), обычным шрифтом подписаны названия блоков программы, “Мета *p-value*” и “Перестановочные *p-value*” отображают применяемые на каждом этапе статистические тесты (Cherenkov, Novikova et al., 2018).

Пакет программ metaRE реализует пятиступенчатый анализ: (1) поиск ДЭГ; (2) поиск цис-регуляторных элементов заданной длины и

структуры; (3) расчет ассоциации между консенсусным присутствием элемента и изменением экспрессии генов; (4) мета-анализ на нескольких экспериментах по изменению экспрессии генов; (5) перестановочный тест (Рисунок 2).

(1) поиск ДЭГ

Входными данными пакета программ metaRE являются транскриптомные данные. Для идентификации ДЭГ в экспериментах, взятых из базы GEO были использованы пакеты *GEOquery* (Davis et al., 2007), *limma* (Ritchie et al., 2015) и *edgeR* (Robinson et al., 2010; McCarthy et al., 2012). Функция *prepareGEO* позволяет загружать и изменять предобработанные таблицы данных из GEO. Функции *processMicroarray* и *processRNAcounts* могут быть использованы для поиска ДЭГ в одиночном эксперименте, они создают новую таблицу для каждого эксперимента индивидуально с заданными пользователем классами ДЭГ, где для каждого гена определяется значение “up” (активация экспрессии гена), “down” (подавление экспрессии гена), или “” (отсутствие изменения экспрессии гена). Функция *preprocessGeneExpressionData* может выполнять такой же анализ одновременно для нескольких экспериментов, она создает финальную таблицу *GeneClassificationMatrix*, в которой содержится информация о ДЭГ со всех экспериментов, взятых в мета-анализ.

(2) Поиск цис-регуляторных элементов

Другими входными данными пакета программ metaRE являются регуляторные последовательности ДНК в *fasta* формате. Мы использовали встроенный в metaRE пакет *Biostrings* (Pagès et al., 2018) для загрузки последовательностей ДНК из BioMart

(<http://www.biomart.org>; Durinck et al., 2005). Для извлечения регуляторных районов генов из *GeneClassificationMatrix* используется функция *metaRE intersect*. Далее пакет аннотирует присутствие потенциальных цис-регуляторных элементов для каждой регуляторной последовательности ДНК. Функция *enumerateOligomers* ищет все возможные олигомеры без учета комплементарных последовательностей ДНК для каждого олигомера, например, в случае гексамеров *metaRE* находит 2080 невырожденных последовательностей ДНК, состоящих из 2016 комплементарных пар и 64 палиндромов вместо 4096 всевозможных комбинаторных вариантов. Помимо аннотации олигомеров возможен систематический анализ регуляторных районов на наличие всевозможных повторов со спейсером разной длины и фиксированной коровой последовательностью парного элемента (*enumerateRepeats*), парных элементов с заданными коровыми последовательностями (*enumerateDyadsWithCore*). Также возможен поиск заданного списка мотивов, описанных пятнадцатibuквенным вырожденным кодом IUPAC (*enumeratePatterns*). В случае функций *enumerateRepeats* и *enumerateDyadsWithCore* можно задавать максимальную и минимальную длину спейсера. В результате программа выдает список векторов в формате *integer*.

(3) *Определение ассоциации между присутствием цис-регуляторного элемента и изменением экспрессии генов.*

На данном этапе для каждой цис-регуляторной последовательности гена для каждого набора транскриптомных данных устанавливается ассоциация с дифференциальной экспрессией генов, отдельно для положительной и отрицательной регуляции. Значимость (*p-value*)

ассоциации рассчитывается из таблицы сопряженности 2x2 (Таблица 2) точным тестом Фишера (Cherenkov, Novikova et al., 2018; Zemlyanskaya et al., 2016). В metaRE процедура реализуется функцией *calculateMassContingencyTablePvalues*.

Таблица 2. Таблица сопряженности 2x2 для расчета обогащения потенциального цис-элемента в промоторах ДЭГ.

		ДЭГ		Всего
		+	-	
Присутствие цис-элемента в регуляторном районе	+	<i>a</i>	<i>b</i>	<i>a+b</i>
	-	<i>c</i>	<i>d</i>	<i>c+d</i>
Всего		<i>a+c</i>	<i>b+d</i>	<i>n=a+b+c+d</i>

(4) Мета-анализ

Функция *calcMetaAssociation* используется для объединения значимостей (*p-value*) для ассоциации одной цис-регуляторной последовательности с изменением экспрессии генов во всех экспериментах. Мы использовали метод Фишера (сложный критерий Фишера) для расчета объединенных значимостей по каждому предсказанному цис-регуляторному элементу гена (*meta p-value*) (Cherenkov, Novikova et al., 2018). Тест на множественное тестирование выполняет функция *calcMetaAssociation*, метод коррекции может быть выбран из следующих: Бонферрони (Bonferroni), Холм-Бонферрони (Bonferroni-Holm), Бенджамини-Хохберг (Benjamini-Hochberg) и Бенджамини-Якутели (Benjamini-Yakuteli). Также пользователь может

задать порог для скорректированной объединенной значимости (meta p -value) ассоциации предсказанных цис-регуляторных последовательностей генов с дифференциальной экспрессией генов.

(5) Перестановочный тест

На последнем этапе в пакете программ metaRE применяется перестановочный тест к цис-элементам с объединенной значимостью (meta p -value), прошедшим пороговое значение. В пакете программ metaRE используется встроенный пакет *doParallel foreach* (CRAN project; <https://cran.r-project.org>) для параллельного выполнения перестановочного теста. Также в программу встроен калькулятор для установки количества перестановок M . Функция *PermutationTest* перемешивает регуляторные районы между генами и пересчитывает объединенные значимости обогащения (meta p -value) каждого цис-элемента. После M перестановок *the function* рассчитывает перестановочную значимость обогащения (permutation p -value) следующим образом: $p = (m + 1) / (M + 1)$, где m количество значимостей (p -values), не превышающих по значению объединенную значимость (meta p -value).

В качестве входных данных в пакете программ metaRE могут быть использованы не только промоторы генов, но и любые другие интересующие последовательности ДНК, например, районы связывания ТФ или участки генома с определенном типом хроматина.

Обогащение цис-элементов в районах связывания ТФ ARF

Для анализа обогащения потенциальных AuxRE в районах связывания ТФ ARF были использованы следующие данные: ARF6 ChIP-seq (Oh et al., 2014; GSM1252254), ARF2 и ARF5 DAP-seq (DNA affinity purification sequencing) (O'Malley et al., 2016; GSM1925138,

GSM1925826 и GSM1925827). Данные DAP-seq были процессированы автором согласно стандартному протоколу (Bardet et al., 2011), процессинг данных ТФ ARF6 провел Ощепков Д.Ю. (ИЦиГ СО РАН). В пиках районов связывания ТФ ARF было проведено картирование потенциальных цис-регуляторных элементов, аналогично процедуре на промоторах (см. выше). Для каждого цис-элемента рассчитывалась ассоциация по модифицированной таблице сопряженности с применением теста Фишера для анализа следующих пропорций: (i) количество положений, занимаемых цис-элементом среди всех 21 098 взятых в анализ промоторов *A.thaliana* (размером 1500 п.н.) к количеству всевозможных положений; (ii) то же самое среди всех пиков связывания ТФ ARF, относящихся к одному набору данных.

Поиск кластеров предсказанных цис-элементов

Для определения кластеризации предсказанных цис-регуляторных элементов мы провели их картирование на промоторные последовательности генов *Arabidopsis*, взятые из базы данных TAIR (сборка генома TAIR10; <https://www.arabidopsis.org>). При картировании парных AuxRE для каждого промотора мы брали расположенный дистально парный AuxRE (самый дальний от сайта старта транскрипции; ССТ) и проверяли есть ли расположенные проксимальнее от него парные элементы на расстоянии от шести до пятнадцати п.н. относительно последнего нуклеотида первого парного элемента, причем было задано условие минимального пересечения парных элементов в одну п.н.

Аннотация предсказанных цис-элементов

Для аннотации предсказанных цис-регуляторных элементов мы использовали он-лайн инструмент ТОМТОМ (<http://meme-suite.org>; Gupta et al., 2007). Сходство последовательностей предсказанных

цис-элементов с известными сайтами связывания ТФ определялось по коэффициенту корреляции Пирсона с порогом значимости меньше 0,05 ($E\text{-value} < 0,05$).

2.3. Растительный материал

В экспериментах использовались растения *A.thaliana* экотипа *Columbia-0*. Растения выращивались на твердой среде $\frac{1}{2}$ Murashige-Skoog, ($\frac{1}{2}$ MS; “Sigma”, США) или в грунте при температуре 22°C, интенсивности освещения 3-5 тыс. люкс и 16/8-часовом фотопериоде. Стерилизация семян производилась путем их инкубации в 70% спирте три раза по три минуты, а затем в стерильной воде также три раза по три минуты. После каждой инкубации в спирте или воде мы отбирали раствор и без перерыва добавляли новый. Для синхронизации прорастания мы подвергали семена яровизации от двух до семи дней при температуре 4°C (Rivero et al., 2014).

2.4. Подготовка РНК-образцов

Выделение тотальной РНК из тканей проводили с использованием реагента TRIZOL (“Invitrogen”, США):

- предварительно охлажденные в жидком азоте ткани растений разрушали в гомогенизаторе TissueLyser II (“Qiagen”, США);
- переносили полученный материал в пластиковую пробирку с реагентом TRIZOL, который инактивирует ферменты, в том числе и РНКазы;
- проводили разделение белков и РНК с помощью добавления хлороформа, интенсивного перемешивания и центрифугирования при 4°C в течение 10 минут;

- к отобранной водной фазе добавляли изопропанол, перемешивали и переносили полученный раствор на колонки из набора реагентов “RNA easy” (“Qiagen”, США);
- выделяли РНК при помощи набора реагентов “RNA easy” (“Qiagen”, США).

Концентрацию выделенной РНК измеряли на спектрофлуориметре NanoDrop (“Thermo Fisher Scientific Inc.”, США).

2.5. Определение уровня экспрессии генов методом

ОТ-количественной ПЦР

Реакция обратной транскрипции осуществлялась с использованием набора реактивов “RevertAid First Strand cDNA Synthesis Kit” (“Thermo Fisher Scientific Inc.”, США). ОТ-количественную ПЦР проводили с интеркалирующими флюорофорами и реакционной смесью SYBR Green qPCR Real Time Master Mix (“Thermo Fisher Scientific”, США) на амплификаторах CFX96 и CFX384 (“BioRad”, США) в соответствии со стандартной методикой (Nolan et al., 2006). Все количественные измерения проводились относительно отработанного на организме и тканях референсного гена *ACTIN2* (Nolan et al., 2006). Для каждого образца измерения проводились в трех биологических и трех технических повторностях, для каждой реакционной смеси был включен положительный и отрицательный контроль с каждой парой праймеров (Nodine et al., 2007). Изначально проводили тестовое измерение изменения количества мРНК 40 отобранных генов (Приложение 1), содержащих выявленные в ходе биоинформатического анализа цис-регуляторные элементы, в ответ на обработку ауксином в раннем (воздействие гормона длительностью до 1 ч.) ответе, по результатам

которого проводили более детальное измерение экспрессии генов после обработки растений ауксином в течение разного времени (длительность обработки: 15, 30, 60, 180 и 360 мин.).

2.6. Метод безлигазного клонирования

Для клонирования промоторов генов в работе был использован метод безлигазного клонирования, который позволяет в короткие сроки осуществить клонирование интересующего фрагмента, при этом эффективность реакции превышает 70%.

В методологии клонирования применяется $3' \rightarrow 5'$ экзонуклеазная активность T4 ДНК-полимеразы для приготовления векторов и вставок со специфичными длинными липкими концами (адаптерами) (Приложение 2), которые легко отжигаются и напрямую трансформируются в бактерию без необходимости ферментативного лигирования (Wendrich et al., 2015).

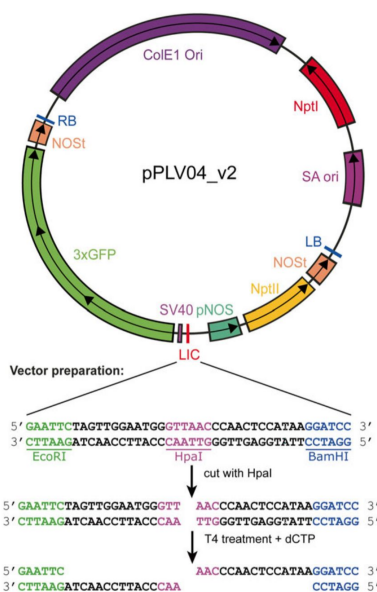


Рисунок 3. Схема плазмидного вектора pPLV04v2, использованного в данной работе, и графическое представление приготовления вектора. Адаптировано из Wendrich et al., 2015.

В работе использовался вектор pPLV04v2 (Рисунок 3) и штамм *Escherichia coli* DH5 α (Taylor et al., 1993). Вектор содержит следующие составляющие: ColE1 Ori – ориджин репликации высокой копийности в бактерии *E. coli*; NptI – кассета бактериальной устойчивости к канамицину; SA or – ориджин репликации низкой копийности в агробактерии; LB/RB – левый и правый сайты для переноса ДНК в геном растения; NOST – NOS сайт терминации транскрипции; pNOS – NOS промотор для экспрессии в растениях; NptII – кассета устойчивости растений к канамицину; SV40 – сигнал локализации в ядре; 3 \times GFP – кассета из трех последовательностей, кодирующих зеленый флуоресцентный белок для усиленной детекции; LIC – сайт безлигазного клонирования исследуемого фрагмента ДНК.

2.7. Мутагенез методом ПЦР с использованием перекрывающихся праймеров

Метод мутагенеза путем ПЦР с использованием перекрывающихся праймеров предоставляет возможность быстро и с экономией средств получить мутации в интересующих позициях последовательности ДНК (Saiki et al., 1985). Метод включает в себя следующие стадии (Рисунок 4): (1,2) амплификация двух мутированных фрагментов интересующей последовательности, в рамках этой работы – промоторов с предсказанным цис-регуляторными элементами, (3) амплификация мутированной последовательности промотора по матрице двух фрагментов, полученных на стадиях (1) и (2) (Ho et al., 1989).

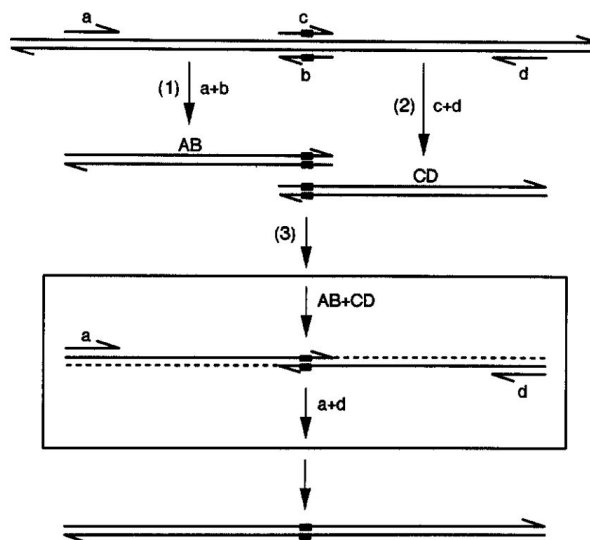


Рисунок 4. Схема ПЦР с перекрывающимися праймерами. Адаптировано из Ho et al., 1989.

Для проведения ПЦР подбираются фланкирующие праймеры и праймеры, несущие мутацию. На первом этапе в двух отдельных ПЦР проводится амплификация с одним фланкирующим праймером и одним праймером с мутацией, таким образом, в результате первого раунда ПЦР мы имеем в наличии два фрагмента мутируемой последовательности. Затем эти два фрагмента используются как матрица во второй ПЦР за счет денатурации и последующего комплементарного соединения перекрывающихся концов и мутированных участков, амплификация проводится с использованием фланкирующих праймеров (Рисунок 4).

Для проверки функциональности отобранных парных AuxRE в ответе на ауксин были сделаны следующие генетические конструкции: конструкция с интактным промотором, конструкция с мутированной последовательностью TGTCNN, конструкция с мутированным элементом-партнером и конструкция с полностью мутированным парным AuxRE (при необходимости). Длина промотора в разных генетических

конструкциях варьировала от 1000 до 3000 п.н. в зависимости от расстояния до предыдущего гена. Все встроенные фрагменты в полученных конструкциях были верифицированы секвенированием (Приложение 3).

2.8. Трансформация растений с помощью агробактерии

Для трансформации растений был использован агробактериальный метод, в котором в качестве переносчика ДНК используется бактерия *Agrobacterium tumefaciens*. Метод основан на способности *Agrobacterium* переносить из своей Ti-плазмиды в растительные клетки фрагменты ДНК, в том числе и гены (Chilton et.al., 1977). Ti-плазмиды могут выступать в качестве биологического вектора для переноса фрагментов ДНК размером до 50 тыс. п.н., так как обладают набором генов вирулентности, консервативными нуклеотидными прямыми повторами, фланкирующими T-ДНК (район T-ДНК может быть заменен на интересующий исследователей ген) и ориджин репликации Ti-плазмиды внутри агробактерии. Искусственно усовершенствованные плазмиды несут гены резистентности к антибиотику (Bevan et.al., 1983). Бинарная векторная система состоит из двух векторов: Ti-плазмиды-помощника и вектора, автономно реплицирующегося как в *E.coli*, так и в агробактерии. Агробактериальный способ обычно приводит к встраиванию только одной или, по крайней мере, ограниченного числа копий гена и при этом, видимо, в определенные локусы хромосом, а другие методы к бессистемному встраиванию множества копий, которые в большинстве случаев сцеплены друг с другом (De Block, 1993). В данной работе была использована бинарная векторная система pGreen/pSoup (Hellens et al., 2000).

2.9. Выявление взаимодействий белок-ДНК в дрожжевой одногибридной системе

Для поиска ТФ, связывающихся с предсказанными цис-элементами, была использована одногибридная система на основе дрожжей *Saccharomyces cerevisiae* (Gaudinier et al., 2011). Базовый компонент дрожжевой одногибридной системы – это активатор транскрипции GAL4, эндогенно экспрессируемый белок, содержащий активирующий транскрипцию домен (активирующий домен, АД). Для проверки взаимодействия между белком и ДНК (в нашем случае – участком промотора гена) создают следующие два типа химер: акцептор, содержащий исследуемую ДНК-последовательность вместе с репортерным геном, и донор, состоящий из АД и ТФ. При взаимодействии акцептора с донором АД запускает транскрипцию репортерного гена (Рисунок 5), в нашей системе – *HIS3*, ответственного за синтез гистидина.

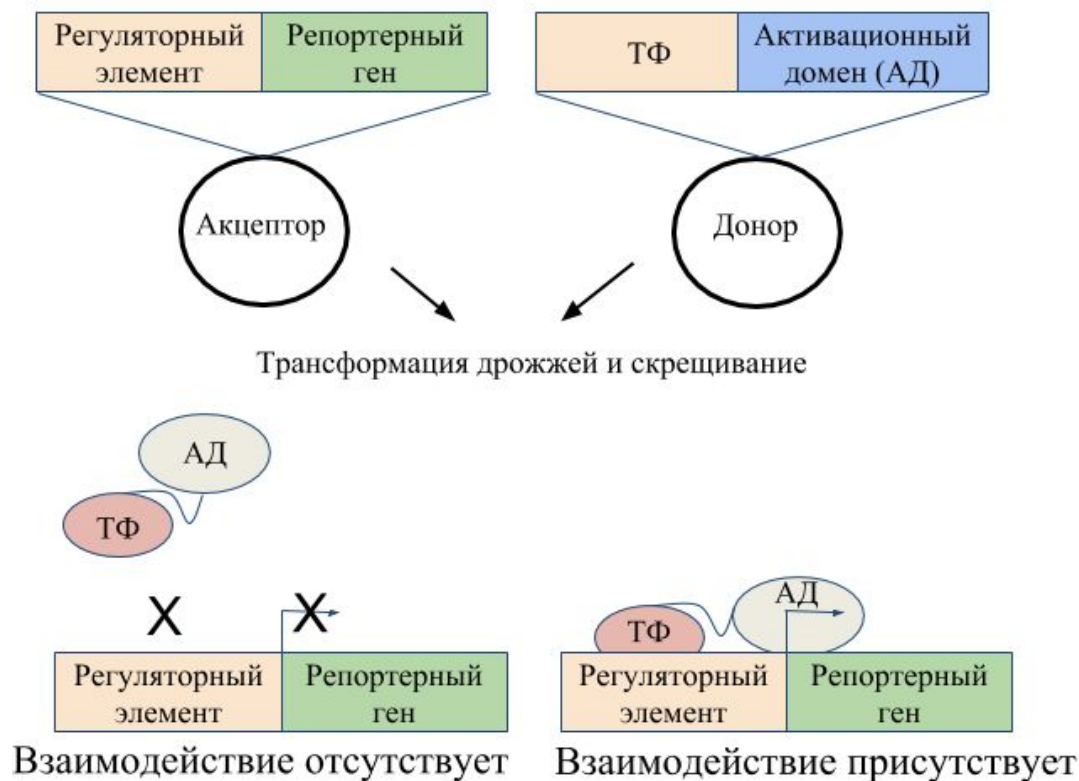


Рисунок 5. Схема дрожжевой одногибридной системы для выявления взаимодействий белок-ДНК.

В данной работе использовались клеточные линии дрожжей PJ69-4A (*MATa trp-901 leu2-3, 112 ura3-52 his3-200 gal4D gal80D GALADE2 LYS::GAL1-HIS3 met::GAL7-lac Z*) и PJ69-4α (*MATa, ade2, trp1-109, leu2-3, 112, ura3-52, his3-200, gal4, gal80, GAL2::ADE2, LYS2::GAL1:HIS3, met2::GAL7:lacZ*) (James et al., 1996). Клонирование троекратного повтора композиционного элемента осуществлялось с помощью набора Gateway Cloning (“Life Technologies“, США) в плазмиды pMW2 и pMW3 (Bass et al., 2017). Трансформация дрожжей производилась по стандартному протоколу (Kawai et al., 2010). Затем дрожжи инкубировали при 30°C в

течении 3-4 дней, после чего проводили их скрещивание и селекцию гибридных колоний с наличием ДНК-белковых взаимодействий на средах разной рестрикции. Для скрининга взаимодействий с ТФ была использована библиотека ТФ *Arabidopsis*, состоящая из 1975 белков (Mitsuda et al., 2010). Плазмиды pMW2 и pMW3 включают гены синтеза гистидина и урацила, соответственно, а pDEST22 – триптофана, таким образом, гибридные колонии способны расти на среде без вышеназванных веществ, следовательно, среда без гистидина, урацила и триптофана (-HIS/URA/TRP) является наименее рестриктивной, и рост дрожжей на ней свидетельствует о наличии размножения и нормальной работе системы. Далее дрожжи высевались на более рестриктивную среду с добавлением 3-АТ (3-амино-триазола), подавляющего синтез гистидина. 3-АТ добавляли в разных концентрациях: 1, 3, 5, 10, 20, 40, 60 и 80 мМ, для полного подавления сигнала (роста колоний) в отрицательном контроле (дрожжах, несущих троекратный повтор парного элемента и “пустую” плазмиду pDEST22, использованную для клонирования ТФ из библиотеки).

2.10. Конфокальная микроскопия

Для изучения паттерна и интенсивности экспрессии белка GFP в репортерных линиях растений использовались микроскопы LSM 510 Meta (Carl Zeiss) и Leica SP8 (Leica Microsystems). Наборы изображений для количественного анализа и сравнений были сняты при одинаковых параметрах съемки для каждой репортерной линии.

ГЛАВА 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ

3.1. Биоинформатический анализ данных

3.1.1. Метод транскриптомного анализа ассоциаций (ТАА)

Сайты связывания большинства ТФ остаются неизученными даже у популярных модельных организмов. Кроме того, количество организмов, для которых существуют модели позиционно-весовых матриц сайтов связывания хотя бы сотни ТФ весьма ограничен (*Saccharomyces cerevisiae*, *Arabidopsis thaliana*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Mus musculus*, *Homo sapiens*). Последовательности парных элементов могут сильно отличаться от индивидуальных сайтов связывания ТФ, поэтому изучение и описание разнообразия одиночных и композиционных сайтов связывания ТФ для разных организмов остается актуальной задачей.

В рамках данной работы был создан метод транскриптомного анализа ассоциаций (ТАА) для предсказания цис-регуляторных элементов, систематически перепредставленных в регуляторных районах дифференциально-экспрессирующихся генов, взятых из серии транскриптомных данных. Метод основан на полном переборе консенсусных символьных последовательностей ДНК и анализе значимости перепредставленности каждого цис-элемента в каждом наборе данных из серии. Такой метод позволяет проводить анализ обогащения всех возможных последовательностей и хорош для описания разнообразия мотивов, характеризующих регуляторные районы генов, изменяющих свою экспрессию при разных условиях (в ответ на стимул, в

мутированных линиях, на разных стадиях развития). Метод направлен на поиск известных и новых цис-элементов, но при этом не учитывает вырожденность сайтов связывания ТФ.

В случае оценки статистической значимости множества последовательностей одновременно, необходимо делать поправку на множественное сравнение, поэтому необходимо соблюсти баланс между желанием оценить как можно больше мотивов и емкостью метода оценки значимости. Для соблюдения этого баланса, было решено принять за идентичные обратно комплементарные последовательности потенциальных цис-элементов их прямым последовательностям. В случае поиска обогащения гексамерами вместо $2^6=4096$ возможных последовательностей осуществлялся поиск 2080, состоящих из 2016-ти комплементарных и 64-х палиндромных последовательностей.

При использовании поправки Бонферрони на множественное сравнение метод ТАА позволяет эффективно находить перепредставленные гексамеры и гептамеры, а также парные элементы, в которых одна последовательность фиксирована, а ее элементом-партнером может быть любой, полученный путем полного перебора возможных вариантов, образованных четырехнуклеотидным алфавитом, также возможно варьирование спейсера в парных элементах.

Обогащение потенциальных цис-регуляторных элементов в промоторах ДЭГ сначала оценивается для каждого отдельного цис-элемента в каждом отдельном наборе данных из одного транскриптомного эксперимента по точному тесту Фишера. Затем значимости (p -value) обогащения отдельного цис-регуляторного элемента по данным разных экспериментов объединяются методом Фишера (см. раздел 2.2). И наконец, для выбора цис-элементов, значимо

ассоциированных с ответом на исследуемый фактор, проводится поправка на множественное сравнение и выбираются элементы с скорректированной значимостью (p -value adj.) выше порогового значения.

Идея метода ТАА не нова, и отдельные подходы метода уже были реализованы в разных алгоритмах и программах, чаще на одном организме (van Helden et al., 1998; van Helden et al., 2000; He et al., 2016). Новизной метода ТАА в сравнении с аналогами было: (1) учет большого количества транскриптомных экспериментов, и (2) способ учета значимости обогащения для отдельного цис-элемента. Обычно, авторы оценивают обогащение цис-элементов в одном или нескольких списках генов независимо, результаты обогащения между списками не сравниваются (например, Bargmann et al., 2013; Berendzen et al., 2012). В этом случае информация о различиях в степени обогащения одного и того же цис-элемента в разных наборах данных нивелируются, что может приводить к пере- и недо-предсказаниям. Метод ТАА решает эту проблему.

К достоинствам предложенного метода ТАА, и разработанного на его основе пакета программ metaRE (<https://github.com/cheburechko/metaRE>) является то, что они применимы к любым организмам и не ограничены поиском обогащения мотивов в регуляторных последовательностях ДНК. Изначально метод разрабатывался для поиска обогащения цис-регуляторных элементов в промоторах ДЭГ, но также в него заложена опция поиска обогащения цис-элементов в районах связывания ТФ и регуляторных районах генов с определенным типом хроматина (Cherenkov, Novikova et al., 2018).

На основе предложенного метода ТАА в группе Мироновой В.В. был создан пакет программ metaRE (см. методы), который включает в себя пять ступеней анализа: поиск ДЭГ в нескольких транскриптомных экспериментах; консенсусный поиск цис-регуляторных элементов заданной длины и структуры в последовательностях ДНК; расчет ассоциации между консенсусным присутствием элемента и изменением экспрессии генов; мета-анализ ассоциации на нескольких экспериментах по изменению экспрессии генов; перестановочный тест для проверки независимости результата от сторонних факторов. Далее описана апробация метода ТАА и пакета программ metaRE в поиске ауксин-чувствительных элементов (AuxRE).

3.1.2. Поиск и функциональная аннотация простых ауксин-чувствительных элементов в промоторах генов

A.thaliana

В результате биоинформатического анализа транскриптомных данных (Таблица 1) с помощью метода ТАА, реализованного в пакете программ metaRE, было идентифицировано 27 и 140 одиночных AuxRE, достоверно ассоциированных с ранним и поздним ответом на ауксин. Из них в раннем ответе 24 одиночных AuxRE ассоциированы с повышением транскрипционной активности генов, три – понижением; в позднем ответе 78 – с повышением, 121 – с понижением (Таблица 3).

Среди них было обнаружено обогащение классического ARF-связывающего AuxRE TGTCTC в промоторах генов, экспрессия которых активировалась ауксином в раннем ответе, что согласуется с предыдущими исследованиями (Ulmasov et al., 1995) и может рассматриваться как свидетельство корректности биоинформатического

метода ТАА. Также мы обнаружили перепредставленность последовательности TGTGGG, которая может быть нетипичным сайтом связывания белков ARF.

Таблица 3. Сводная таблица по выявленным простым AuxRE, обогащенным в промоторах ауксин-чувствительных генов.

Изменение экспрессии	Ранний ответ (<=2ч)	Поздний ответ (>2ч)
активация	24*	78
ингибирование	3	121
(1) Из них, без A/T-богатых цис-элементов		
активация	8*	16
ингибирование	3	26
(2) из них, специфичные по времени ответа		
активация	6*	18
ингибирование	0	59
(3) из них, обогащенные в районах связывания ARF		
активация	11*	25
ингибирование	3	37

Примечание: * включая ARF-связывающий цис-элемент TGTCTC

Аннотация предсказанных одиночных AuxRE осуществлялась при помощи анализа литературы и инструмента ТОМТОМ (<http://meme-suite.org>; Gupta et al., 2007). В последнем реализуется поиск сходства заданной пользователем последовательности нуклеотидов с тремя базами данных сайтов связывания ТФ: DAP-seq (O'Malley et al., 2016), PBM (Franco-Zorrilla et al., 2014) и CIS-BP (Weirauch et al., 2014). Сходство цис-элементов определялось по коэффициенту корреляции

Пирсона с пороговым значением значимости меньше 0,05 ($E\text{-value} < 0,05$), рассчитываемой в программе ТОМТОМ для каждого распознанного сайта связывания ТФ.

Интересно, что наиболее достоверно ассоциированными с ответом на ауксин оказались не классические ARF-связывающие AuxRE. ТАТА-подобные гексамеры (ТАТААА, ТАТАТА, АТАТАТ, АТАТАС, АТАТАГ, АТАТАА), bZIP-связывающие сайты (АСГТАТ) и bHLH-связывающие E-бокс (САСАТГ) были ассоциированы с положительной регуляцией ауксином больше, чем последовательность TGTCTC (по значениям *meta p-value* и *permutation p-value*). Связь цис-элементов ТАТАТА и САСАТГ с ранним ответом на ауксин была показана ранее (Trenner et al., 2016).

В случае отрицательной регуляции ауксином были выявлены потенциальные bZIP-связывающие цис-элементы (АСГТГТ, АСГТГГ) и MYB-связывающий I-бокс сайт (GATAAG). Связь ТФ bZIP и MYB с регуляцией ауксином была также продемонстрирована ранее (Weiste and Droge-Laser et al., 2014), однако на данный момент не было выявлено ни одного AuxRE, который бы обеспечивал репрессию в ответ на ауксин.

А/Т-богатые одиночные AuxRE оказались наиболее распространенными и достоверно связанными с ответом на ауксин в нашем исследовании (Таблица 3, 4). Вместе с ТАТА-подобными последовательностями, две третьих предсказанных одиночных AuxRE оказались А/Т-богатыми. У *Arabidopsis* многие ТФ связываются с цис-регуляторными элементами, включающими более пяти нуклеотидов А/Т подряд: REM, ABI3-VP1, C2H2, HB, Homeobox, MYB (O'Malley et al., 2016). Возможно, какие-то из этих семейств ТФ действительно вносят большой вклад в транскрипционный ответ на ауксин, но на данный

момент это неизвестно. Также перепредставленность А/Т-богатых цис-регуляторных последовательностей в результатах нашего анализа может быть связана с их ролью в эпигенетической регуляции ответа на ауксин. Этот вопрос требует дальнейшего анализа.

Таблица 4. Аннотация выявленных одиночных AuxRE, связанных с ранним ответом на ауксин.

Одиночный AuxRE	Анализ транскриптома				Chip(DAP)-seq анализ			Описание (из литературных данных)
	Ранний ответ (<=2ч.)		Поздний ответ (>2ч.)		Обогащение в областях связывания с ТФ ARF			
	+	-	+	-	2	5	6	
Известные и потенциальные сайты связывания белка ARF								
TGTCTC	**				***	***	***	классический AuxRE (Ulmasov et al., 1997).
TGTCCC	***				***	***	***	AuxRE (Xu et al., 1997; Weiste and Droge-Laser, 2014).
GTCCCC	**				***	***	***	потенциальный AuxRE или сайт связывания белка TCP.
TGTGGG	***						***	
bHLH- и BZR1/BES1-связывающие сайты								
CACATG	***			***	***	***	***	HUD мотив, обогащен в районе связывания ТФ ARF6 (Oh et al., 2014).
Потенциальный сайт связывания ТФ MYB								
GATAAG	***		**				***	Коровый сайт связывания ТФMYB, I-бокс (Rose et al., 1999).
Известные и потенциальные сайты связывания ТФ bZIP								
TACGTA	**		***	***			***	А-бокс, связывается с ТФ bZIP (Izawa et al., 1993).
ACGTAT	***		***	***			***	А-бокс-подобный
ACGTAG	**						***	
ACGTGT		***	***	***		**	***	G-бокс-подобный, ABRE, сайт связывания ТФ AREB/ABF (Yamaguchi-Shinozaki and Shinozaki, 2005).
ACGTGG		***		***			***	
А/Т-богатые последовательности								

TATAAA	***	***	***			Классический ТАТА-бокс (Heard et al., 1993)
TATATA, ATATAT, ATATAC, ATATAG	***	***	***			ТАТА-подобный (Yamamoto et al., 2009). Обогащен рядом со стартом транскрипции.
ATATAA	***	***	***	**		
AACATT	**			***	***	Неизвестные А/Т-богатые последовательности
CATAAT, GATTA	***	***	***			
ACATA, TATTA	**	***	***			
ATTAGA	**	***				
AAATAC						
CATATT	**		***			
CATTAT	**	***	***		*	
TAATTA	**	***	***			Потенциальный АТНВ-связывающий сайт.

Примечание: ≤ 2 ч. – ранний транскрипционный ответ до двух часов обработки ауксином включительно, > 2 ч. – поздний ответ при обработке ауксином длительностью более двух часов, “+” – активация транскрипции, “-” – подавление транскрипции, 2/5/6 – ТФ ARF2/ARF5/ARF6, Chip(DAP)-seq – использованные в анализе полногеномные эксперименты по связыванию ТФ с ДНК, ***/**/* – уровень значимости (*p*-value) меньше 0,001/0,01/0,05, соответственно.

Ингибирование транскрипционной активности генов в ответ на ауксин остается малоизученным, в данном исследовании были выявлены 127 одиночных AuxRE, связанных с подавлением экспрессии ауксином, из них только три представлены в промоторах генов раннего ответа и 124 – позднего. Более 100 одиночных AuxRE, связанных с подавлением экспрессии ауксином и поздним ответом, оказались А/Т-богатыми. Многие последовательности были выявлены как связанные с активацией и подавлением экспрессии генов одновременно. Например, bZIP-связывающий сайт ACGTAT обогащен в промоторах генов,

транскрипционная активность которых повышается в ответ на ауксин в раннем ответе, и промоторах генов, транскрипционная активность которых подавляется в ответ на ауксин в раннем и позднем ответе (Таблица 4).

Таким образом, в промоторах ауксин-чувствительных генов было выявлено обогащение сайтами связывания ТФ bZip, MYB, bHLH и некоторыми другими. Помимо известных сайтов связывания ТФ, неизвестные последовательности оказались ассоциированы с ответом на ауксин, их роль в регуляции транскрипции ауксином еще предстоит выяснить. Впервые была продемонстрирована взаимосвязь А/Т-богатых последовательностей с ответом на ауксин, большинство из которых также не были аннотированы как сайты связывания ТФ.

3.1.3. Обогащение предсказанных одиночных AuxRE в районах связывания ТФ ARF

Ранее было показано, что ТФ ARF могут гетеродимеризоваться с другими ТФ (Shin et al., 2007; Varaud et al., 2011; Oh et al., 2014). Возможно, что некоторые из выявленных в нашем исследовании одиночных AuxRE, ассоциированных с ответом на ауксин, являются сайтами связывания этих ТФ. Для проверки этой гипотезы мы провели анализ представленности выявленных одиночных AuxRE в районах связывания белков ARF2/5/6, взятых из данных экспериментов DAP-seq (O'Malley et al., 2016) и ChIP-seq (Oh et al., 2014) (см. методы). Анализ был осуществлен при помощи пакета программ metaRE. Для каждого цис-элемента рассчитывалась ассоциация по точному тесту Фишера, которым сравнивалась частота встречаемости предсказанной последовательности в промоторах исследуемых генов с его частотой

встречаемости в пиках связывания одного из ТФ ARF по отношению ко всем возможным положениям. Полученные уровни значимости обогащения цис-регуляторных последовательностей в районах связывания ТФ ARF (p -value) были скорректированы по Бонферрони.

Действительно, в районах связывания ТФ ARF 2/5/6 оказались перепредставлены ARF-связывающие сайты TGTCTC и TGTCCC, что подтверждает адекватность использованного метода. Не А/Т-богатые одиночные AuxRE, ассоциированные с ранним ответом на ауксин, были обогащены в районах связывания ТФ ARF6; также сайты связывания ТФ bHLH (CACATG) и ТФ bZIP (ACGTGT) обогащены в районах связывания ТФ ARF5 (Таблица 4).

Большая часть не А/Т-богатых одиночных AuxRE, ассоциированных с поздним ответом на ауксин, была обогащена в районах связывания хотя бы одного из ТФ ARF, за исключением потенциального сайта связывания ТФ EIN3 (ATGTA[T/C]) (Kosugi and Ohashi, 2000), сайта AACCCCT и других семи не описанных ранее последовательностей. Полученные результаты подтверждают продемонстрированное ранее обогащение сайтов связывания ТФ bZIP, MYB, и bHLH рядом с TGTC-содержащими AuxRE (Shin et al., 2007; Berendzen et al., 2012; Walcher and Nemhauser, 2012; Oh et al., 2014).

Некоторые А/Т-богатые последовательности также оказались обогащены в районах связывания ТФ ARF (Рисунок 6). Так первая группа А/Т-богатых гексамеров была обогащена в районах связывания ТФ ARF2/5. При выравнивании этих цис-элементов относительно друг друга они составили сайт связывания ТФ YAB5 (TAAACAT[A/T][A/T]) (ТОМТОМ E -value<0,05). Входящая в его состав последовательность AACATT связана с ранней активацией экспрессии генов в ответ на

ауксин. Вторая группа поли-А/Т последовательностей оказалась обогащена в районах связывания белка ARF2 (Рисунок 6). Из этого можно предположить, что у ТФ ARF2 есть ТФ-партнер с А/Т-богатым сайтом связывания или для взаимодействия ТФ ARF2 с ДНК необходимо определенное состояние хроматина.

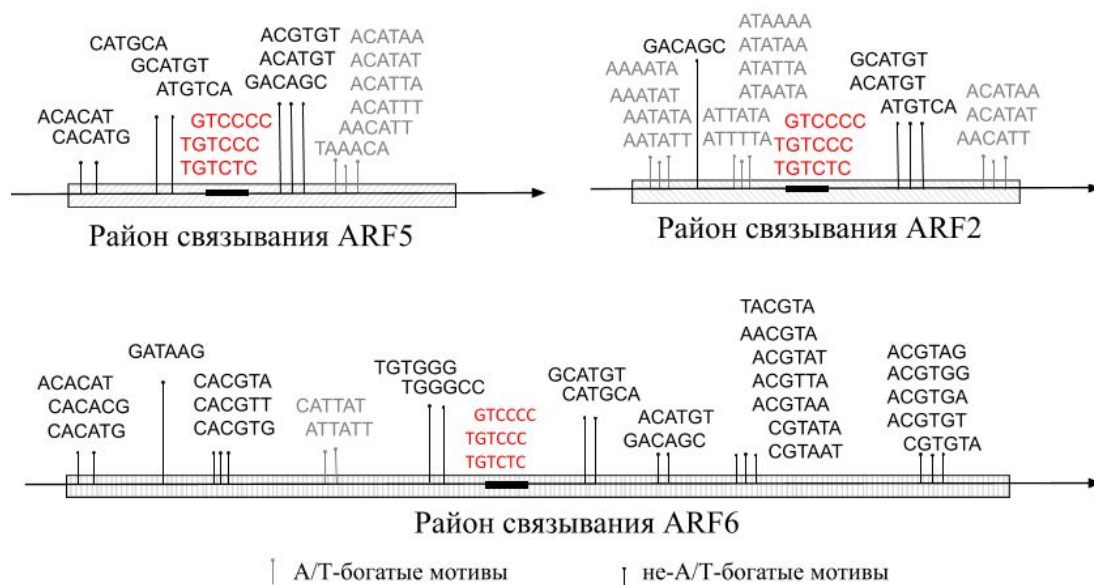


Рисунок 6. Схема представленности одиночных AuxRE в районах связывания ТФ ARF. Все перечисленные на схеме одиночные AuxRE были значимо перепредставлены (p -value adj.<0.05) в районах связывания ARF2/5/6 по данным DAP(Chip)-Seq. Красным шрифтом обозначены ARF-связывающие AuxRE, серым – предсказанные А/Т-богатые цис-регуляторные элементы, черным – все остальные предсказанные одиночные AuxRE.

В анализ были взяты данные DAP-seq только трех ТФ ARF, которые не отображают районы всех представителей семейства белков ARF, таким образом, не выявленные обогащенными в районах связывания ТФ

ARF2/5/6 предсказанные одиночные AuxRE могут быть партнерами для других белков ARF или посредниками для связывания ТФ ARF с ДНК.

3.1.4. Поиск и функциональная аннотация парных AuxRE в промоторах генов *A.thaliana*

В связи с описанными в предыдущей главе результатами возник вопрос: есть ли в промоторах ауксин-чувствительных генов обогащение парными AuxRE определенной структуры, состоящими из последовательности TGTCNN и элемента-партнера. Так как в предыдущем анализе (раздел 3.1.2) не учитывалось расстояние и расположение последовательности TGTCNN относительно ее потенциального элемента-партнера, то описание разнообразия парных AuxRE со спейсером определенной длины, обогащенных в промоторах ауксин-чувствительных генов, стало следующим этапом исследования. При помощи метода ТАА, реализованного в пакете программ metaRE, был проведен поиск обогащения потенциальных парных AuxRE в промоторах ауксин-чувствительных генов на транскриптомных данных (Таблица 1). В поиске парных AuxRE учитывалась ориентация элемента-партнера относительно мотива TGTCNN, была задана длина спейсера от 0 до 15 п.н. Всего было проверено 31200 комбинаций парных цис-регуляторных элементов, из которых обогащенные были отобраны по порогу скорректированной обобщенной значимости меньше $1E-5$ (meta p -value adj. $<1E-5$).

В результате биоинформатического анализа было выявлено 592 разных потенциальных парных AuxRE (meta p -value adj. $<1E-5$) (Рисунок 7), среди них 579 обогащены в промоторах генов, транскрипция которых активируется в ответ на ауксин, 561 – в промоторах генов, транскрипция

которых подавляется в ответ на ауксин. Таким образом, 13 парных AuxRE встречались исключительно в промоторах активируемых генов, 23 – подавляемых. Среди предсказанных элементов-партнеров последовательности TGTCNN был выявлен 21 элемент, найденный ранее в работе Cherenkov, Novikova et al., 2018 как одиночный AuxRE. Причем эти последовательности встречаются в парных AuxRE, связанных и с активацией, и с ингибированием экспрессии в ответ на ауксин (Рисунок 7А). В пересечение также попал канонический ARF-связывающий AuxRE TGTCTC.

Всего же 2% (14 шт.) предсказанных парных AuxRE оказались представлены гомодимерами, состоящими из повторенной последовательности TGTCNN или последней в паре с TGTC-содержащим элементом (Рисунок 7Б). Такие парные AuxRE отличаются по ориентации (Рисунок 7А). Элементы-партнеры представлены тремя последовательностями TGTCCC, TGTCTC и TGTCTG, расположенными на разном расстоянии от последовательности TGTCNN (Рисунок 7Б). Интересно, что спейсеры между элементами в составе гомодимерного парного AuxRE представлены 1, 5, 6, 13, 15 п.н., хотя ранее продемонстрирована повышенная аффинность связывания белка ARF с повторами TGTCNN со спейсером 7-8 п.н. (Boer et al., 2014). В случае TGTC-содержащих элементов мы заметили, что некоторые из них являются одними и теми же элементами-партнерами со сдвигом в один нуклеотид (Рисунок 7В).

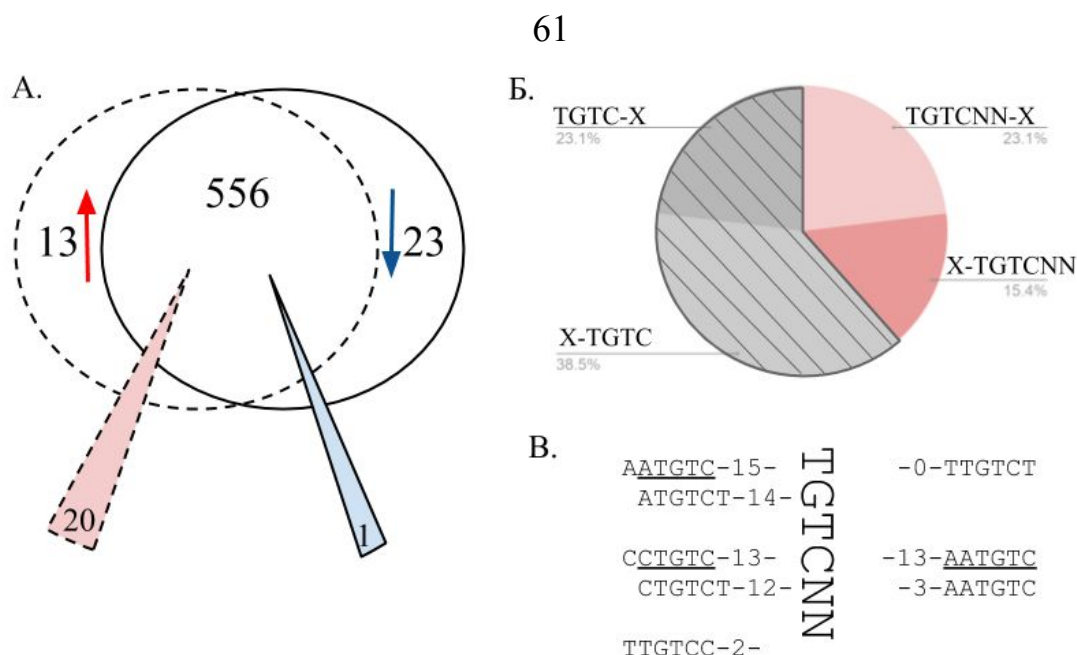


Рисунок 7. Выявленные парные AuxRE в промоторах генов *Arabidopsis*, ассоциированные с транскрипционным ответом на ауксин. А. Пересечение парных AuxRE, ассоциированных с положительной и отрицательной регуляцией активности генов, с одиночными AuxRE, выявленными в разделе 3.1.2 (Таблица 4), и выделенными в красный треугольник при ассоциации с активацией экспрессии генов, в голубой – с подавлением. Б. Аннотация выявленных гомодимерных парных AuxRE; серым отмечены парные AuxRE с TGTC-содержащим элементом-партнером, розовым – с последовательностью TGTCNN. В. TGTC-содержащие элементы-партнеры в составе парных AuxRE.

Для аннотации предсказанных элементов-партнеров в составе парных AuxRE была использована программа ТОМТОМ (<http://meme-suite.org>; Gupta et al., 2007). 7.3% предсказанных элементов-партнеров оказались известными сайтами связывания ТФ (E -value < 0,05) (Рисунок 8), представителей семейств ТФ AP2, bHLH, bZIP, C2H2, TCP, WRKY, MYB и др. Такой низкий процент

распознавания сайтов программой ТОМТОМ (<http://meme-suite.org>; Gupta et al., 2007) связан с тем, что эта программа узнает только наиболее приближенный к оптимальной последовательности сайта связывания мотив, пакет metaRE же находит все потенциальные цис-элементы, обогащенные в промоторах ауксин-чувствительных генов, многие из которых отличаются от последовательности оптимального сайта связывания ТФ или захватывают только его часть. Последнее значит, что несколько предсказанных пакетом программ metaRE цис-элементов зачастую представляют собой один цис-регуляторный элемент длиннее шести нуклеотидов, как, например, в случае YAB5-связывающего сайта (см. раздел 3.1.3). Данный вопрос будет подробнее обсуждаться в следующем разделе.

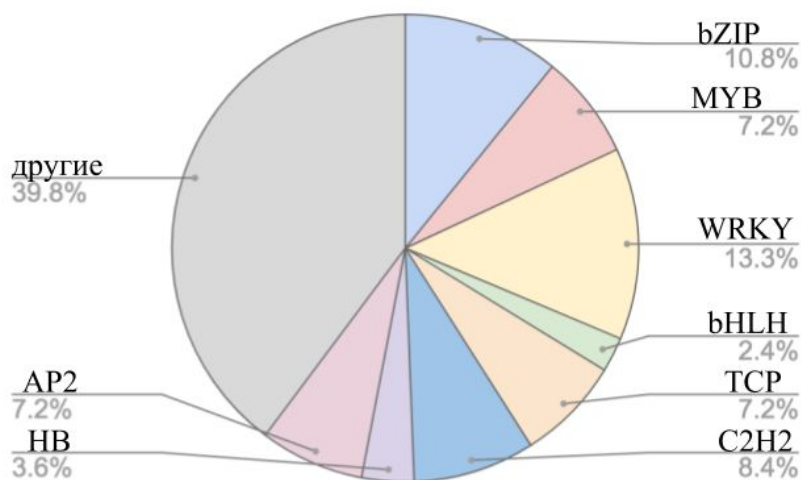


Рисунок 8. Аннотация предсказанных элементов-партнеров последовательности TGTCNN с помощью программы ТОМТОМ.

Опубликованные свидетельства участия ТФ бZIP, бHLH, MYB в передаче сигнала от ауксина обсуждались ранее (см. раздел 3.1.2). В рамках поиска парных AuxRE, эти сайты связывания ТФ были выявлены

как элементы-партнеры последовательности TGTCNN с определенной ориентацией и длиной спейсера, что является косвенным свидетельством их совместного функционирования на промоторах ауксин-чувствительных генов в виде парных AuxRE. 7,2% аннотированных программой TOMTOM сайтов относятся к сайтам связывания ТФ TCP. Ранее было продемонстрировано взаимодействие ТФ TCP4 и TCP18 с ТФ ARF3 (Simonini et al., 2016). Также известно, что ТФ TCP3 участвует в регуляции ответа на ауксин, его биосинтеза и транспорта (Nicolas and Cubas, 2016). ТФ TCP17 и его близкий гомолог, ТФ TCP13, регулируют биосинтез ауксина через PIF-зависимый и PIF-независимый пути, в последнем осуществляется регуляция белков YUCCA (Zhou et al., 2017).

Значительную долю аннотированных элементов-партнеров составили сайты связывания ТФ WRKY (13,3%). Для них также не было продемонстрировано взаимодействия с ТФ ARF, но известно, что ТФ WRKY23 влияет на распределение ауксина в корне через локальный контроль биосинтеза флавоноидов (Grunewald et al., 2012), в свою очередь регуляция транскрипционной активности ТФ WRKY23 осуществляется в зависимости от белков SLR-ARF7/ARF19 (Grunewald et al., 2008). Возможно, что участие представителей семейства ТФ WRKY в регуляции процессов ауксином – более общее явление, для подтверждения которого необходимы дальнейшие исследования.

Помимо перечисленных цис-регуляторных элементов, связывающихся с семействами ТФ, в ходе аннотации была выявлена разрозненная группа известных сайтов связывания ТФ (Таблица 5). Для большинства из ТФ, связывающихся с этими элементами-партнерами в составе парных AuxRE, взаимосвязь с ответом на ауксин ранее

продемонстрирована не была. Однако они участвуют в различных защитных реакциях, например, белок AHL20 подавляет защитный ответ на воздействие патогенных организмов у *Arabidopsis* (Lu et al., 2010); ТФ CBF3 участвует в регуляции ответа на холодовой стресс (Lee et al., 2005); ТФ NTL9 и DEL1 вовлечены в регуляцию процессов гормоном стресса – салициловой кислотой (Perrot-Rechenmann, 2010; Berckmans et al., 2011); ТФ HSFA1E является ключевым регулятором сигнального пути HSF в ответ на стресс (Nishizawa-Yokoi et al., 2011). ТФ SPL вовлечены в регуляцию многих процессов *Arabidopsis*, среди которых развитие побега, репродуктивный рост, толерантность к абиотическому стрессу (Xu et al., 2016). Связь с ауксином было показана только для белка STY1, который может активировать транскрипцию белков YUC4 и YUC8 через последовательность АСТТАС/А (Eklund et al., 2010).

Таблица 5. Аннотация выявленных элементов-партнеров в составе парных AuxRE.

Элемент-партнер TGTCNN	Парный AuxRE	Предсказание программы TOMTOM
GCCAAA	GCCAAA-0-TGTCNN	E2FC, DEL1, DEL2
CCATCA	TGTCNN-15-CCATCA	HAP3
TGTACG	TGTACG-11-TGTCNN	SPL3, SPL5, SPL1
CGTAGG	CGTAGG-8-TGTCNN TGTCNN-0-CGTAGG	TOE2
GTACGA	GTACGA-10-TGTCNN	TOE2, TOE1
TACGTG	TACGTG-14-TGTCNN	ANAC55
AGCTAG	TGTCNN-9-AGCTAG	STY1
CTAGGG	CTAGGG-15-TGTCNN	STY1
GACATA	GACATA-11-TGTCNN	CBF3
ATGCAT	TGTCNN-10-ATGCAT TGTCNN-8-ATGCAT	LEC2
AATTAА	AATTAА-1-TGTCNN	Высоко-мобильные белки (HMG (high mobility group box protein))
ATTTGA	ATTTGA-12-TGTCNN	WLIM2A
TAGAAG	TAGAAG-4-TGTCNN	HSFA1E
ATATAT	TGTCNN-10-ATATAT	AHL20
ATTAAT	TGTCNN-10-ATTAAT	ANL2
ACTTAA	TGTCNN-5-ACTTAA	NTL9
TTACTT	TTACTT-10-TGTCNN	NTL9
TGAATT	TGTCNN-1-TGAATT	TSO1-подобные белки (Tesmin/TSO1-like CXC domain-containing protein)
TTGAAT	TTGAAT-5-TGTCNN	TSO1-подобные белки (Tesmin/TSO1-like CXC domain-containing protein)
TTTAAA	TTTAAA-5-TGTCNN	TSO1-подобные белки (Tesmin/TSO1-like CXC domain-containing protein)

Таким образом, были идентифицированы парные AuxRE, достоверно ассоциированные ($\text{meta } p\text{-value} < 1\text{E-}5$) с ответом на ауксин. Среди элементов-партнеров в составе парных AuxRE были выявлены как известные сайты связывания ТФ, так и большое количество неизвестных последовательностей. Среди известных элементов-партнеров были выявлены сайты связывания ТФ, участие которых в регуляции ответа на ауксин в литературе не упоминалось. Многие из идентифицированных цис-элементов при выравнивании образуют последовательности большей длины более приближенные к реальным сайтам связывания ТФ, что будет рассмотрено в следующем разделе.

3.1.5. Поиск регуляторных модулей, ассоциированных с ответом на ауксин

Интересно, что количество элементов-партнеров, входящих в состав парных AuxRE, более чем в два раза превышает количество предсказанных в разделе 3.1.2 одиночных AuxRE. Мы предположили, что часть парных AuxRE может представлять собой фрагменты более протяженных цис-регуляторных элементов, находящихся рядом с последовательностью TGTCNN в регуляторных районах генов. Для проверки этой гипотезы предсказанные парные AuxRE были картированы на геном *Arabidopsis* (см. методы). По результатам выравнивания были обнаружены кластеры парных AuxRE, в которых было локализовано более одного предсказанного парного AuxRE (далее, регуляторные модули). Всего было выявлено 3462 гена, содержащих такие регуляторные модули. Из них 3114 гена содержали по одному модулю,

279 – по два, 41 – по три, 14 – по четыре и более. В последнюю группу попали как известные гены, вовлеченные в регуляцию транскрипционного ответа на ауксин, так и гены, связь которых с ауксином не была ранее продемонстрирована (Таблица 6). Среди известных ауксин-чувствительных генов оказались представители семейства *SAUR* (*Small Auxin-Up RNA*), *GH3* и *IAA* (*indole-3-acetic acid inducible*). *SAUR* – это ауксин-чувствительные гены, регулирующие элонгацию клеток (Spartz et al., 2014), активация которых происходит в раннем ответе на ауксин; гены семейства *GH3* производят конъюгаты ауксина и жасмоната (Staswick et al., 2005); гены семейства *IAA* – ингибиторы активности ТФ ARF (Weijers and Wagner, 2016). Все эти семейства состоят из генов первичного ответа на ауксин, который реализуется в течение одного часа (Abel et al., 1994), видимо, с этим и связано наличие большого количества парных AuxRE в их промоторах, что позволяет осуществлять быструю и тонкую регуляцию изменений в их экспрессии. К генам, связь которых с ауксином ранее не была показана, но в которых нами были найдены множественные регуляторные модули, относятся *AT3G20340*, *BT5* и *AGP22*.

Таблица 6. Гены, содержащие более четырех регуляторных модулей в своих промоторах.

AGI код	Количество о парных элементов	Количество модулей	Количество отдельных парных элементов	Аннотация гена
<i>AT1G29430</i>	27	5	3	<i>SAUR-подобный</i>
<i>AT1G29440</i>	23	5	2	<i>SAUR63</i>
<i>AT1G29500</i>	22	4	4	<i>SAUR-подобный</i>
<i>AT1G29510</i>	29	6	2	<i>SAUR68</i>
<i>AT1G29520</i>	21	4	0	<i>AWPM-19-подобный</i>
<i>AT1G52830</i>	32	4	2	<i>IAA6</i>
<i>AT1G78100</i>	19	4	5	<i>AUF1</i>
<i>AT2G23170</i>	231	10	0	<i>GH3.3</i>
<i>AT3G15540</i>	40	5	1	<i>IAA19</i>
<i>AT3G20340</i>	57	4	1	Неизвестный, вовлечен в оксидативный стресс
<i>AT4G27260</i>	10	4	0	<i>WES/GH3</i>
<i>AT4G37610</i>	47	5	1	<i>BT5</i>
<i>AT5G47370</i>	33	6	2	<i>HAT2</i>
<i>AT5G53250</i>	13	4	3	<i>AGP22</i>

Рассмотрим строение регуляторных модулей на примере представителей семейства генов *GH3.3* и *IAA* (Рисунок 9). Во-первых, в регуляторных модулях этих генов в непосредственной близости располагается 2 и более парных AuxRE, относительно которых со

сдвигом в несколько нуклеотидов находятся другие парные AuxRE. Причем в некоторых модулях последовательность TGTCNN из одного парного AuxRE может формировать пару с элементом-партнером из расположенного рядом другого парного AuxRE (Рисунок 9А). Как видно из рисунка с выравниванием парных AuxRE, входящие в них элементы-партнеры распознаются программой ТОМТОМ (<http://meme-suite.org>; Gupta et al., 2007) далеко не во всех случаях. В то же время распознавание в составе одного выравнивания для отличающихся в несколько нуклеотидов элементов-партнеров может отличаться. Например, в модуле *GH3.3* на интервале в промоторной последовательности [-79;-53] от ССТ перекрывающиеся элементы-партнеры были распознаны как сайты связывания ТФ bZip, bHLH, TSP, WRKY, AP2. Как функционируют эти ТФ на промоторе остается вопросом: конкурентно или синергично, все или не все они функциональны.

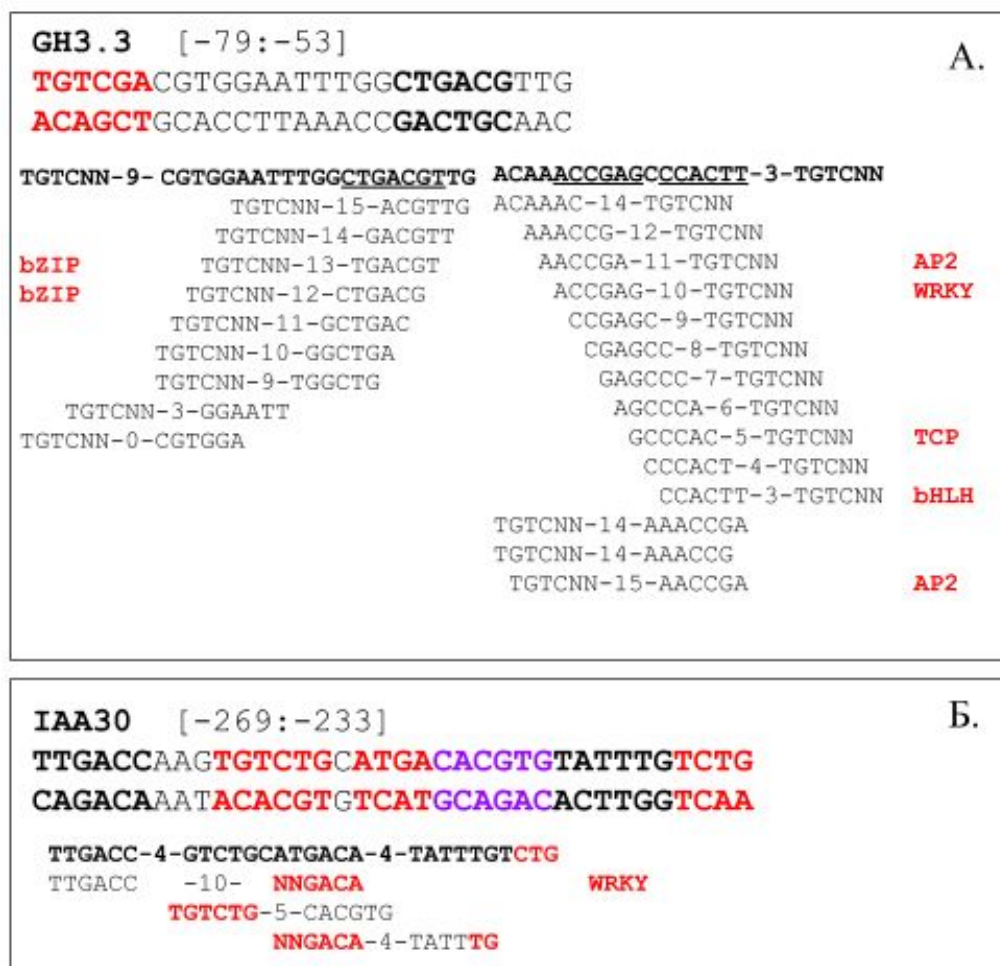


Рисунок 9. Пример регуляторных модулей в промоторах генов *GH3.3* и *IAA30*. А. Схема модуля, расположенного в промоторе гена *GH3.3* на интервале [-79;-53] от ССТ. Б. Схема модуля, расположенного в промоторе гена *IAA30* на интервале [-269;-233] от ССТ.

В ходе анализа регуляторных модулей, во-первых, были выявлены кластеры парных AuxRE, в которых элементы-партнеры последовательности TGTCNN выравниваются относительно друг друга, во-вторых, в рамках одного модуля, как правило, было идентифицировано несколько сайтов связывания ТФ.

3.2. Экспериментальная верификация результатов биоинформатического анализа и биологической функциональности отобранных парных AuxRE

3.2.1. Верификация обнаруженных парных AuxRE методом ОТ-количественной ПЦР

При выборе парных AuxRE для экспериментальной верификации мы руководствовались степенью достоверности их обогащения в промоторах ауксин-чувствительных генов и представленностью в геноме (количество промоторов, в которых они присутствуют). Таким образом, по упомянутым критериям были отобраны 40 парных AuxRE (Таблица 7).

Таблица 7. Парные AuxRE, наиболее значимо обогащенные в промоторах ауксин-чувствительных генов.

Парный элемент*	Количество генов	Значимость (<i>p</i> -value)
GCGGTC-12-TGTCNN	2	<5E-324
CCGCGG-14-TGTCNN	2	<5E-324
GACGTT-0-TGTCNN	25	0,000002
ACATCC-10-TGTCNN	30	0,000025
TCCAAA-6-TGTCNN	105	0,000003
TGTCNN-13-GGACTA	29	0,000075
CAAATA-4-TGTCNN	122	0,000022
GGCAGA-15-TGTCNN	16	<5E-324
AGACGT-1-TGTCNN	20	0,000004
CACATG-4-TGTCNN	91	0,000069

TTCACA-0-TGTCNN	114	0,001469
TGTATA-0-TGTCNN	131	0,001904
GTGAGT-11-TGTCNN	28	0,001489
TGTCNN-13-AATATC	113	0,000427
GGCAGA-15-TGTCNN	16	<5E-324
TGTCNN-6-TACGAA	35	0,002538
TAATAC-15-TGTCNN	75	0,003565
TGTCNN-11-TTTATA	192	0,001648
TGTCNN-2-CCACGT	35	0,000069
TAAAGA-14-TGTCNN	108	0,001205
GGCGTG-0-TGTCNN	11	0,000016
CGGGTA-10-TGTCNN	5	0,000256
TGTCNN-14-CTCGGC	11	0,000146
GACACA-15-TGTCNN	36	0,000112
TGTCNN-1-TTGCCG	8	0,000773
TGTCNN-14-GTTAGG	35	0,00008
GTACGA-10-TGTCNN	9	0,000638
TGTCNN-13-GCTCGG	8	0,000676
TGTCNN-8-TATATC	99	0,000395
TGTCNN-0-CGTAGG	11	0,000617
TGTCNN-14-CGTAAT	29	0,000382
TGTCNN-3-GGCATC	12	0,001557
TGTCNN-4-TTAACA	64	0,000982

ATAAGA-12-TGTCNN	82	0,001559
TGTCNN-10-ATTAAT	190	0,00012
AGCAGC-0-TGTCNN	14	0,000088
TTTAGC-8-TGTCNN	71	0,001288
CATTAC-7-TGTCNN	55	0,001059
TTGTAA-5-TGTCNN	125	0,001756

Примечание: * – парный AuxRE, в котором на расстоянии X п.н. от TGTCNN располагается элемент-партнер.

Для первичной верификации предсказаний методом ОТ-количественной ПЦР была проверена экспрессия 40 генов (Таблица 8), промоторы которых содержат потенциальные парные AuxRE из списка (Таблица 7). Если парный AuxRE был аннотирован как связанный с активацией/подавлением транскрипции, то предполагалось, что и экспрессия генов, которые он регулирует, должна изменяться соответственно. Измерения проводились в одной биологической повторности на трех типах органов трехдневного проростка *A.thaliana* (корень, семядольные листья, весь проросток), обработанного 1 мкМ 2,4-Д (более устойчивый к деградации аналог ауксина) в течение одного часа.

Таблица 8. Изменение экспрессии ауксин-чувствительных генов после обработки 2,4-Д в течение одного часа (Приложение 4).

название гена	количество элементов из таблицы 8, X=TGTCNN	относительное изменение экспрессии		
		корень	семядольные листья	проросток
<i>ICR1</i>	2	1,082	0,744	0,422
<i>HAT2</i>	2	4,444	0,411	0,608
<i>SAUR63</i>	2	4,963	4,075	4,375
<i>IAA30</i>	1 (CAAATA-4-X)	15,143	0,154	4,038
<i>IAA10</i>	1 (TCCAAA-6-X)	2,638	1,713	0,458
<i>TCP20</i>	1 (X-13-GGACTA)	1,087	1,039	0,341
<i>GH3.3</i>	10	78,602	0,702	14,572
<i>HEC1</i>	1 (ACATCC-10-X)	0,636	0,934	1,54
<i>LBD29</i>	3	44,996	1,363	36,3
<i>CRF2</i>	1 (CAAATA-4-X)	1,489	2,366	0,631
<i>ABCG8</i>	1 (X-6-TACGAA)	10,217	1,34	0,403
<i>PBP1</i>	1 (GGCAGA-15-X)	13,615	0,995	3,998
<i>PUCHI</i>	1 (X-13-AATATC)	3,618	1,206	0,231
<i>EBF2</i>	1 (X-11-TTTATA)	1,571	1,23	0,450
<i>HAM2</i>	1 (X-11-TTTATA)	2,908	1,99	0,707
<i>WRKY23</i>	1 (TGTATA-0-X)	0,973	0,91	0,845
<i>bHLH093</i>	1 (X-11-TTTATA)	1,296	1,046	0,645
<i>CLV1</i>	1 (TAATAC-15-X)	1,087	1,766	1,268
<i>CLE44</i>	1 (TGTATA-0-X)	5,425	1,077	2,064

<i>IAA29</i>	1 (GTGAGT-11-X)	23,114	1,037	1,756
<i>AGP22</i>	1 (X-8-TATATC)	0,623	0,887	1,111
<i>HMT3</i>	1 (GACACA-15-X)	0,887	1,601	1,459
<i>WRKY60</i>	1 (X-14-GTTAGG)	1,234	0,806	0,712
<i>ERANT1</i>	1 (GACACA-15-X)	0,784	4,526	0,581
<i>ROPGEF4</i>	1 (GACACA-15-X)	1,043	1,125	0,429
<i>WIP3</i>	1 (X-8-TATATC)	0,79	1,155	0,555
<i>SHA1</i>	1 (CGGGTA-10-X)	0,923	5,197	1,002
<i>APL</i>	1 (X-8-TATATC)	1,329	0,93	0,665
<i>HKT1</i>	2	0,665	1,201	0,617
<i>BT5</i>	2	1,149	2,249	0,485
<i>GASA1</i>	1 (TTTAGC-8-X)	0,74	4,337	0,663
<i>CYP710A1</i>	1 (TTTAGC-8-X)	0,998	1,005	0,223
<i>TBL37</i>	1 (TTTAGC-8-X)	0,756	0,997	0,187
<i>ARPI</i>	1 (ATAAGA-12-X)	0,5473	0,964	0,119
<i>EXPBI</i>	1 (ATAAGA-12-X)	0,119	0,965	0,089
<i>MYB40</i>	1 (X-3-GGCATC)	0,349	1,0284	0,255
<i>FRA8</i>	1 (TTGTAA-5-X)	0,96	0,994	0,116
<i>BRL2</i>	1 (ATAAGA-12-X)	0,77	6,987	0,137
<i>YUC3</i>	1 (TTTAGC-8-X)	0,784	4,839	0,726
<i>PLT6</i>	2	0,768	206,868	0,946

Примечание: измерения осуществлялись в одной биологической повторности. Красным закрашены ячейки со значением, отражающим подавление экспрессии, зеленым – активацию.

Такого рода измерения позволили, во-первых, выбрать гены-кандидаты для дальнейшей работы и, во-вторых, выбрать тип ткани для измерения изменений экспрессии генов на следующем этапе исследования.

По нашим результатам чувствительность к ауксину 80% активируемых и 60% ингибируемых ауксином генов, имеющих в своих промоторах предсказанный парный AuxRE, была подтверждена хотя бы в одной из тканей (Рисунок 10А, Приложение 4). Из каждой группы для дальнейшего анализа были отобраны по 10 генов, транскрипционная активность которых активировалась или подавлялась в ответ на ауксин.

Далее для 20 генов были исследованы изменения экспрессии при различном времени обработки проростков ауксином методом ОТ-количественной ПЦР. Трехдневные проростки *A. thaliana* были обработаны 1 мкМ 2,4-Д в течение 15, 30, 60, 180 и 360 минут. Измерения изначально осуществлялись на корнях и семядольных листьях, так как по результатам ОТ-количественной ПЦР на проростках были получены неоднозначные данные (Таблица 8, Приложение 4). В дальнейшем мы также отказались от измерений на семядольных листьях, так как на первой временной точке наблюдалось резкое повышение экспрессии достаточно большого числа генов. Мы предположили, что это вызвано реакцией на стресс, которым могла быть обработка фитогормоном в жидкой среде. Исходя из этого, далее приводятся результаты, полученные на образцах РНК корней.

В эксперименте с разным временем обработки растений ауксином мы проверили, во-первых, динамику изменений транскрипционной активности генов, промоторы которых содержат предсказанные парные AuxRE, а во-вторых, их связь с ранним или поздним ответом на ауксин.

Мы установили, что по результатам ОТ-количественной ПЦР нельзя проследить четкого разграничения между ранним и поздним ответом. Большинство генов изменяли свою экспрессию как в раннем, так и в позднем ответе, с отличием только в величине изменения экспрессии - в позднем ответе оно было более значительным. Так экспрессия генов *LBD29*, *IAA30* и *HAT2*, промоторы которых содержат парные AuxRE, связанные с ранним ответом на ауксин, повышалась во всех временных точках, в том числе и в точках позднего ответа (180 и 360 минут) (Рисунок 10Б).

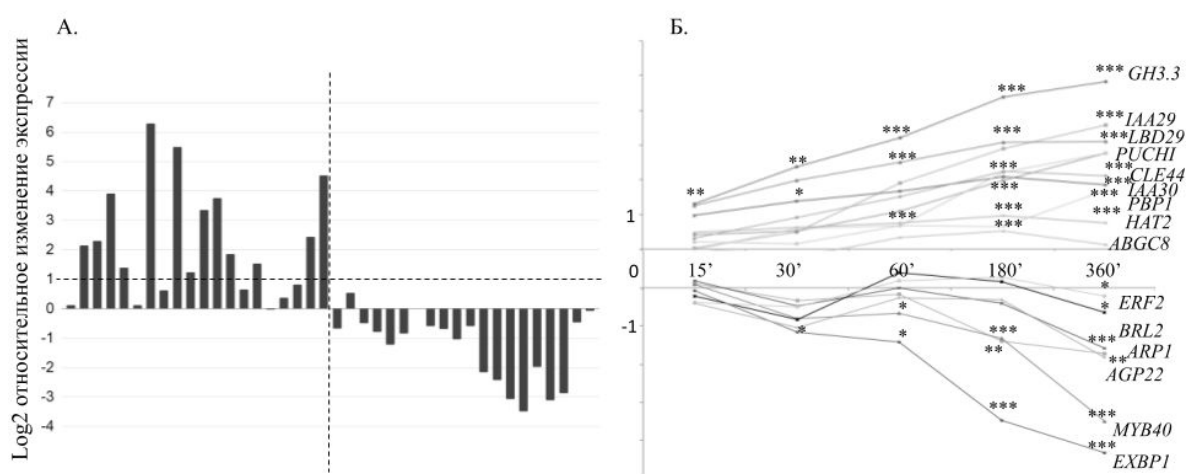


Рисунок 10. Изменение экспрессии генов, содержащих в промоторах предсказанные парные AuxRE, в ответ на ауксин (1мкМ 2,4-Д).

А. Гистограмма изменения экспрессии генов в ответ на ауксин при обработке целых проростков, корней, семядольных листьев экзогенным ауксином в течение 60 минут (Приложение 4). Б. График изменения экспрессии генов в ответ на ауксин в корнях *A.thaliana* на разных временных точках обработки: 15, 30, 60, 180 и 360 мин. Пороги значимости: * – p -value < 0,05; ** – p -value < 0,01; *** – p -value < 0,001.

В случае генов, парные AuxRE которых связаны с поздним ответом на ауксин, также наблюдалось изменение экспрессии и на ранних временных интервалах (Рисунок 10Б). Таким образом, четкой корреляции с ранним и поздним ответом выявлено не было. Например, экспрессия гена *AGP22* подавлялась на поздних временных точках (Рисунок 10Б), хотя в соответствии с данными микрочип экспериментов, которые были использованы в биоинформатическом анализе, он связан с ранним ответом. В случае генов позднего ответа, экспрессия половины из них подавлялась и на ранних временных точках (Рисунок 10Б).

Можно заключить, что разделение на ранний и поздний ответ на ауксин является условным и свидетельствует лишь о чувствительности микрочип метода в детекции слабых изменений экспрессии генов.

3.2.2. Проверка функциональности предсказанных AuxRE методом направленного мутагенеза

По результатам измерения экспрессии генов после разной длительности обработки проростков ауксином были отобраны пять парных AuxRE для верификации их функциональности методом направленного мутагенеза (Таблица 9). Три из отобранных парных AuxRE ассоциированы с активацией экспрессии генов в ответ на ауксин и два – с подавлением. Также при выборе мы руководствовались стабильной экспрессией генов, промоторы которых содержат эти элементы. Элементы-партнеры в составе двух отобранных парных AuxRE (TATTTG-4-TGTCNN, TGTCNN-8-TATATC) были выявлены как обогащенные в промоторах ауксин-чувствительных генов одиночные AuxRE (Cherenkov, Novikova et al., 2018).

Таблица 9. Отобранные для проверки направленным мутагенезом парные AuxRE и их краткая характеристика.

Парный AuxRE	Аннотация элемента-партнера*	Тип транскрипционного ответа
TGTCTG-5-CACGTG	G-бокс	активация
TTGACC-10-TGTCNN	WRKY	активация
CAAATA-4-TGTCNN	неизвестный	активация
TGTCNN-8-TATATC	неизвестный	репрессия
ATAAGA-12-TGTCNN	MYB-подобный	репрессия

Примечание: *парный AuxRE идентифицирован как значимо совпадающий с известным сайтом связывания ТФ программой ТОМТОМ (<http://meme-suite.org>; Gupta et al., 2007).

Проверка функциональности выбранных парных AuxRE осуществлялась на последовательностях промоторов двух генов, в которых был найден один и тот же предсказанный цис-регуляторный элемент (Таблица 10).

Таблица 10. Характеристика промоторов генов, отобранных для верификации методом направленного мутагенеза.

Ген	Парный AuxRE	Позиция парного AuxRE относительно ССТ	Количество TGTCNN в промоторе гена	Длина промотора
<i>IAA30</i>	TTGACC-10-TGTCNN TATTTG-4-TGTCNN, TGTCTG-5-CACGTG	-236	21	1887
<i>GATA23</i>	TATTTG-4-TGTCNN	-701	8	1039
<i>AGP22</i>	TGTCNN-8-TATATC	-949	9	1435
<i>AGP7</i>	TGTCNN-8-TATATC	-797	9	1942
<i>BRL2</i>	ATAAGA-12-TGTCNN	-734	13	3190
<i>ARPI</i>	ATAAGA-12-TGTCNN	-196	19	2898

Для проверки функциональности отобранных парных AuxRE в ответе на ауксин были сделаны и трансформированы в растения *A.thaliana* агробактериальным методом необходимые наборы генетических конструкций (см. раздел 2.7; раздел 2.8). В результате этой работы было создано 30 генетических конструкций. После их трансформации в растения была проведена селекция 250 независимых линий растений, из которых было отобрано около 60-ти линий для дальнейшего анализа.

3.2.3. Паттерны экспрессии белка GFP в репортерных линиях при стандартных условиях

Для шести генов, отобранных для изучения функциональности парных AuxRE, были получены репортерные линии с интактным промотором и набором мутаций последовательности TGTCNN и

элемента-партнера в составе парного AuxRE. В репортерных линиях (Рисунок 11) было получено более двух тысяч трансформантов третьего поколения (Т3), из них были отобраны 2-3 независимые линии со стабильным паттерном экспрессии репортера GFP. По этим линиям ниже приведены результаты конфокальной микроскопии. Анализ проводился на растениях в стандартных условиях выращивания (Рисунок 12 - 14).

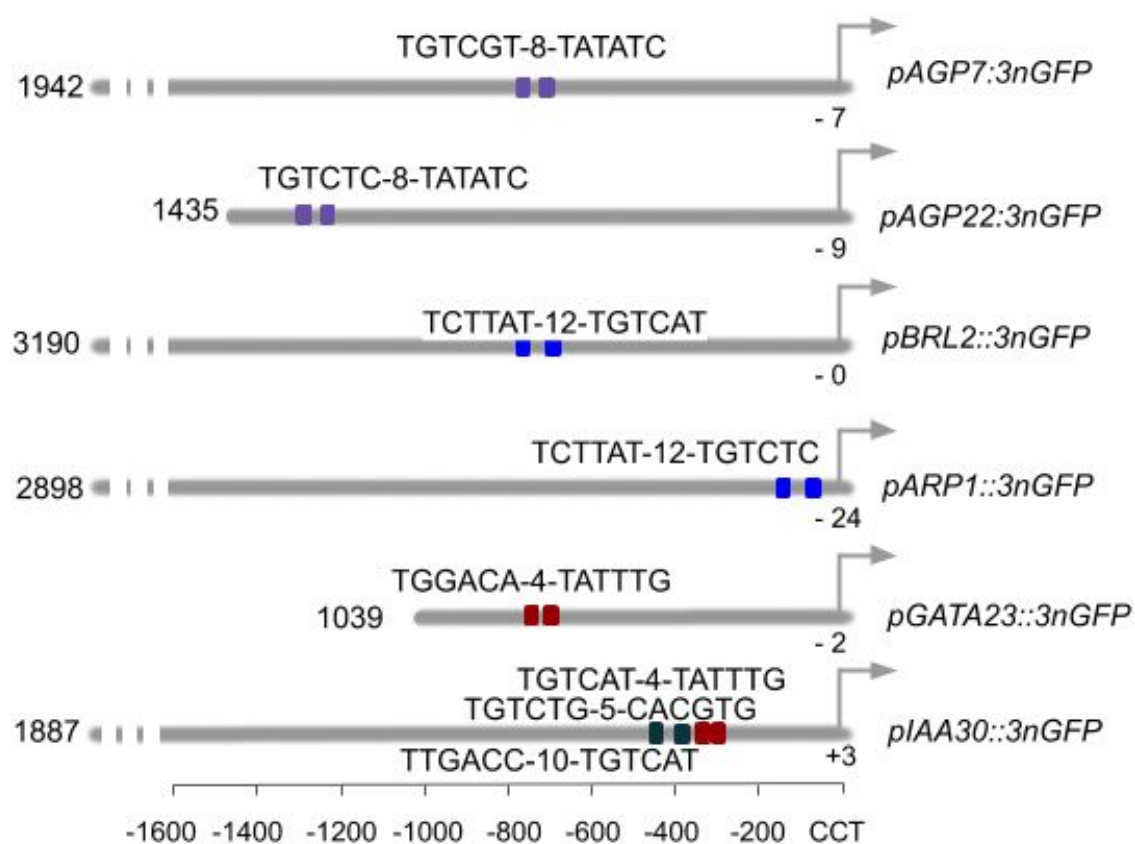


Рисунок 11. Упрощенная схема промоторов, клонированных в генетические векторы для создания репортерных линий. Цветными квадратами обозначены мутированные парные AuxRE в последовательностях промоторов генов.

Белок GFP в конструкции *pIAA30::n3GFP* экспрессируется в кончике корня (Рисунок 12А) в линиях с интактным промотором и промотором с

мутированными элементами (Рисунок 12 Б-Е), причем в случае мутации двух последовательностей TGTCTG1 и CACGTG (Рисунок 12 Б, Г) наблюдается небольшое расширение паттерна и увеличение интенсивности свечения (Рисунок 12 Б-Г) по сравнению с линией, содержащей интактный промотор (Рисунок 12 А).

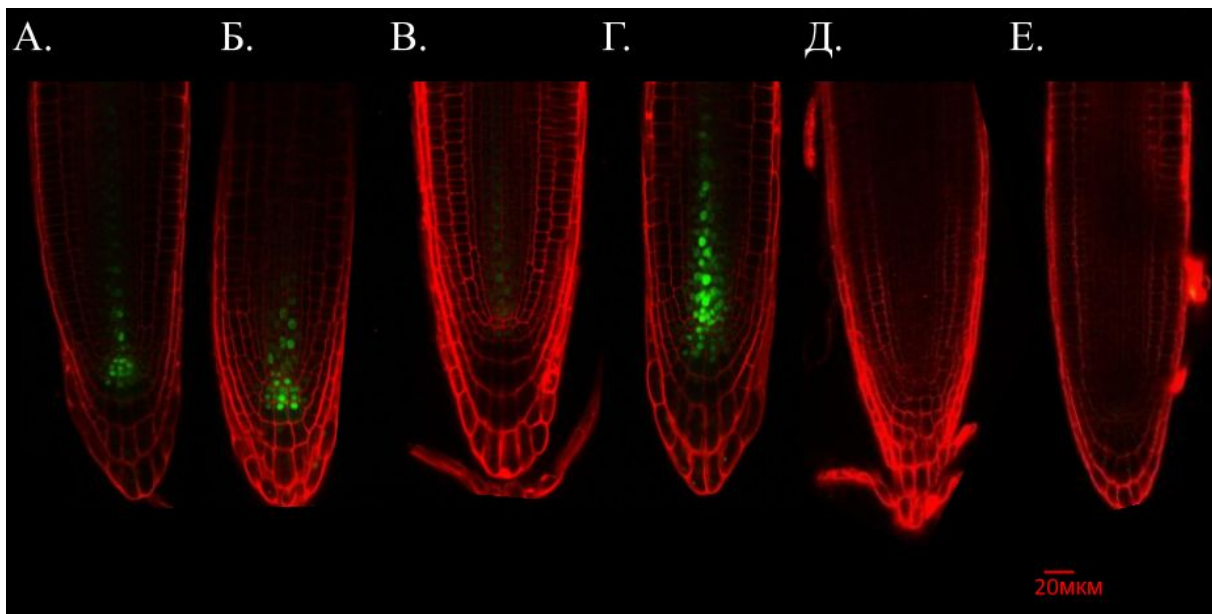


Рисунок 12. Паттерны экспрессии белка GFP под нативным и мутированными промоторами гена *IAA30*. А. Линия с интактным промотором. Б. Линия с мутированной последовательностью CACGTG. В. Линия с мутированной последовательностью ТАТТТГ. Г. Линия с мутированной последовательностью TGTCTG1. Д. Линия с мутированной последовательностью TGTCTG2. Е. Линия с мутированной последовательностью ТТGACC. Клеточные стенки окрашены йодидом пропидия (красный).

В репортерных линиях с промотором гена *GATA23* экспрессия белка GFP наблюдалась вдоль проводящего пучка в зонах возникновения

боковых корней (Рисунок 13). Если ген *GATA23* с промотором дикого типа экспрессируется в зоне возникновения боковых корней, то при мутации в парном AuxRE паттерн экспрессии GFP был “размазан” вдоль проводящего пучка.

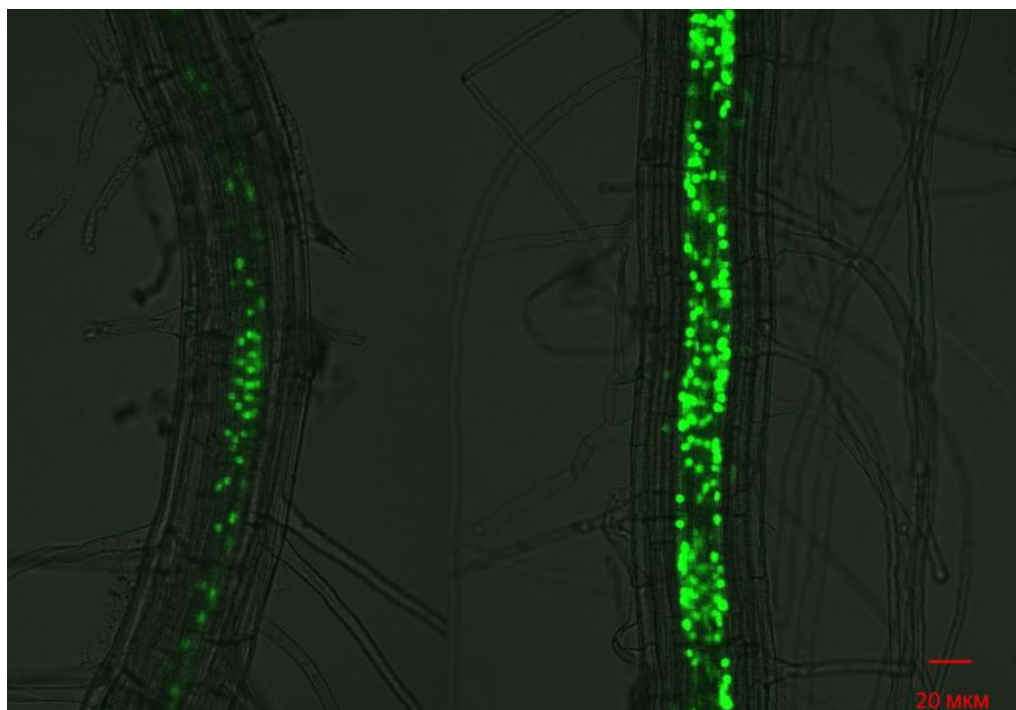


Рисунок 13. Паттерн экспрессии гена *GATA23* с интактным (А) и мутированным промотором (Б) в зоне всасывания корня. Клеточные стенки не окрашены.

В случае репортерной линии, в которой GFP экспрессируется под интактным промотором гена *AGP22*, флуоресценция наблюдалась в эпидермисе и кортексе кончика корня, а также, начиная с зоны роста, в кортексе и проводящем пучке (Рисунок 14А). Паттерн экспрессии GFP под промотором гена *AGP7* не отличался от такового в линии с интактным промотором гена *AGP22* за исключением отсутствия характерного паттерна в кончике корня.

Паттерн экспрессии GFP под промотором гена *BRL2*, как интактного, так и мутированного, сильно не отличался по расположению и интенсивности свечения (Приложение 5). В вытянутых вдоль всего проводящего пучка ядрах наблюдалась умеренная экспрессия GFP (Рисунок 14Б).

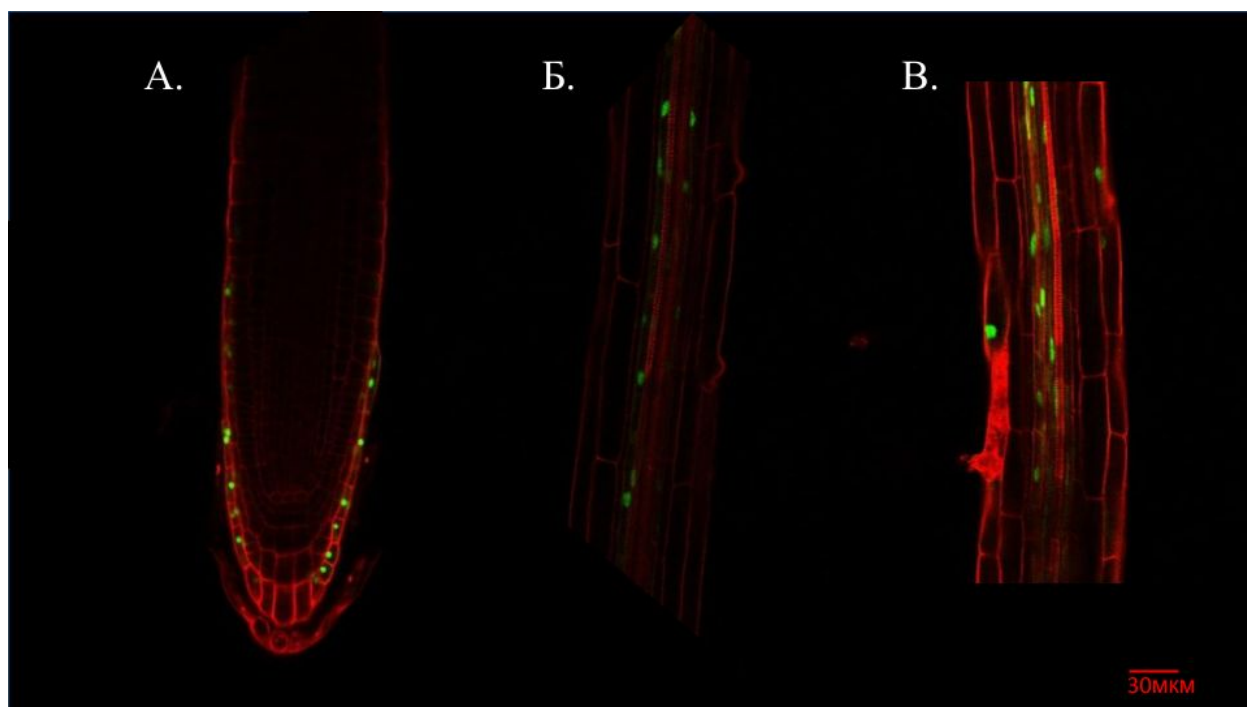


Рисунок 14. Паттерны экспрессии генов *AGP22* (А), *BRL2* (Б), *ARPI* (В) под интактным и мутированным промотором. Клеточные стенки окрашены йодидом пропидия (красный).

При изучении экспрессии GFP в репортерной линии гена *ARPI* с интактным промотором было выявлено два паттерна экспрессии (Рисунок 14В): во всех полученных 42 линиях наблюдалась слабая экспрессия в проводящем пучке, в шести из них экспрессия GFP присутствовала в кончике корня. Разницы в степени экспрессии в корне при стандартных

условиях визуально не наблюдалось при рассмотрении линий с интактным промотором и мутированным (Приложение 5).

Таким образом, были изучены и описаны паттерны экспрессии 30 репортерных линий. Паттерны экспрессии генов *AGP22*, *AGP7*, *ARPI* были изучены *in vivo* впервые. В линиях с мутациями в парных AuxRE в промоторах генов *IAA30* и *GATA23* наблюдалось расширение паттерна экспрессии GFP по сравнению с репортерными линиями, несущими интактные промоторы. В остальных линиях систематических отличий между экспрессией GFP под мутированными и интактными промоторами в нормальных условиях выявлено не было (Приложение 5).

3.2.4. Изменение паттернов экспрессии белка GFP в репортерных линиях *Arabidopsis* при обработке растений ауксином

Следующим этапом были изучены паттерны экспрессии гомозиготных линий третьего поколения в ответ на обработку ауксином. В случае линий, промоторы которых содержат парные AuxRE, ассоциированные с подавлением экспрессии (*pAGP22::n3GFP*, *pAGP7::n3GFP*, *pARPI::n3GFP*, *pBRL2::n3GFP*), значимых отличий интенсивности флуоресценции репортера GFP после обработки ауксином в корне обнаружено не было (Приложение 5). Согласно анализу изображений в программе ImageJ, в линиях с ожидаемой активацией экспрессии GFP (*pIAA30::n3GFP*, *pGATA23::n3GFP*) были обнаружены значимые ($p\text{-value} < 0,05$) отличия между контрольной линией (белок GFP экспрессируется под интактным промотором) и линией с мутированными AuxRE.

В промоторе гена *IAA30* в составе одного регуляторного модуля расположены три парных AuxRE (Рисунок 15), в состав которых вошли

следующие элементы-партнеры: G-бокс, A/T-богатый элемент и потенциальный сайт связывания WRKY. Причем последовательности G-бокс и A/T-богатого элемента пересекаются с последовательностями TGTCNN (Рисунок 15).

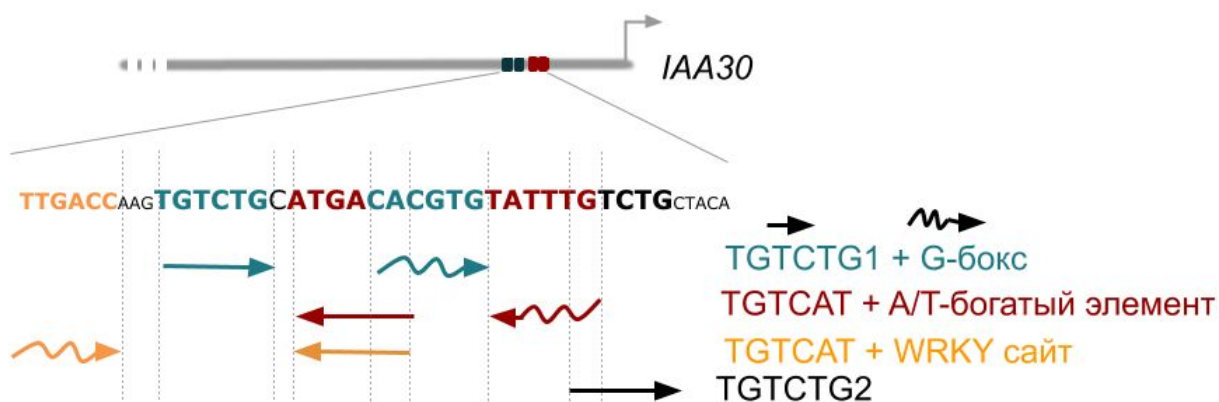


Рисунок 15. Схема предсказанного регуляторного модуля гена *IAA30*. Последовательности TGTCNN обозначены прямой стрелкой, последовательности элементов-партнеров – волнистой. Разными цветами обозначены стрелки, соответствующие разным предсказанным парным AuxRE.

Наблюдалась активация экспрессии репортера GFP под интактным промотором гена *IAA30* (*pIAA30::n3GFP*), как и ожидалось в соответствии с приведенными ранее результатами ОТ-количественной ПЦР по этому гену. В линиях с мутированными последовательностями в составе парных AuxRE в промоторе гена *IAA30* ответ на ауксин отличался от контрольной группы. При мутации последовательностей TGTCTG1 и CACGTG наблюдалось расширение паттерна экспрессии GFP (*pIAA30m1::n3GFP,-TGTCTG1*; *pIAA30m2::n3GFP, -CACGTG*; Рисунок 16) по сравнению с контролем, при мутации TGTCTG2 и TTGACC паттерн экспрессии GFP не идентифицировался как в контрольной

группе, так и в случае обработки ауксином (*pIAA30m4::n3GFP, -TTGACC*; *pIAA30m3::n3GFP, -TGTCTG2*; Рисунок 16).

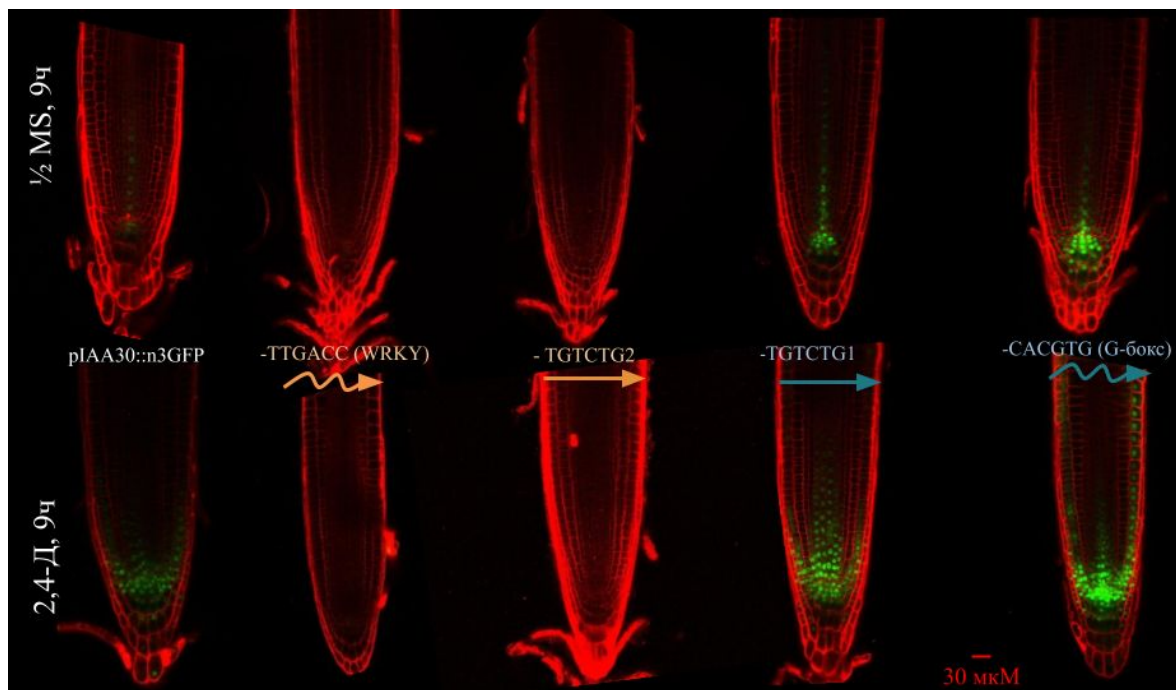


Рисунок 16. Паттерны экспрессии GFP под промотором гена *IAA30* с мутированными последовательностями TGTCTG, CACGTG, TTGACC в контроле ($\frac{1}{2}$ MS) и после обработки аналогом ауксина (2,4-Д) в течение 9 ч. Подписи репортерных линий растений соответствуют рисунку 14. Клеточные стенки окрашены пропидия йодидом (красный).

В случае мутации последовательности TATTTG паттерн экспрессии GFP сохранялся без значимых изменений в ответ на обработку ауксином (Рисунок 17), если же была мутирована вся последовательность TGTCTG2, то паттерн экспрессии GFP исчезал полностью, что говорит о ключевой функциональной роли в меристеме корня именно последовательности TGTCTG2, а не – TATTTG, в которую входят два нуклеотида от канонической последовательности AuxRE (TGTCTG). Отличия интенсивности флуоресценции GFP и изменения паттерна

экспрессии в изучаемых репортерных линиях с промотором гена *IAA30* были подтверждены количественно с помощью анализа изображений при помощи программы ImageJ (Рисунок 17). Для линий растений с мутированными последовательностями TGTCTG2 и TTGACC количественный анализ не проводился, так как экспрессия репортерного белка GFP в этих линиях полностью исчезала.

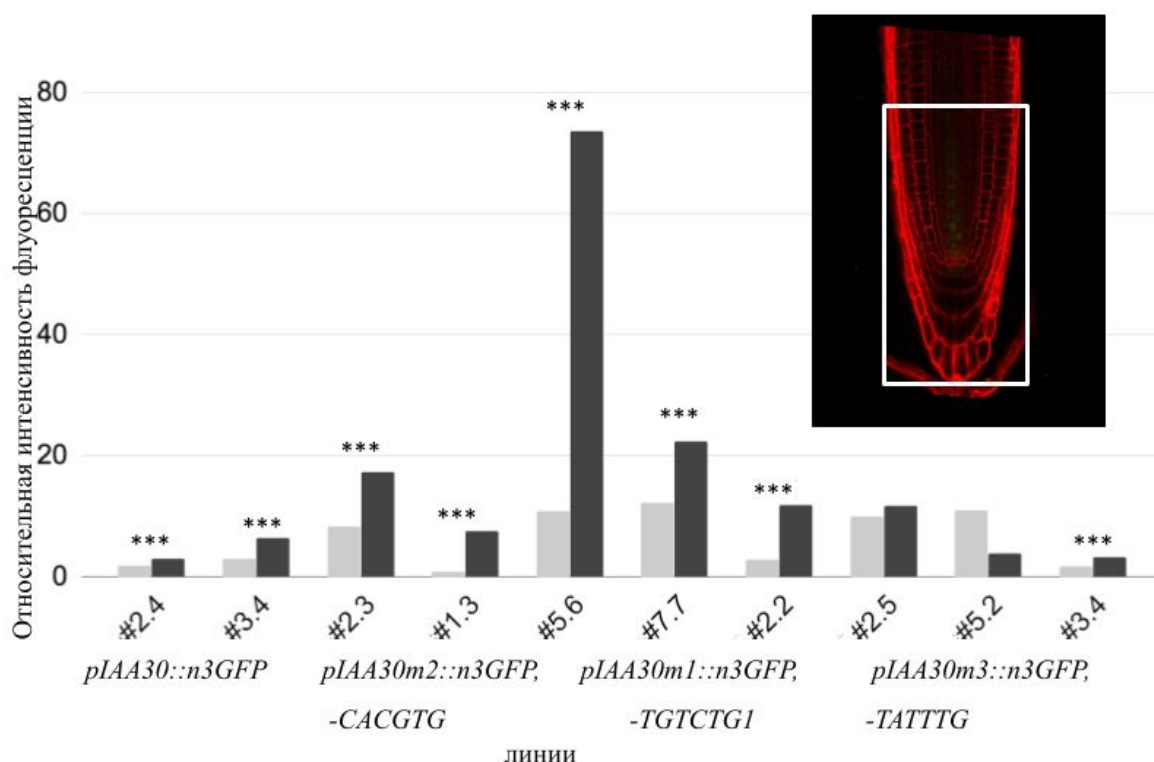


Рисунок 17. Количественный анализ интенсивности флуоресценции GFP в линиях *pIAA30::n3GFP*; *pIAA30m1::n3GFP, -TGTCTG1*; *pIAA30m2::n3GFP, -CACGTG*; *pIAA30m3::n3GFP, -TATTTG*; независимые линии пронумерованы. Зона анализируемого участка корня выделена белой рамкой. Клеточные стенки окрашены пропидия йодидом (красный).

Следует отметить, что при мутации последовательностей TTGACC и TGTCTG2 наблюдалась потеря флуоресценции белка GFP, что является косвенным свидетельством их совместного функционирования. Однако

эти два элемента были выявлены в исследуемом регуляторном модуле в составе разных парных AuxRE (TGTCNN-10-TTGACC, TGTCNN-4-TATTTG), хотя согласно экспериментальным данным функционируют совместно. Это свидетельство комплексной сложности функционирования сайтов связывания TF в составе регуляторных модулей и необходимости изучения не одиночных и парных AuxRE отдельно, а – с учетом контекста.

Функциональность парного элемента TGTCCA-4-TATTTG также была изучена в промоторе гена *GATA23*. В связи со сложностью изучения паттерна экспрессии гена *GATA23* в корне (в норме экспрессия гена *GATA23* наблюдается в зоне возникновения боковых корней, при обработки растений ауксином паттерн расширяется на весь проводящий пучок, чем осложняется выбор зоны корня для анализа), исследование проводили с помощью метода ОТ-количественной ПЦР: после обработки ауксином гомозиготных по мутации растений третьего поколения были собраны корни, выделена РНК и относительная экспрессия гена *GFP* была измерена методом ОТ-количественной ПЦР, по результатам которой было обнаружено снижение экспрессии в линиях с мутированным промотором (Рисунок 18).

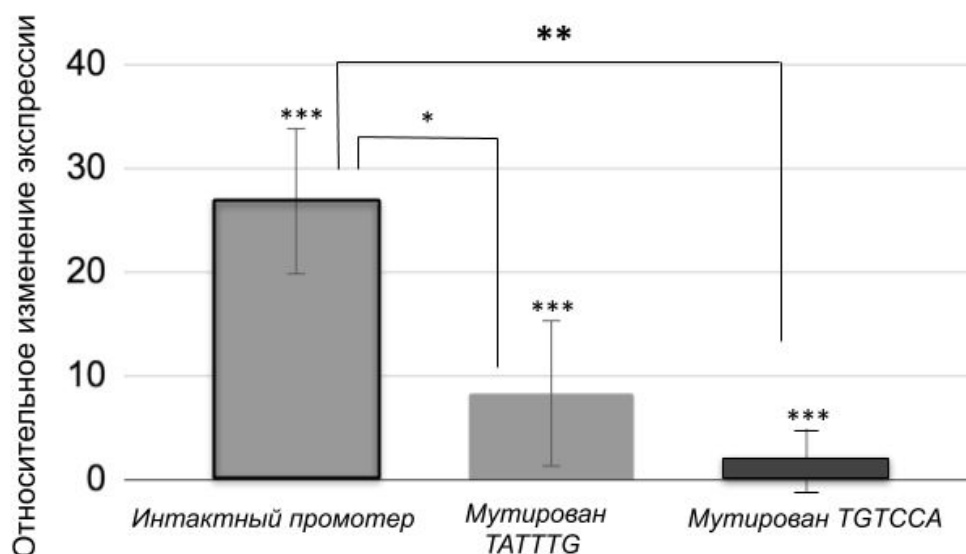


Рисунок 18. Изменение экспрессии гена *GFP* под промотором гена *GATA23* в эксперименте с обработкой 1мкМ 2,4-Д в течение 6 ч. Измерение осуществлялось в одной биологической повторности для каждой независимой линии для каждого типа промотора. Пороги значимости: * – p -value < 0,05; ** – p -value < 0,01; *** – p -value < 0,001.

В экспериментах *in vivo* мы показали, что парные AuxRE TGTCCA-4-TATTTG и TGTCTG-5-CACGTG, TTGACC-24-TGTCTG участвуют в регуляции транскрипции генов *GATA23* и *IAA30* ауксином, соответственно. Причем при мутировании парного AuxRE TGTCTG-5-CACGTG наблюдалась активация экспрессии репортера, что говорит об ингибирующей роли данного парного AuxRE в регуляции активности гена *IAA30*. В случае же мутирования парного AuxRE TTGACC-24-TGTCTG экспрессия репортера полностью подавлялась, значит этот парный AuxRE отвечает за активацию экспрессии гена *IAA30*. Возможно, что наличие элемента-партнера последовательности TGTCNN

и определяет то, что одни те же последовательности TGTCNN могут обеспечивать как подавление, так и активацию транскрипции генов.

3.2.5. Выявление транскрипционных факторов, связывающихся с идентифицированными парными AuxRE в промоторе гена *IAA30*

Для определения ТФ, связывающихся с предсказанным регуляторным модулем в промоторе гена *IAA30* (Рисунок 19А) в геном дрожжей *Saccharomyces cerevisiae* был встроен троекратный повтор изучаемого регуляторного модуля, состоящего из предсказанных парных AuxRE. Дрожжи, несущие повтор, скрещивались с дрожжами, несущими один из 1957 ТФ *Arabidopsis*; при взаимодействии ТФ с повтором регуляторного модуля происходила активация экспрессии гистидина. Селекция взаимодействий проводилась на среде без гистидина и дополнительным подавлением его синтеза с помощью 3-АТ (3-амино-1,2,4-триазол).

После скрининга было выявлено 23 ТФ-ДНК взаимодействия, из них – 14 с ТФ семейства bZip, пять – ТФ LBD, а также ТФ WOX3, MADS, WRKY7, ARF11 и А-ацетилтрансферазой (Рисунок 19Б; Приложение 6).

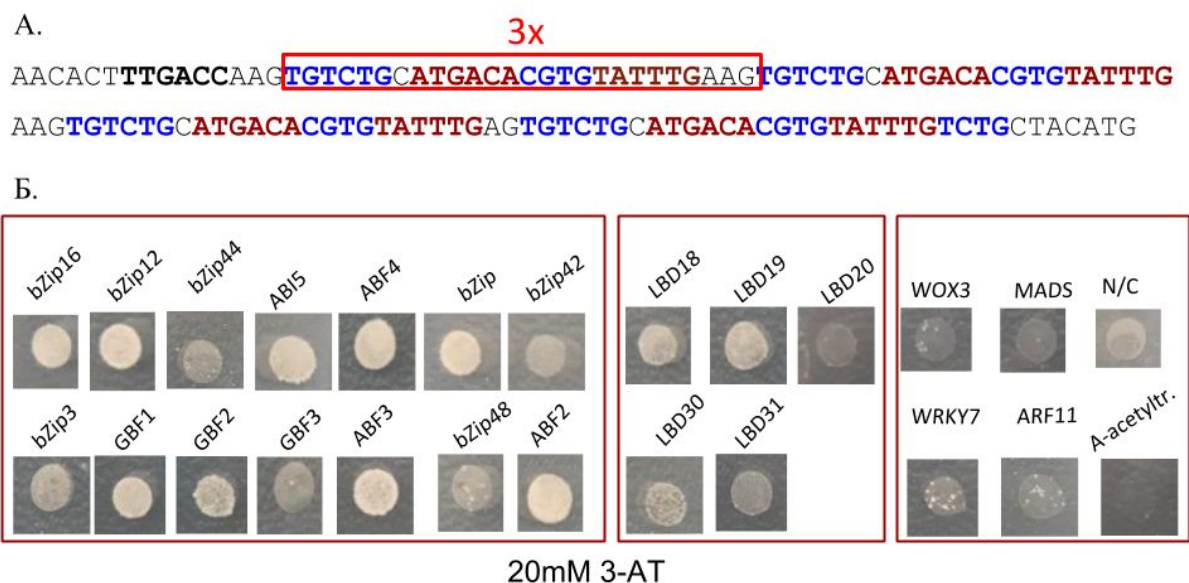


Рисунок 19. Выявление связывающихся с регуляторным модулем промотора гена *IAA30* ТФ в дрожжевой одногибридной системе. А. Схема повтора регуляторного района *IAA30*, взятого в анализ; черным выделен сайт связывания ТФ WRKY, синим – парный AuxRE с сайтом связывания ТФ bZip (G-бокс), бордовым выделен парный AuxRE с А/Т-богатым элементом-партнером. Б. Выявленные взаимодействия предсказанного ауксин-чувствительного района гена *IAA30* на селективной среде (-HIS/URA/TRP, 20mM 3-AT). N/C – отрицательный контроль.

Для описания уже известных взаимодействий 23 ТФ и А-ацетилтрансферазы с 23-мя ТФ ARF (Приложение 7) была использована база данных STRING (<https://string-db.org>). В базе данных STRING на основании опубликованных экспериментальных данных, коэкспрессии, гомологии и анализа текстов была построена сеть взаимодействий (Рисунок 20).

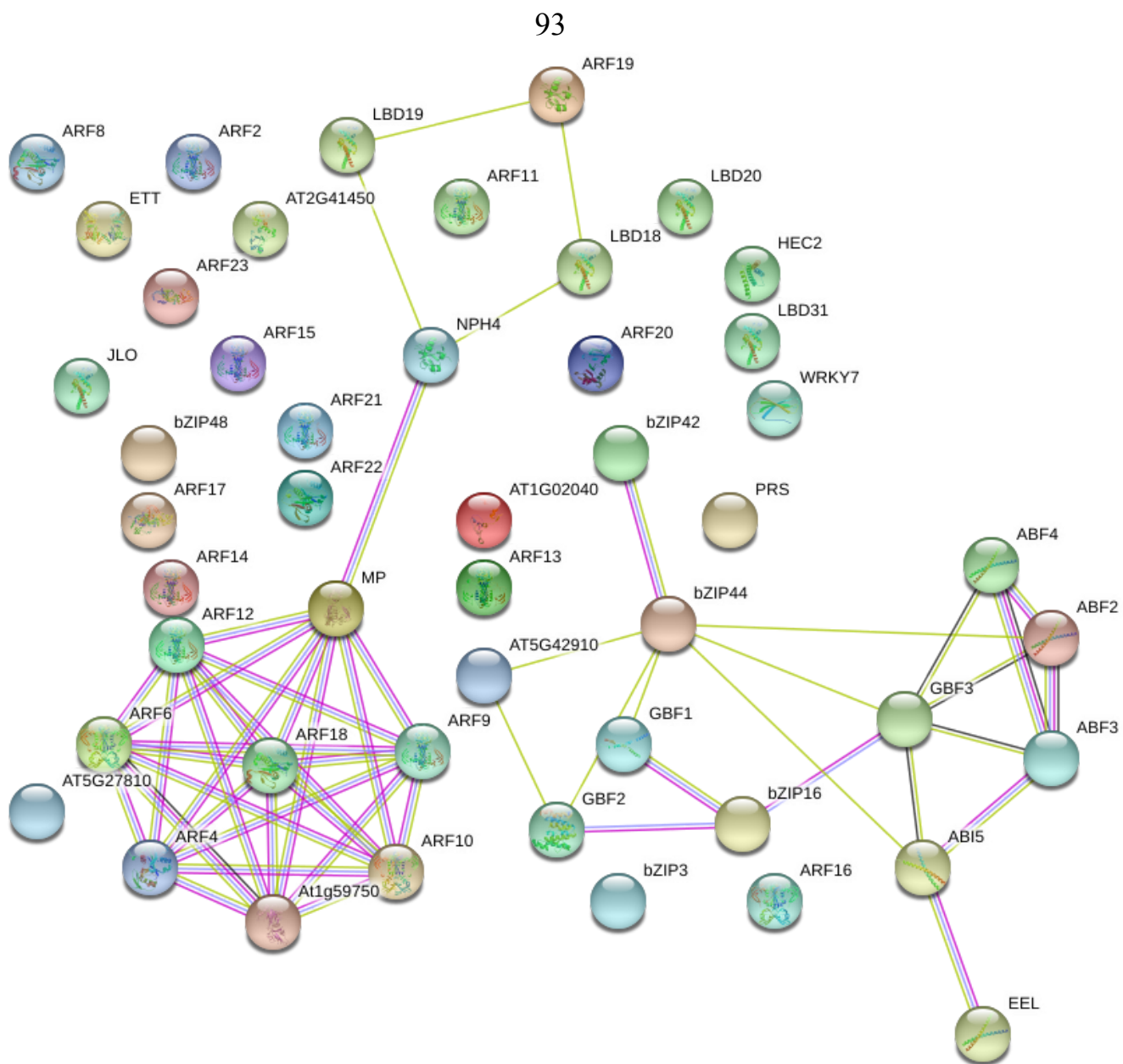


Рисунок 20. Сеть взаимодействий ТФ на основе данных, полученных в дрожжевой одногибридной системе, с семейством белков ARF по данным базы STRING. Розовыми линиями обозначены экспериментально подтвержденные взаимодействия, светло-зеленым – связь на основе анализа текстовых данных, темно-зеленым – коэкспрессия, синим – гомология генов.

Среди выявленных по базе данных STRING взаимодействий были определены подтвержденные ранее физические взаимодействия белков ARF между собой (Legrand et al., 2016; Vernoux et al., 2011; Wang и Estelle, 2014; Piya et al., 2014; Li et al., 2011; Korasick et al., 2015) и ТФ bZip, таких как ABF2, ABF3, ABF4, GBF1, GBF2, GBF3 и др. (Tezuka et al., 2013; Shaikhali et al., 2012; Yoshida et al., 2010; Lynch et al., 2012). Среди гетеродимерных взаимодействий ТФ с белками ARF было показано взаимодействие ТФ LBD18/19 с ТФ ARF7/19 (Okushima et al., 2007; Lee et al., 2009). Для ТФ bZIP44 ранее была продемонстрирована регуляция активности ауксин-чувствительных генов путем связывания с гистон-деацетилазой ADA2b и изменения структуры хроматина, сайты связывания ТФ bZIP44 были расположены вблизи AuxRE в промоторе гена *GH3.3* (Weiste и Droge-Laser, 2014). Возможно, что выявленные в одногибридной дрожжевой системе взаимодействия ТФ семейства bZip с ауксин-чувствительным регуляторным районом, содержащим предсказанный элемент-партнер TGTCNN – CACGTG (см. раздел 3.2.2) в промоторе гена *IAA30*, функционируют аналогичным образом. Для оставшихся ТФ (MADS (AT5G27810), WRKY7, WOX3, NEC2) свидетельств взаимодействия с ТФ ARF в литературе обнаружено не было.

Выявленные в данном исследовании ТФ являются потенциальными партнерами ТФ ARF в регуляции транскрипционной активности генов в ответ на ауксин, причем партнерство может выражаться как в физическом взаимодействии или конкуренции белков, так и опосредованной роли ТФ, не входящих в семейство ТФ ARF, в ответе на ауксин, например, изменении структуры хроматина.

3.2.6. ARF5-зависимое подавление экспрессии гена *WOX4*

Функциональность парных AuxRE, ассоциированных с подавлением экспрессии генов ауксином, верифицировать в рамках данной работы не удалось, так как репортерные конструкции оказались нечувствительными к ауксину (см. раздел 3.2.4). В связи с этим мы решили исследовать цис-элементы, обеспечивающие репрессию транскрипции генов ауксином, прицельно на генах с доказанной ауксин-зависимой репрессией.

В сотрудничестве с исследователями из университета Гейдельберга (Германия), мы провели анализ промотора гена *WOX4*. Репрессия транскрипции гена *WOX4* ауксином была продемонстрирована ранее (Suer et al., 2011). В промоторе гена *WOX4* нами с использованием позиционно-весовых матриц были идентифицированы неканонические AuxRE на расстоянии 335 п.н. (CAGACA) и 2175 п.н. (TGTCATtaCCGACA) от ССТ. Связывание ТФ ARF5 с этими сайтами было проверено методом анализа электрофоретической подвижности белков в группе Thomas Greb. Для одиночного AuxRE было показано связывание с ТФ ARF5 (Brackmann et al., 2018). Таким образом, возможно, ТФ ARF5 подавляет экспрессию гена *WOX4* через последовательность TGTCTG.

Мы провели более подробный анализ ARF5-связывающего участка промотора гена *WOX4* [-435; -235] (Рисунок 21). Анализ состоял из двух этапов: (1) пересечение исследуемого участка с районами связывания 531 ТФ, определенными на основе анализа данных эксперимента Dap-seq (O'Malley et al., 2016) (реализовано совместно с Левитским В.Г., ИЦиГ

СО РАН), (2) распознавание сайтов связывания ТФ в исследуемом районе при помощи программы MEME (<http://meme-suite.org>; Gupta et al., 2007).

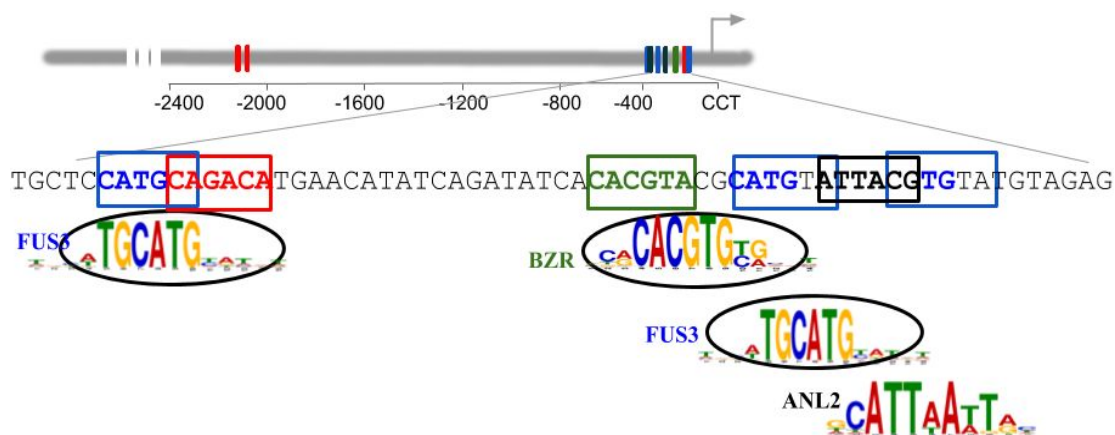


Рисунок 21. Схема промотора гена *WOX4* с выявленными сайтами связывания ТФ; в рамки взяты сайты, выявленные при анализе районов связывания ТФ, полученных на основе анализа данных эксперимента Dap-seq: синим обозначены сайты-связывания ТФ FUS3, зеленым – ТФ BZR, черным – ТФ ANL2, красным – ТФ ARF5.

По результатам первого этапа анализа рядом с ARF5-связывающим AuxRE были выявлены районы связывания ТФ FUS3, ANL2 и BZR. При помощи инструмента MEME позиционно-весовыми матрицами были распознаны сайты связывания тех же ТФ, что и на первом этапе. Таким образом, по результатам двухэтапного анализа рядом с верифицированным сайтом связывания ТФ ARF5 были предсказаны сайты связывания ТФ ANL2, FUS3 и BZR. Возможно, именно, за счет наличия этих сайтов связывания ТФ рядом с последовательностью TGTCTG обеспечивается негативная регуляция транскрипционной активности гена *WOX4* через ТФ ARF5 совместно с ТФ ANL2, FUS3 и BZR. Механизм функционирования описанного регуляторного модуля остается неизвестен.

ЗАКЛЮЧЕНИЕ

Ауксин регулирует большое количество физиологических процессов растений, при этом механизмы, обеспечивающие специфичность в регуляции этих процессов, до сих пор не были выявлены. В рамках данной работы мы предположили, что регуляция ауксином разнообразных процессов осуществляется за счет взаимодействия ТФ ARF с другими ТФ, которые реализуются в регуляции конкретных генов, благодаря наличию в их промоторах парных цис-элементов для обоих ТФ. Для выявления ТФ, участвующих в передаче сигнала от ауксина, мы провели поиск их сайтов связывания в промоторах ауксин-чувствительных генов, объединив биоинформатические и генетические методы для поиска новых типов AuxRE.

В рамках данной работы был разработан метод транскриптомного анализа ассоциаций (ТАА) для предсказания цис-регуляторных элементов в промоторах генов, причем эти цис-регуляторные элементы присутствуют в промоторах разных наборов ДЭГ по данным разных транскриптомных экспериментов. Метод ТАА был взят в основу пакета программ metaRE, разработанного моими коллегами. В данной работе метод был апробирован для предсказания одиночных и парных AuxRE в геноме *A. thaliana*: были предсказаны и аннотированы новые простые и парные AuxRE. Мы провели детальный анализ новых AuxRE путем адаптации пакета программ metaRE к разным задачам, таким как поиск обогащения цис-элементов в районах связывания ТФ, ассоциация цис-элементов с положением относительно ССТ, временем ответа на

обработку растений ауксином (Cherenkov, Novikova et al., 2018). Функциональность шести парных AuxRE была проверена экспериментально.

Для экспериментальной верификации было выбрано несколько неизвестных ранее парных AuxRE. Чувствительность генов с кандидатными AuxRE в регуляторных районах к ауксину была проверена методом ОТ-количественной ПЦР, по результатам которой были отобраны предсказанные цис-элементы для проверки их функциональности *in vivo*. В результате изучения активности репортеров в созданных трансгенных линиях растений с интактными промоторами и промоторами с мутированными парными AuxRE была подтверждена функциональность трех парных элементов. Репортерные линии с парными AuxRE, связанными с подавлением экспрессии ауксином, оказались нечувствительны к обработке фитогормоном. Поэтому мы задались вопросом, могут ли ТФ ARF обеспечивать подавление экспрессии в ответ на ауксин, ведь в нашем анализе одиночные ARF-связывающие AuxRE не были ассоциированы с подавлением транскрипционной активности генов (Таблица 4). В связи с этим был проведен более детальный анализ регуляторного района гена *WOX4*, экспрессия которого значительно изменялась в ответ на ауксин (Suer et al., 2011). В сотрудничестве с группой Thomas Greb мы показали, что подавление экспрессии гена *WOX4* осуществляется путем связывания ТФ ARF5 с одиночным AuxRE в промоторе этого гена.

По результатам верификации биологической функциональности парных AuxRE возник вопрос, с помощью каких ТФ осуществляется регуляция через подтвержденные парные AuxRE и какой механизм лежит в основе. Чтобы ответить на этот вопрос, было проверено связывание

регуляторного модуля в промоторе гена *IAA30*, содержащего три предсказанных парных AuxRE, с библиотекой ТФ *Arabidopsis* в дрожжевой одногибридной системе. В скрининге с 1957 ТФ были выявлены взаимодействия 23-х ТФ и А-ацетилтрансферазы с троекратным повтором исследуемого регуляторного района длиной 120 п.н. (Рисунок 19А). Механизм функционирования этих белков на промоторе гена *IAA30* остается неизвестным. Полученные результаты демонстрируют комбинаторную сложность организации регуляторных районов генов. Видимо, в регуляции транскрипции генов в ответ на ауксин присутствует конкуренция и кооперация ТФ, причем в пределах небольшого участка одного промотора.

Актуальной задачей генетики является дальнейшее изучение структурной сложности цис-регуляторных районов генов и выявление механизмов их функционирования. Данная работа продемонстрировала перспективность применения биоинформатических подходов с последующей экспериментальной верификацией для решения подобных задач.

ВЫВОДЫ

1. Разработан метод Транскриптомного Анализа Ассоциаций (ТАА) для поиска цис-регуляторных элементов, систематически обогащенных в промоторах разных наборов дифференциально экспрессирующихся генов.
2. С помощью метода ТАА предсказаны 147 одиночных и 592 парных ауксин-чувствительных элемента (AuxRE) в промоторах генов *Arabidopsis thaliana*. 27% предсказанных одиночных AuxRE являются известными сайтами связывания транскрипционных факторов, среди неизвестных элементов 66% представлены А/Т-богатыми последовательностями. Среди элементов-партнеров в составе парных AuxRE 7% распознаны как известные сайты связывания транскрипционных факторов.
3. Выявлена композиционная сложность регуляторных районов ауксин-чувствительных генов *Arabidopsis thaliana*, когда в непосредственной близости друг от друга расположены два и более парных AuxRE.
4. В промоторе гена *IAA30* найден и верифицирован *in vivo* AuxRE сложной структуры длиной 35 п.н. С этим AuxRE *in vitro* связываются 23 транскрипционных фактора из семейств bZip, MADS, WOX, LBD, WRKY, bHLH и одна А-ацетилтрансфераза.
5. Активность гена *WOX4 Arabidopsis thaliana* негативно регулируется транскрипционным фактором ARF5 через последовательность TGTCTG в его промоторной области.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

По материалам диссертации опубликовано 9 научных работ, из них – три статьи в журналах из Перечня ВАК (все индексируются в РИНЦ, Scopus и Web of Science), в том числе три – в зарубежных журналах; а также шесть тезисов конференций.

Статьи в журналах

1. Pavel Cherenkov#, **Daria Novikova**#, Nadya Omelyanchuk, Victor Levitsky, Ivo Grosse, Dolf Weijers and Victoria Mironova. Diversity of cis-regulatory elements associated with auxin response in *Arabidopsis thaliana* //Journal of experimental botany. – 2017. – Т. 69. – №. 2. – С. 329-339. doi: 10.1093/jxb/erx254, # – equal contribution
2. Omelyanchuk, N. A., Wiebe, D. S., **Novikova D. D.**, Levitsky, V. G., Klimova, N., Gorelova, V., C. Weinholdt, G. V. Vasiliev, E. V. Zemlyanskaya, N. A. Kolchanov, A. V. Kochetov, I. Grosse, V. V. Mironova. 2017. Auxin regulates functional gene groups in a fold-change-specific manner in *Arabidopsis thaliana* roots. /Scientific reports. – 2017. – Т. 7. – №. 1. – С. 2489. doi:10.1038/s41598-017-02476-8
3. Klaus Brackmann, Jiyan Qi, Michael Gebert, Virginie Jouannet, Theresa Schlamp, Karin Grünwald, Eva-Sophie Wallner, **Daria D. Novikova**, Victor G. Levitsky, Javier Agustí, Pablo Sanchez, Jan U. Lohmann and Thomas Greb. Spatial specificity of auxin responses coordinates wood formation //Nature communications. – 2018. – Т. 9. – №. 1. – С. 875. doi: 10.1038/s41467-018-03256-2

Тезисы конференций

4. **Novikova D.D.**, Mironova V.V., Omelyanchuk N.A., Auxin-responsive transcriptome of *Arabidopsis thaliana* roots, Bioinformatics of Genome Regulation and Structure\Systems Biology, 23-28 of June 2014, Novosibirsk
5. **Novikova D.D.**, Mironova V.V., Omelyanchuk N.A., Cherenkov P.A. Meta-analysis of transcriptome data to investigate auxin response mechanisms in *Arabidopsis thaliana* L., International Conference on Systems Biology, 16 -20 of September 2016, Barcelona, Spain
6. **Novikova D.D.**, Mironova V.V., Cherenkov P.A., Functional annotation of auxin responsive cis-regulatory elements, II Всероссийская конференция с международным участием "Высокопроизводительное секвенирование в геномике", 18-23 июня 2017, Новосибирск, с. 67
7. **Novikova D.D.**, Weijers D., Mironova V.V., Diversity of auxin responsive cis-regulatory elements, 1–5 of July 2018, Auxins and Cytokinins in Plant Development, Prague, Czech Republic, p. 78
8. **Novikova D.D.**, Cherenkov P.A., Tkachev K.U., Levitsky V.G., Mironova V.V., MetaRE: search for cis-regulatory elements via meta-analysis of transcriptomic data, 20-25 of August 2018, Bioinformatics of Genome Regulation and Structure\Systems Biology, Novosibirsk, p.179
9. **Novikova D.D.**, Weijers D., Mironova V.V., Diversity of AuxREs in *Arabidopsis thaliana* L, International Conference on Plant Growth Substances, 25 - 29 of June 2019, Paris, France, p.36

СПИСОК ЛИТЕРАТУРЫ

1. Дерфлинг К. Гормоны растений: Системный подход. – М.: Мир, 1985. – 304 С.
2. Ежова Т. А. *Arabidopsis thaliana*-модельный объект генетики растений. – М.:МАКС Пресс, 2003. – 219 С.
3. Ковганко Н. В. Брассиностероиды в растительном мире //Химия природных соединений. – 1991. – Т. 2. – С. 159-173.
4. Кравец В. С. и др. Регуляторы роста растений: внутриклеточная гормональная сигнализация и применение в аграрном производстве. Второй международный симпозиум (Киев, Украина, 8-12 октября 2007 г.) //Физиология растений. – 2008. – Т. 55. – №. 4. – С. 629-640.
5. Кулаева О. Н. Этилен в жизни растений //Соросовский образовательный журнал. – 1998. – Т. 11. – С. 78-84.
6. Кулаева О. Н. Как регулируется жизнь растений //Соросовский образовательный журнал. – 1995. – №. 1. – С. 20-27.
7. Медведев С. С. Физиология растений: учебник. — СПб.: БХВ-Петербург, 2012. – 512 С.
8. Ракитина Т. Я. и др. Абсцизовая кислота и этилен в мутантах *Arabidopsis thaliana*, различающихся по устойчивости к ультрафиолетовой (УФ-Б) радиации //Физиол. растений. – 1994. – Т. 41.
9. Тузова Р., Ковалев Н. Молекулярно-генетические механизмы эволюции органического мира. Генетическая и клеточная инженерия.- М.: Litres, 2017. – 398С.

10. Abe H. et al. Role of Arabidopsis MYC and MYB homologs in drought-and abscisic acid-regulated gene expression //The Plant Cell. – 1997. – T. 9. – №. 10. – C. 1859-1868.
11. Abel S., Oeller P. W., Theologis A. Early auxin-induced genes encode short-lived nuclear proteins //Proceedings of the National Academy of Sciences. – 1994. – T. 91. – №. 1. – C. 326-330.
12. Abel S., Theologis A. Early genes and auxin action //Plant Physiology. – 1996. – T. 111. – №. 1. – C. 9.
13. Bailey T. L. et al. MEME SUITE: tools for motif discovery and searching //Nucleic acids research. – 2009. – T. 37. – №. suppl_2. – C. W202-W208.
14. Bardet A. F. et al. A computational pipeline for comparative ChIP-seq analyses //Nature protocols. – 2012. – T. 7. – №. 1. – C. 45.
15. Bargmann B. O. R. et al. A map of cell type-specific auxin responses //Molecular systems biology. – 2013. – T. 9. – №. 1. – C. 688.
16. Bargmann B. O. R. et al. TARGET: a transient transformation system for genome-wide transcription factor target discovery //Molecular plant. – 2013. – T. 6. – №. 3. – C. 978.
17. Barrett T. et al. NCBI GEO: mining millions of expression profiles—database and tools //Nucleic acids research. – 2005. – T. 33. – №. suppl_1. – C. D562-D566.
18. Benjamini Y., Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing //Journal of the royal statistical society. Series B (Methodological). – 1995. – C. 289-300.
19. Benjamini Y. et al. The control of the false discovery rate in multiple testing under dependency //The annals of statistics. – 2001. – T. 29. – №. 4. – C. 1165-1188.

20. Berckmans B. et al. Light-dependent regulation of DEL1 is determined by the antagonistic action of E2Fb and E2Fc //Plant physiology. – 2011. – T. 157. – №. 3. – C. 1440-1451.
21. Berendzen K. W. et al. Bioinformatic cis-element analyses performed in Arabidopsis and rice disclose bZIP-and MYB-related binding sites as potential AuxRE-coupling elements in auxin-mediated transcription //BMC plant biology. – 2012. – T. 12. – №. 1. – C. 125.
22. Berg O. G., von Hippel P. H. Selection of DNA binding sites by regulatory proteins: Statistical-mechanical theory and application to operators and promoters //Journal of molecular biology. – 1987. – T. 193. – №. 4. – C. 723-743.
23. Bevan M. W., Flavell R. B., Chilton M. D. A chimaeric antibiotic resistance gene as a selectable marker for plant cell transformation //Nature. – 1983. – T. 304. – №. 5922. – C. 184.
24. Blenau W., Rademacher E., Baumann A. Plant essential oils and formamidines as insecticides/acaricides: what are the molecular targets? //Apidologie. – 2012. – T. 43. – №. 3. – C. 334-347.
25. Boer D. R. et al. Structural basis for DNA binding specificity by the auxin-dependent ARF transcription factors //Cell. – 2014. – T. 156. – №. 3. – C. 577-589.
26. Brackmann K. et al. Spatial specificity of auxin responses coordinates wood formation //Nature communications. – 2018. – T. 9. – №. 1. – C. 875.
27. Bass J. I. F., Reece-Hoyes J. S., Walhout A. J. M. Generating bait strains for yeast one-hybrid assays //Cold Spring Harbor Protocols. – 2016. – T. 2016. – №. 12. – C. pdb. prot088948.
28. Chapman E. J., Estelle M. Mechanism of auxin-regulated gene

expression in plants //Annual review of genetics. – 2009. – T. 43. – C. 265-285.

29. Cherenkov P. et al. Diversity of cis-regulatory elements associated with auxin response in *Arabidopsis thaliana* //Journal of experimental botany. – 2017. – T. 69. – №. 2. – C. 329-339.

30. Chilton M. D. et al. Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis //Cell. – 1977. – T. 11. – №. 2. – C. 263-271.

31. Choi J. et al. Cytokinins and plant immunity: old foes or new friends? //Trends in plant science. – 2011. – T. 16. – №. 7. – C. 388-394.

32. Chomczynski P., Mackey K. Short technical reports. Modification of the TRI reagent procedure for isolation of RNA from polysaccharide- and proteoglycan-rich sources //Biotechniques. – 1995. – T. 19. – №. 6. – C. 942-945.

33. Chuang H., Zhang W., Gray W. M. *Arabidopsis* ETA2, an apparent ortholog of the human cullin-interacting protein CAND1, is required for auxin responses mediated by the SCFTIR1 ubiquitin ligase //The Plant Cell. – 2004. – T. 16. – №. 7. – C. 1883-1897.

34. Darwin, C. *The Power of Movement in Plants*. John Murray, London. – 1880.

35. Davis S., Meltzer P. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. // *Bioinformatics* - 2007 - 14 - P. 1846–1847.

36. De Block, M., D. Debrouwer, and T. Moens. "The development of a nuclear male sterility system in wheat. Expression of the barnase gene under the control of tapetum specific promoters." *Theoretical and Applied Genetics* 95.1-2 (1997): 125-131.

37. Devlin P. F., Yanovsky M. J., Kay S. A. A genomic analysis of the

shade avoidance response in *Arabidopsis* //Plant Physiology. – 2003. – T. 133. – №. 4. – C. 1617-1629.

38. Dharmasiri N. et al. Plant development is regulated by a family of auxin receptor F box proteins //Developmental cell. – 2005. – T. 9. – №. 1. – C. 109-119.

39. Dombrecht B. et al. MYC2 differentially modulates diverse jasmonate-dependent functions in *Arabidopsis* //The Plant Cell. – 2007. – T. 19. – №. 7. – C. 2225-2245.

40. Dreher K. A. et al. The *Arabidopsis* Aux/IAA protein family has diversified in degradation and auxin responsiveness //The Plant Cell. – 2006. – T. 18. – №. 3. – C. 699-714.

41. Dreos R., Ambrosini G., Bucher P. Influence of rotational nucleosome positioning on transcription start site selection in animal promoters //PLoS computational biology. – 2016. – T. 12. – №. 10. – C. e1005144.

42. Durinck S. et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis //Bioinformatics. – 2005. – T. 21. – №. 16. – C. 3439-3440.

43. Edgar R., Domrachev M., Lash A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository //Nucleic acids research. – 2002. – T. 30. – №. 1. – C. 207-210.

44. Eklund D. M. et al. The *Arabidopsis thaliana* STYLISH1 protein acts as a transcriptional activator regulating auxin biosynthesis //The Plant Cell. – 2010. – T. 22. – №. 2. – C. 349-363.

45. Fernández-Calvo P. et al. The *Arabidopsis* bHLH transcription factors MYC3 and MYC4 are targets of JAZ repressors and act additively with MYC2 in the activation of jasmonate responses //The Plant Cell. – 2011. – T. 23. – №. 2. – C. 701-715.

46. Flores-Sandoval E., Eklund D. M., Bowman J. L. A simple auxin transcriptional response system regulates multiple morphogenetic processes in the liverwort *Marchantia polymorpha* // *PLoS genetics*. – 2015. – T. 11. – №. 5. – C. E1005207.
47. Foster R., Izawa T., Chua N. H. Plant bZIP proteins gather at ACGT elements // *The FASEB Journal*. – 1994. – T. 8. – №. 2. – C. 192-200.
48. Franco-Zorrilla J. M. et al. DNA-binding specificities of plant transcription factors and their potential to define target genes // *Proceedings of the National Academy of Sciences*. – 2014. – T. 111. – №. 6. – C. 2367-2372.
49. Gaudinier, A., Zhang, L., Reece-Hoyes, J. S., Taylor-Teeples, M., Pu, L., Liu, Z., Brady, S. M. Enhanced Y1H assays for *Arabidopsis* // *Nature methods*. – 2011. – T. 8. – №. 12. – C. 1053.
50. Guilfoyle T. J. The PB1 domain in auxin response factor and Aux/IAA proteins: a versatile protein interaction module in the auxin response // *The Plant Cell*. – 2015. – T. 27. – №. 1. – C. 33-43.
51. Goda H, Sawa S, Asami T, Fujioka S, Shimada Y, Yoshida S. Comprehensive comparison of auxin-regulated and brassinosteroid-regulated genes in *Arabidopsis* // *Plant physiology*. – 2004. – T. 134. – №. 4. – C. 1555-1573.
52. Gray WM, Ostin A, Sandberg G, Romano CP, Estelle M. High temperature promotes auxin-mediated hypocotyl elongation in *Arabidopsis* // *Proceedings of the National Academy of Sciences*. – 1998. – T. 95. – №. 12. – C. 7197-7202.
53. Gray W. M. et al. *Arabidopsis* SGT1b is required for SCFTIR1-mediated auxin response // *The Plant Cell*. – 2003. – T. 15. – №. 6. – C. 1310-1319.
54. Goossens J. et al. Jasmonates: signal transduction components and

their roles in environmental stress responses //Plant molecular biology. – 2016. – T. 91. – №. 6. – C. 673-689.

55. Grunewald W. et al. A role for AtWRKY23 in feeding site establishment of plant-parasitic nematodes //Plant physiology. – 2008. – T. 148. – №. 1. – C. 358-368.

56. Guern J. Regulation from within: the hormone dilemma //Annals of botany. – 1987. – C. 75-102.

57. Gupta S. et al. Quantifying similarity between motifs //Genome biology. – 2007. – T. 8. – №. 2. – C. R24.

58. Hagen G., Guilfoyle T. Auxin-responsive gene expression: genes, promoters and regulatory factors //Plant molecular biology. – 2002. – T. 49. – №. 3-4. – C. 373-385.

59. He H. et al. Network-based meta-analyses of associations of multiple gene expression profiles with bone mineral density variations in women //PloS one. – 2016. – T. 11. – №. 1. – C. e0147475.

60. Hellens, R. P., Edwards, E. A., Leyland, N. R., Bean, S., & Mullineaux, P. M. pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation //Plant molecular biology. – 2000. – T. 42. – №. 6. – C. 819-832.

61. Hellmann H. et al. Arabidopsis AXR6 encodes CUL1 implicating SCF E3 ligases in auxin regulation of embryogenesis //The EMBO journal. – 2003. – T. 22. – №. 13. – C. 3314-3325.

62. Hehl R., Steffens N. O., Wingender E. Isolation and analysis of gene regulatory sequences //Handbook of Plant Biotechnology. – 2004.

63. van Helden J., André B., Collado-Vides J. Extracting regulatory sites from the upstream region of yeast genes by computational analysis of oligonucleotide frequencies //Journal of molecular biology. – 1998. – T. 281. –

№. 5. – C. 827-842.

64. Helden J., Rios A. F., Collado-Vides J. Discovering regulatory elements in non-coding sequences by analysis of spaced dyads //Nucleic acids research. – 2000. – T. 28. – №. 8. – C. 1808-1818.

65. Ho S. N. et al. Site-directed mutagenesis by overlap extension using the polymerase chain reaction //Gene. – 1989. – T. 77. – №. 1. – C. 51-59.

66. Izawa T., Foster R., Chua N. H. Plant bZIP protein DNA binding specificity //Journal of molecular biology. – 1993. – T. 230. – №. 4. – C. 1131-1144.

67. Jakoby M. et al. bZIP transcription factors in Arabidopsis //Trends in plant science. – 2002. – T. 7. – №. 3. – C. 106-111.

68. James P., Halladay J., Craig E. A. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast //Genetics. – 1996. – T. 144. – №. 4. – C. 1425-1436.

69. Jung C. et al. Overexpression of AtMYB44 enhances stomatal closure to confer abiotic stress tolerance in transgenic Arabidopsis //Plant physiology. – 2008. – T. 146. – №. 2. – C. 623-635.

70. Kato H. et al. Auxin-mediated transcriptional system with a minimal set of components is critical for morphogenesis through the life cycle in *Marchantia polymorpha* //PLoS genetics. – 2015. – T. 11. – №. 5. – C. E1005084.

71. Kim M. et al. Examination of the auxin hypothesis of phytoalexin action in classical auxin assay systems in maize //Journal of plant physiology. – 2016. – T. 190. – C. 67-71.

72. Kim S. Y., Kim Y. S. Genome-wide prediction of transcriptional regulatory elements of human promoters using gene expression and promoter analysis data //BMC bioinformatics. – 2006. – T. 7. – №. 1. – C. 330.

73. Kawai S., Hashimoto W., Murata K. Transformation of *Saccharomyces cerevisiae* and other fungi: methods and possible underlying mechanism // *Bioengineered bugs*. – 2010. – T. 1. – №. 6. – C. 395-403.
74. Kepinski S., Leyser O. The Arabidopsis F-box protein TIR1 is an auxin receptor // *Nature*. – 2005. – T. 435. – №. 7041. – C. 446.
75. Klee H., Estelle M. Molecular genetic approaches to plant hormone biology // *Annual review of plant biology*. – 1991. – T. 42. – №. 1. – C. 529-551.
76. Ko S., Kamada H. Isolation of carrot basic leucine zipper transcription factor using yeast one-hybrid screening // *Plant molecular biology reporter*. – 2002. – T. 20. – №. 3. – C. 301-301.
77. Korasick D. A., Jez J. M., Strader L. C. Refining the nuclear auxin response pathway through structural biology // *Current opinion in plant biology*. – 2015. – T. 27. – C. 22-28.
78. Kosugi S., Ohashi Y. Cloning and DNA-binding properties of a tobacco Ethylene-Insensitive3 (EIN3) homolog // *Nucleic Acids Research*. – 2000. – T. 28. – №. 4. – C. 960-967.
79. Lavy M. et al. Constitutive auxin response in *Physcomitrella* reveals complex interactions between Aux/IAA and ARF proteins // *Elife*. – 2016. – T. 5. – C. E13325.
80. Lee B., Henderson D. A., Zhu J. K. The Arabidopsis cold-responsive transcriptome and its regulation by ICE1 // *The Plant Cell*. – 2005. – T. 17. – №. 11. – C. 3155-3175.
81. Lee H. W. et al. LBD18/ASL20 regulates lateral root formation in combination with LBD16/ASL18 downstream of ARF7 and ARF19 in Arabidopsis // *Plant physiology*. – 2009. – T. 151. – №. 3. – C. 1377-1389.
82. Legrand J. et al. Modelling the influence of dimerisation sequence

dissimilarities on the auxin signalling network //BMC systems biology. – 2016. – T. 10. – №. 1. – C. 22.

83. Lenhard J. Models and statistical inference: The controversy between Fisher and Neyman–Pearson //The British journal for the philosophy of science. – 2006. – T. 57. – №. 1. – C. 69-91.

84. Leyser O. Auxin signaling //Plant Physiology. – 2018 – T.176.– №. 1. – C. 465-479.

85. Li W. et al. DNA methylation and histone modifications regulate de novo shoot regeneration in Arabidopsis by modulating WUSCHEL expression and auxin signaling //PLoS genetics. – 2011. – T. 7. – №. 8. – C. e1002243.

86. Li C. et al. Verification of DNA motifs in Arabidopsis using CRISPR/Cas9-mediated mutagenesis //Plant biotechnology journal. – 2018. – T. 16. – №. 8. – C. 1446-1451.

87. Liao C.-Y., Smet W, Brunoud G., Yoshida S., Vernoux T., Weijers D. Reporters for sensitive and quantitative measurement of auxin response //Nature methods. – 2015. – T. 12. – №. 3. – C. 207.

88. Liu J. S. The collapsed Gibbs sampler in Bayesian computations with applications to a gene regulation problem //Journal of the American Statistical Association. – 1994. – T. 89. – №. 427. – C. 958-966.

89. Liu X. S., Brutlag D. L., Liu J. S. An algorithm for finding protein–DNA binding sites with applications to chromatin-immunoprecipitation microarray experiments //Nature biotechnology. – 2002. – T. 20. – №. 8. – C. 835.

90. Lu H., Zou Y., Feng N. Overexpression of AHL20 negatively regulates defenses in Arabidopsis //Journal of integrative plant biology. – 2010. – T. 52. – №. 9. – C. 801-808.

91. Lynch T., Erickson B. J., Finkelstein R. R. Direct interactions of

ABA-insensitive (ABI)-clade protein phosphatase (PP) 2Cs with calcium-dependent protein kinases and ABA response element-binding bZIPs may contribute to turning off ABA response //Plant molecular biology. – 2012. – T. 80. – №. 6. – C. 647-658.

92. dos Santos Maraschin F., Memelink J., Offringa R. Auxin-induced, SCFTIR1-mediated poly-ubiquitination marks AUX/IAA proteins for degradation //The Plant Journal. – 2009. – T. 59. – №. 1. – C. 100-109.

93. Martínez-García J. F., Huq E., Quail P. H. Direct targeting of light signals to a promoter element-bound transcription factor //Science. – 2000. – T. 288. – №. 5467. – C. 859-863.

94. McCarthy J.D., Chen Y., Smyth G.K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation // Nucleic Acids Research - 2012 - 40(10) - P. 4288–4297.

95. Mitsuda N. et al. Efficient yeast one-/two-hybrid screening using a library composed only of transcription factors in *Arabidopsis thaliana* //Plant and Cell physiology. – 2010. – T. 51. – №. 12. – C. 2145-2151.

96. Mironova, V.V., Omelyanchuk, N.A., Wiebe, D.S., Levitsky, V.G. Computational analysis of Auxin Responsive Elements in the *Arabidopsis thaliana* L. genome // BMC Genomics - 2014 г. -N 15

97. Mironova V. et al. The systems biology of auxin in developing embryos //Trends in plant science. – 2017. – T. 22. – №. 3. – C. 225-235.

98. Mockaitis K., Estelle M. Auxin receptors and plant development: a new signaling paradigm //Annual review of cell and developmental biology. – 2008. – T. 24. – C. 55-80.

99. Mutte S. K. et al. Origin and evolution of the nuclear auxin response system //Elife. – 2018. – T. 7. – C. e33399.

100. Nemhauser J. L., Feldman L. J., Zambryski P. C. Auxin and ETTIN in *Arabidopsis gynoecium* morphogenesis // *Development*. – 2000. – T. 127. – №. 18. – C. 3877-3888.
101. Nicolas M., Cubas P. TCP factors: new kids on the signaling block // *Current opinion in plant biology*. – 2016. – T. 33. – C. 33-41.
102. Nishizawa-Yokoi A. et al. HsfA1d and HsfA1e involved in the transcriptional regulation of HsfA2 function as key regulators for the Hsf signaling network in response to environmental stress // *Plant and Cell Physiology*. – 2011. – T. 52. – №. 5. – C. 933-945.
103. Nodine M. D., Yadegari R., Tax F. E. RPK1 and TOAD2 are two receptor-like kinases redundantly required for *Arabidopsis* embryonic pattern formation // *Developmental cell*. – 2007. – T. 12. – №. 6. – C. 943-956.
104. Nolan T., Hands R. E., Bustin S. A. Quantification of mRNA using real-time RT-PCR // *Nature protocols*. – 2006. – T. 1. – №. 3. – C. 1559.
105. Oh E. et al. Cell elongation is regulated through a central circuit of interacting transcription factors in the *Arabidopsis* hypocotyl // *Elife*. – 2014. – T. 3. – C. E03031.
106. Okushima Y. et al. ARF7 and ARF19 regulate lateral root formation via direct activation of LBD/ASL genes in *Arabidopsis* // *The Plant Cell*. – 2007. – T. 19. – №. 1. – C. 118-130.
107. O'Malley R. C. et al. Cistrome and epicistrome features shape the regulatory DNA landscape // *Cell*. – 2016. – T. 165. – №. 5. – C. 1280-1292.
108. Orchard K., May G. E. An EMSA-based method for determining the molecular weight of a protein-DNA complex // *Nucleic acids research*. – 1993. – T. 21. – №. 14. – C. 3335.
109. Pagès H. et al. Biostrings: Efficient manipulation of biological strings // *R package version*. – 2017. – T. 2. – №. 0.

110. Paponov I. A. et al. Comprehensive transcriptome analysis of auxin responses in *Arabidopsis* //Molecular Plant. – 2008. – T. 1. – №. 2. – C. 321-337.
111. Pan Y. J. et al. A novel WRKY-like protein involved in transcriptional activation of cyst wall protein genes in *Giardia lamblia* //Journal of Biological Chemistry. – 2009. – T. 284. – №. 27. – C. 17975-17988.
112. Péret B. et al. Auxin regulates aquaporin function to facilitate lateral root emergence //Nature cell biology. – 2012. – T. 14. – №. 10. – C. 991.
113. Piya S. et al. Protein-protein interaction and gene co-expression maps of ARFs and Aux/IAAs in *Arabidopsis* //Frontiers in plant science. – 2014. – T. 5. – C. 744.
114. Pufky J., Qiu Y., Rao M. V., Hurban P., Jones A.M. The auxin-induced transcriptome for etiolated *Arabidopsis* seedlings using a structure/function approach //Functional & integrative genomics. – 2003. – T. 3. – №. 4. – C. 135-143.
115. Rademacher E. H. et al. Different auxin response machineries control distinct cell fates in the early plant embryo //Developmental cell. – 2012. – T. 22. – №. 1. – C. 211-222.
116. Raschke A. et al. Natural variants of ELF3 affect thermomorphogenesis by transcriptionally modulating PIF4-dependent auxin response genes //BMC plant biology. – 2015. – T. 15. – №. 1. – C. 197.
117. Ripoll J. J. et al. microRNA regulation of fruit growth //Nature Plants. – 2015. – T. 1. – №. 4. – C. 15036.
118. Ritchie M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies //Nucleic acids research. – 2015. – T. 43. – №. 7. – C. e47-e47.

119. Rivero L. et al. Handling Arabidopsis plants: growth, preservation of seeds, transformation, and genetic crosses //Arabidopsis protocols. – Humana Press, Totowa, NJ, 2014. – C. 3-25.
120. Perrot-Rechenmann C. Cellular responses to auxin: division versus expansion //Cold Spring Harbor perspectives in biology. – 2010. – T. 2. – №. 5. – C. a001446.
121. Del Pozo J. C. et al. AXR1-ECR1–dependent conjugation of RUB1 to the Arabidopsis cullin AtCUL1 is required for auxin response //The Plant Cell. – 2002. – T. 14. – №. 2. – C. 421-433.
122. Prigge M. J. et al. Physcomitrella patens auxin-resistant mutants affect conserved elements of an auxin-signaling pathway //Current Biology. – 2010. – T. 20. – №. 21. – C. 1907-1912.
123. Prigge M. J. et al. The Arabidopsis auxin receptor F-box proteins AFB4 and AFB5 are required for response to the synthetic auxin picloram //G3: Genes, Genomes, Genetics. – 2016. – T. 6. – №. 5. – C. 1383-1390.
124. Robinson M.D., McCarthy D.J., Smyth G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data // Bioinformatics - 2010 - 26 - P. 139–140.
125. Rose A., Meier I., Wienand U. The tomato I-box binding factor LeMYBI is a member of a novel class of Myb-like proteins //The Plant Journal. – 1999. – T. 20. – №. 6. – C. 641-652.
126. Saiki R. K. et al. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia //Science. – 1985. – T. 230. – №. 4732. – C. 1350-1354.
127. Salehin M., Bagchi R., Estelle M. SCFTIR1/AFB-based auxin perception: mechanism and role in plant growth and development //The Plant Cell. – 2015. – T. 27. – №. 1. – C. 9-19.

128. Santamaria M. et al. The promoter of a basic PR1-like gene, AtPRB1, from *Arabidopsis* establishes an organ-specific expression pattern and responsiveness to ethylene and methyl jasmonate // *Plant Molecular Biology*. – 2001. – T. 47. – №. 5. – C. 641-652.
129. Sequeira-Mendes J. et al. The functional topography of the *Arabidopsis* genome is organized in a reduced number of linear motifs of chromatin states // *The Plant Cell*. – 2014. – T. 26. – №. 6. – C. 2351-2366.
130. Sessions A. et al. ETTIN patterns the *Arabidopsis* floral meristem and reproductive organs // *Development*. – 1997. – T. 124. – №. 22. – C. 4481-4491.
131. Shaikhali J. et al. Redox-mediated mechanisms regulate DNA binding activity of the G-group of basic region leucine zipper (bZIP) transcription factors in *Arabidopsis* // *Journal of Biological Chemistry*. – 2012. – T. 287. – №. 33. – C. 27510-27525.
132. Shin R. et al. The *Arabidopsis* transcription factor MYB77 modulates auxin signal transduction // *The Plant Cell*. – 2007. – T. 19. – №. 8. – C. 2440-2453.
133. Simonini S. et al. A noncanonical auxin-sensing mechanism is required for organ morphogenesis in *Arabidopsis* // *Genes & Development*. – 2016. – T. 30. – №. 20. – C. 2286-2296.
134. Spartz A. K. et al. SAUR inhibition of PP2C-D phosphatases activates plasma membrane H⁺-ATPases to promote cell expansion in *Arabidopsis* // *The Plant Cell*. – 2014. – T. 26. – №. 5. – C. 2129-2142.
135. Spensley M. et al. Evolutionarily conserved regulatory motifs in the promoter of the *Arabidopsis* clock gene LATE ELONGATED HYPOCOTYL // *The Plant Cell*. – 2009. – T. 21. – №. 9. – C. 2606-2623.
136. Smalle J., Vierstra R. D. The ubiquitin 26S proteasome proteolytic

pathway //Annu. Rev. Plant Biol. – 2004. – T. 55. – C. 555-590.

137. Staswick P. E. et al. Characterization of an Arabidopsis enzyme family that conjugates amino acids to indole-3-acetic acid //The Plant Cell. – 2005. – T. 17. – №. 2. – C. 616-627.

138. Stigliani, A., Martin-Arevalillo, R., Lucas, J., Bessy, A., Vinos-Poyo, T., Mironova, V., Parcy, F. Capturing auxin response factors syntax using DNA binding models //Molecular plant. – 2018.

139. Stirnberg P. et al. FHY3 promotes shoot branching and stress tolerance in Arabidopsis in an AXR1-dependent manner //The Plant Journal. – 2012. – T. 71. – №. 6. – C. 907-920.

140. Stormo G. D. DNA binding sites: representation and discovery //Bioinformatics. – 2000. – T. 16. – №. 1. – C. 16-23.

141. Suer S. et al. WOX4 imparts auxin responsiveness to cambium cells in Arabidopsis //The Plant Cell. – 2011. – T. 23. – №. 9. – C. 3247-3259.

142. Suzuki J. Y. et al. Unique architecture of the plastid ribosomal RNA operon promoter recognized by the multisubunit RNA polymerase in tobacco and other higher plants //The Plant Cell. – 2003. – T. 15. – №. 1. – C. 195-205.

143. Szemenyei H., Hannon M., Long J. A. TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis //Science. – 2008. – T. 319. – №. 5868. – C. 1384-1386.

144. Tan X. et al. Mechanism of auxin perception by the TIR1 ubiquitin ligase //Nature. – 2007. – T. 446. – №. 7136. – C. 640.

145. Taylor R. G., Walker D. C., McInnes R. R. E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing //Nucleic acids research. – 1993. – T. 21. – №. 7. – C. 1677.

146. Tezuka K. et al. A novel abi5 allele reveals the importance of the

conserved Ala in the C3 domain for regulation of downstream genes and salt tolerance during germination in *Arabidopsis* // *Plant signaling & behavior*. – 2013. – T. 8. – №. 3. – C. e23455.

147. Thimann K. V. Auxins and the growth of roots // *American Journal of Botany*. – 1936. – T. 23. – №. 8. – C. 561-569.

148. Tiwari S. B., Hagen G., Guilfoyle T. J. Aux/IAA proteins contain a potent transcriptional repression domain // *The Plant Cell*. – 2004. – T. 16. – №. 2. – C. 533-543.

149. Trenner J. et al. Auxin-induced expression divergence between *Arabidopsis* species may originate within the TIR1/AFB–AUX/IAA–ARF module // *Journal of experimental botany*. – 2016. – T. 68. – №. 3. – C. 539-552.

150. Udvardi M. K., Czechowski T., Scheible W. R. Eleven golden rules of quantitative RT-PCR // *The Plant Cell*. – 2008. – T. 20. – №. 7. – C. 1736-1737.

151. Ulmasov T., Hagen G., Guilfoyle T. J. Dimerization and DNA binding of auxin response factors // *The Plant Journal*. – 1999. – T. 19. – №. 3. – C. 309-319.

152. Ulmasov T., Murfett J., Hagen G., Guilfoyle T.J. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements // *The Plant Cell*. – 1997. – T. 9. – №. 11. – C. 1963-1971.

153. Valouev A., Johnson D.S., Sundquist A. Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data // *Nature methods*. – 2008. – T. 5. – №. 9. – C. 829.

154. Vanneste S., Friml J. Auxin: a trigger for change in plant development // *Cell*. – 2009. – T. 136. – №. 6. – C. 1005-1016.

155. Varaud E. et al. AUXIN RESPONSE FACTOR8 regulates Arabidopsis petal growth by interacting with the bHLH transcription factor BIGPETALp //The Plant Cell. – 2011. – T. 23. – №. 3. – C. 973-983.
156. Vernoux T. et al. The auxin signalling network translates dynamic input into robust patterning at the shoot apex //Molecular systems biology. – 2011. – T. 7. – №. 1. – C. 508.
157. Vert G. et al. Integration of auxin and brassinosteroid pathways by Auxin Response Factor 2 //Proceedings of the National Academy of Sciences. – 2008. – T. 105. – №. 28. – C. 9829-9834.
158. Villalobos L. I. A. C. et al. A combinatorial TIR1/AFB–Aux/IAA co-receptor system for differential sensing of auxin //Nature chemical biology. – 2012. – T. 8. – №. 5. – C. 477.
159. Vityaev E. E. et al. Computer System Gene Discovery for Promoter Structure Analysis //In silico biology. – 2002. – T. 2. – №. 3. – C. 257-262.
160. Vityaev E. E. et al. Software for analysis of gene regulatory sequences by knowledge discovery methods //Bioinformatics of Genome Regulation and Structure II. – Springer, Boston, MA, 2006. – C. 491-498.
161. Walcher C. L., Nemhauser J. L. Bipartite promoter element required for auxin response //Plant Physiology. – 2012. – T. 158. – №. 1. – C. 273-282.
162. Wang R., Estelle M. Diversity and specificity: auxin perception and signaling through the TIR1/AFB pathway //Current opinion in plant biology. – 2014. – T. 21. – C. 51-58.
163. Wasserman W. W., Sandelin A. Applied bioinformatics for the identification of regulatory elements //Nature Reviews Genetics. – 2004. – T. 5. – №. 4. – C. 276.
164. Weijers D., Wagner D. Transcriptional responses to the auxin

hormone //Annual review of plant biology. – 2016. – T. 67. – C. 539-574.

165. Weirauch M. T. et al. Determination and inference of eukaryotic transcription factor sequence specificity //Cell. – 2014. – T. 158. – №. 6. – C. 1431-1443.

166. Wendrich J. R. et al. Ligation-independent cloning for plant research //Plant Functional Genomics. – Humana Press, New York, NY, 2015. – C. 421-431.

167. Weiste C., Dröge-Laser W. The Arabidopsis transcription factor bZIP11 activates auxin-mediated transcription by recruiting the histone acetylation machinery //Nature communications. – 2014. – T. 5. – C. 3883.

168. Wu M. F. et al. Auxin-regulated chromatin switch directs acquisition of flower primordium founder fate //Elife. – 2015. – T. 4. – C. E09269.

169. Xu M. et al. Developmental functions of miR156-regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) genes in Arabidopsis thaliana //PLoS genetics. – 2016. – T. 12. – №. 8. – C. e1006263.

170. Yamaguchi-Shinozaki K., Shinozaki K. Organization of cis-acting regulatory elements in osmotic-and cold-stress-responsive promoters //Trends in plant science. – 2005. – T. 10. – №. 2. – C. 88-94.

171. Yamamoto Y. Y. et al. Heterogeneity of Arabidopsis core promoters revealed by high-density TSS analysis //The Plant Journal. – 2009. – T. 60. – №. 2. – C. 350-362.

172. Yoshida T. et al. AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation //The Plant Journal. – 2010. – T. 61. – №. 4. – C. 672-685.

173. Yu X. et al. A brassinosteroid transcriptional network revealed by

genome-wide identification of BES1 target genes in *Arabidopsis thaliana* //The Plant Journal. – 2011. – T. 65. – №. 4. – C. 634-646.

174. Zambryski P. et al. Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity //The EMBO journal. – 1983. – T. 2. – №. 12. – C. 2143-2150.

175. Zemlyanskaya E. V. et al. Meta-analysis of transcriptome data identified TGTCNN motif variants associated with the response to plant hormone auxin in *Arabidopsis thaliana* L //Journal of bioinformatics and computational biology. – 2016. – T. 14. – №. 02. – C. 1641009.

176. Zhang Z. W. et al. The roles of two transcription factors, ABI4 and CBFA, in ABA and plastid signalling and stress responses //Plant molecular biology. – 2013. – T. 83. – №. 4-5. – C. 445-458.

ПРИЛОЖЕНИЯ

Приложение 1. Используемые для ОТ-количественной ПЦР последовательности праймеров.

Матрица	Направление праймера	Последовательность праймера
AT2G23170	F	TCTGCGATCTCCGATGAT
AT2G23170	R	TAATCACTTTCTTCTGAACGAA
AT3G58190	F	ACGGTTCAACACGGGGAT
AT3G58190	R	TCACGAGAAGGAGATGTAGC
AT4G23750	F	CCGAGAATCAAATTCACAGAG
AT4G23750	R	GGTTTGGTGAGTGTTGGT
AT5G67060	F	AGTCCGATCACTCAATGA
AT5G67060	R	GCTGTGTTGGGAGAGATA
AT3G27010	F	AAATGTAGCTGGGTTTGG
AT3G27010	R	CAGGCATCTGATTATGATTC
AT4G32280	F	GGTGCTGCGTCTTGTTTG
AT4G32280	R	AATTTTCCTCCACATTCCCTAAC
AT2G47260	F	TTTGGGTTTCCTATTGAT
AT2G47260	R	CTATTAAGGTA CT CGTTGA
AT5G65640	F	CACCGTTACAACCACTGT
AT5G65640	R	AGGGATACGACATCGGAT
AT3G60630	F	CAACAACAACGACGACACTAAC
AT3G60630	R	GAGAGAGACGGTGGTGATAATG
AT5G18560	F	ATTCAAGCGTTTGGTAATAG
AT5G18560	R	GTTTAAGGCAGTGGTTTG
AT1G75820	F	CATTTGGTGAATCTAACTTT
AT1G75820	R	GAATGTTCCAGTAAGGTT
AT5G25350	F	TCTGACGATTGATTCTTG
AT5G25350	R	GCTCCTGATAGAGATAGAT

AT4G10310	F	TAGAAGGAATGAGTTCGTA
AT4G10310	R	AAGTGTATGGAGGAAGATA
AT4G37610	F	CAATACATCTACTTGTGTTATC
AT4G37610	R	TCTATGACAGAGAAGACC
AT5G17400	F	CCATTGTGTTAGTCGGTTC
AT5G17400	R	GCAGAGGTAGTAATACTCCAA
AT2G45890	F	TGCCAGTCAGATACATAA
AT2G45890	R	CAGTTGCTTGTAATCAC
AT3G13360	F	CTCGGTTTCTGTTGGTAA
AT3G13360	R	TTCGCTTCTGCTATATGTT
AT1G79430	F	ATCTACGAAAGGAAATCA
AT1G79430	R	TCATCATAGGACTCAATG
AT5G14340	F	TCTCTATTTGAAACGCAATG
AT5G14340	R	ATCCGTCATATCCAAGTC
AT2G28110	F	CAGTGGCTCTGATCATGTCTT
AT2G28110	R	CTCCATCAGCAATAGCTCTATCC
AT2G01950	F	CGCTGCTTCGGTTTGTATTT
AT2G01950	R	CGGCTCTTTCTCCTTCTCTATC
AT4G36670	F	TTATGAGTGGAGCGATGG
AT4G36670	R	GAGCAATGATCCGACAAG
AT2G25000	F	TCGAGAACATCACCAACGATAAA
AT2G25000	R	GCCCGTATTTCCCTCCATTGATA
AT3G54770	F	TCGCCGTTCCAAAGGATATG
AT3G54770	R	GCCATTGATGATCGGAGTAGAG
AT1G75750	F	AAGCTCTTATCGCTTCTCTTCTC
AT1G75750	R	CGATCTTCTTTGCGTAACCATTT
AT2G345000	F	CGGACATCCCTTTGGTAAGAAG
AT2G345000	R	GTGCCTTGGGAGTGAAGTTAG
AT2G200750	F	CGTCAGGACCATCTGCTAAAG
AT2G200750	R	CGGCGGTAGAGGATGTTTAAT
AT2G340070	F	GGACCTCCTCGTCTTCAATTC

AT2G340070	R	CTCATCAGTGAAGACCCATCTC
AT1G004610	F	GGCTTAAGGACAACGACTTCT
AT1G004610	R	TGAACCCTACCGCATACAATC
AT1G041000	F	AGGTTCGATCGTCTTTGTCTTC
AT1G041000	R	CCTACAGCAACTTGTCTGGTT
AT1G29440	F	ACCATTCGATTCGGCCTTT
AT1G29440	R	AGAGCATCTAGCACTTGAGATTG
AT5G47370	F	CGATCCTACATCAGATCTTCGC
AT5G47370	R	GATTCCTTCTCTCTCACTTCTCTTC
AT3G621000	F	TGAAAGTGAACATGGAAGGAGTT
AT3G621000	R	ATCTTCTTCTTCAGCCCAAAGG
AT1G17140	F	CCAAGACCAAGAGTTTCAGAGT
AT1G17140	R	AGTGATCAGACGGTTGGAATG
AT5G54490	F	AAGGGTTCGAGCTTCTTATGG
AT5G54490	R	GAACATCGTCGTCCGTTAGAT
AT4G13195	F	CTACAACCTTCCACTCCATCTT
AT4G13196	R	TCAGCTCTGAACTCTCTCCT
AT5G52860	F	TGCTCCCAAACCCATCAATAG
AT5G52860	R	TTCTTTCGCCACCGGATAAG
AT5G53250	F	GCCGATTGCTCAATCTCATTC
AT5G53250	R	GCCAAAGCCACCATCATTAG
AT3G22740	F	CAGATGACACGCAAACCAATC
AT3G22740	R	CTCTGACTCTCCCTCTGACTTA
AT5G63780	F	GCTATCCCTGACCAACCTTTAG
AT5G63780	R	ACCTTCCACAGCCCATTTAG
GATA23-GFP	F	CAAACCCAACCTCCATAAG
GATA23-GFP	R	ATGAACTTCAGGGTCAG

Приложение 2. Использованные для клонирования последовательности праймеров.

Матрица	Направление праймера	Последовательность праймера
AGP22prom	F	TAGTTGGAATGGGTGATCCACGGGAGAA ACTTTCT
AGP22prom	R	TTATGGAGTTGGGTTTTTATCTCCGGTGAT TATTCC
AGP22TATATC	F	CTTTGCTGCtttttATAGCTTA
AGP22TATATC	R	GAAAGAACTAAGCTATAaaaaaGCAGCAA
AGP22TGTCTC	F	CTTATAGCACtttttTTTGCTGCTT
AGP22TGTCTC	R	GAATATCGTGaaaaaGTGCTA
AGP22TATATC-TGTC TC	F	GCTATAaaaaaGCAGCAAAaaaaaGTGCTT
AGP22TATATC-TGTC TC	R	GCACtttttTTTGCTGCtttttATA
IAA30prom	F	TAGTTGGAATGGGTTTACTATTTGTAATCC GTAAAAC
IAA30prom	R	TTATGGAGTTGGGTTCATTTTTTTTTATTTC TTTTAC
IAA30TGTCTG	F	CTTTGACCAAGtttttCATGA
IAA30TGTCTG	R	GTCATGaaaaaCTTGGTCAA
IAA30CACGTG	F	TGTCTGCATGAtttttTATTT
IAA30CACGTG	R	GACAAATAaaaaaTCATGC
IAA30TATTTG	F	CACGTGtttttTCTGCTAC
IAA30TATTTG	R	GCAGACaaaaaCACGTGTCA
IAA30TTGACC	F	CCAACACTAAAAAAAAAGTGTCTGCATGAC A
IAA30TTGACC	R	GCAGACACTTTTTTTTAGTGTTGGTTTTGA CATGACACGTGTATTAAAAACTACATGG ACA
IAA30tatt-TGTCTG	F	ACA
IAA30tatt-TGTCTG	R	GTCCATGTAGTTTTTTAATACACGTGTCAT

ARP1prom	F	TAGTTGGAATGGGTTCGAAAAGTGAAAAC GAAAGGTAAAACC
ARP1prom	R	TTATGGAGTTGGGTTCGAAAGAAAGAAAG GGCAGAGATA
ARP1-TTGACA	F	CTTTtttttTTATTAAGCTTATCTTATACT TAGTATAAGATAAGCTTAATAAaaaaaaAAA G
ARP1-TTGACA	R	CTTTTTGACATTATTAAGCTTAttttttACT TAGTaaaaaaTAAGCTTAATAATGTCAAAAA G
ARP1-TCTTAT	F	CTTTtttttTTATTAAGCTTAttttttACT
ARP1-TCTTAT	R	TAGTaaaaaaTAAGCTTAATAAaaaaaaAAAG
ARP1-TCTTAT-TTGA CA	F	TAGTTGGAATGGGTTCGAACAATTATTTGT TATAGGTAGACG
ARP1-TCTTAT-TTGA CA	R	TTATGGAGTTGGGTTCGAAGTGTAAGTGA CATTATTAATC
BRL2prom	F	CAAttttttATTCCTCCGACTCTTATTA CTAATAAGAGTCGGAAGGTAAaaaaaaTTGT AGAATCGTCT
BRL2prom	R	CAACGGACAATTCCTCCGACtttttTA CTAaaaaaaGTCGGAAGGTAATGTCCGTTGTA GAATCGTCT
BRL2-TGTCCG	F	CAAttttttATTCCTCCGACTtttttTA
BRL2--TGTCCG	F	CTAaaaaaaGTCGGAAGGTAAaaaaaaTTGTAGA ATCGTCT
BRL2-TCTTAT	R	TAGTTGGAATGGGTGTTACGGGGATTAG CCTTTGC
BRL2-TCTTAT	F	TTATGGAGTTGGGTTTGGGCGCTCTTTTGA TGAACA
BRL2-TGTCCG-TCTT AT	R	CCACAAtttttTGTATTGTTCA
BRL2-TGTCCG-TCTT AT	F	GAGATCTGAACAAATAACAAaaaaaaTGTG CCACATGGACATTGTtttttTTCA
GATA23prom	R	
GATA23prom	F	
GATA23-TGGACA	R	
GATA23-TGGACA	F	
GATA23-TATTTG	R	

GATA23-TATTTG	F	GAGATCTGAAaaaaaaACAATGTCCATGTG
GATA23-TGGACA-TA TTTG	R	CCACAtttttTTGTTATTTGTTCA
GATA23-TGGACA-TA TTTG	F	GAGATCTGAACAAATAACAAaaaaaaTGTG TAGTTGGAATGGGTTTAGACAATGAAAAT TATGTATGGAGAAGCTTACTCG TTATGGAGTTGGGTTCACTATAATGTAGTC GCATATATGT
AGP7prom	F	GCACATGGGATATACGATAATAtttttT GAAAGAACGAaaaaaaTATTATCGTATATCCC AT
AGP7prom	R	GCACATGGtttttCGATAATAACGACAT GAAAGAACGATGTCGTTATTATCGaaaaaaCC AT
AGP7-ACGACA	F	GCACATGGtttttCGATAATAtttttT
AGP7-ACGACA	R	GAAAGAACGAaaaaaaTATTATCGaaaaaaCCAT
AGP7-GATATA	F	
AGP7-GATATA	R	
AGP7-ACGACA-GAT ATA	F	
AGP7-ACGACA-GAT ATA	R	

Приложение 3. Используемые для секвенирования последовательности праймеров.

Матрица	Направление праймера	Последовательность праймера
AGP22seq1	F	TCGAGGTCGACGGTATCGATAAGCTTGATATCGA ATTCTAGTTG
AGP22seq1	R	CGAAGAAGGAATCTTTTGGTTTTGTAGAATCTCA
IAA30seq1	F	GTAACGCCAGGGTTTTCCCA
IAA30seq1	R	CCCGAAACATCCAAACAATG
IAA30seq2	F	GTTTTTGGCTCAATGATAAAGACAC
IAA30seq2	R	CAGAGACCAAACACATCATATTG
IAA30seq3	F	ATTTTGAATACTAACGACGGATAAGACTAAGATC ACC
IAA30seq3	R	GGGTTCGAAGATTTGAACGTGTGTATTATATA TAGAGAATGG
GATA23seq1	F	GCGTCATCGGCGGGGGTCATAACGTGACTCCCTT AAT
GATA23seq1	R	AAAGTGGCAGTACGCACGACCG
BRL2seq1	F	CTCCCTTAATTCTCCGCTCA
BRL2seq1	R	CCATCCGACCACTATAGTTTCA
BRL2seq2	F	ACACAGAAATCCACCTAATGATTC
BRL2seq2	R	CCAACGAGAAAATGCTATTTTG
BRL2seq3	F	TTGAATGGTAGCTAGGGTATAACC
BRL2seq3	R	GGGAATTGTATCTATGAGATGTTTC
AGP7seq1	F	CTGCAAGGCGATTAAGTTGG
AGP7seq1	R	CGAAACCAACCGACGATTAT
AGP7seq2	F	GCATATGTAACAAAGCAAGTTATAGTGC
AGP7seq2	R	GGTTCACTATAATGTAGTCGCATATATG
ARP1seq1	F	ACAAAAGTTAATGCGTCTGCTT
ARP1seq1	R	AACACAAATATACCCGAATAGAGG
ARP1seq2	F	GACGTAACCGAAGGGCTCGA

ARP1seq2	R	TTGGGCCTAACAGCAACT
		TCTAAACAATCAAACAAAAAGGAACAGATGGAG
ARP1seq3	F	TACC
ARP1seq3	R	ACCAATCCCGCAAATGAAAA

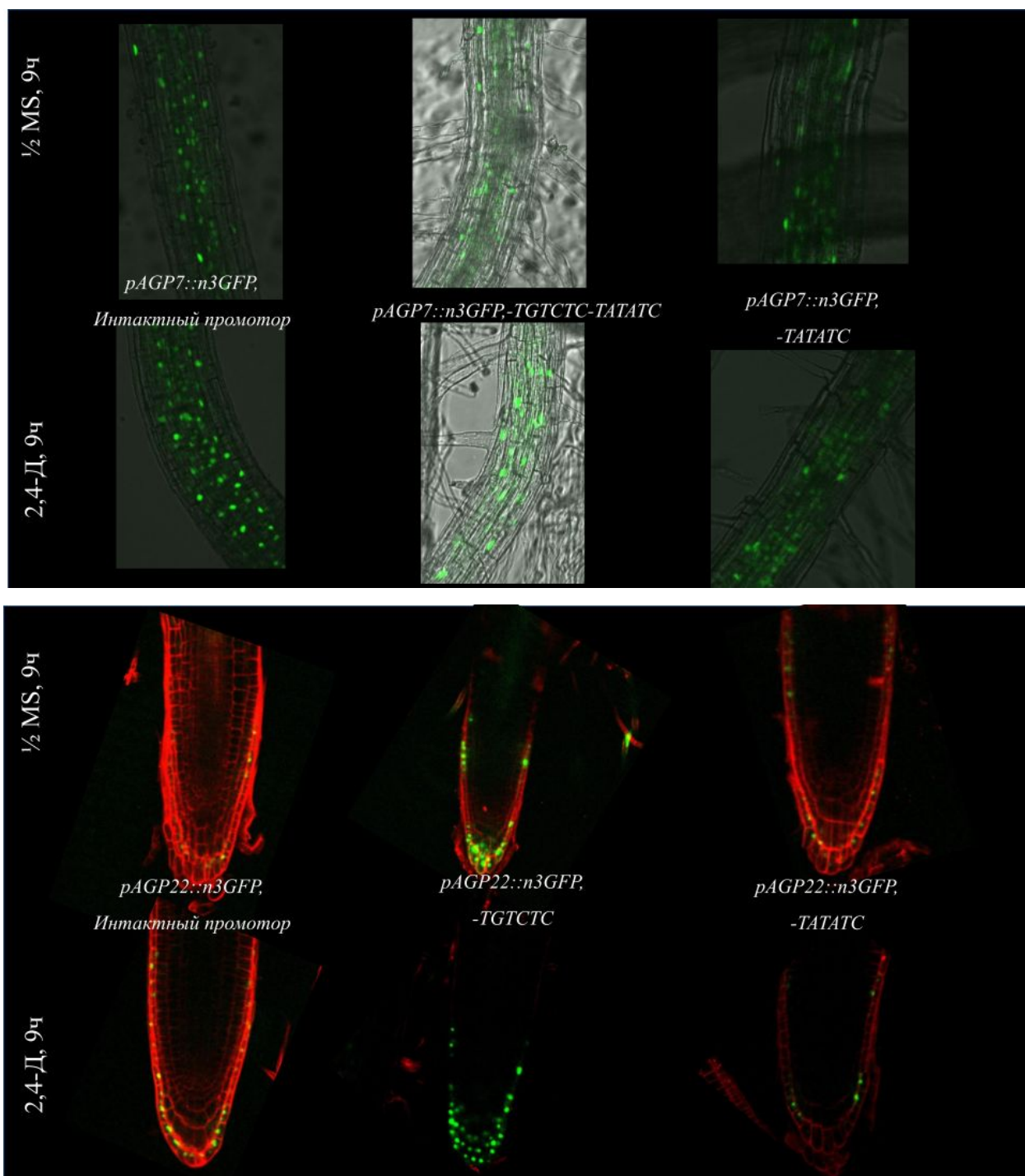
Приложение 4. Результаты ОТ-количественной ПЦР на семядольных листьях, корнях и целых трехдневных проростках *Arabidopsis*, обработанных аукином (1мкМ 2,4-Д) в течение 1 часа.

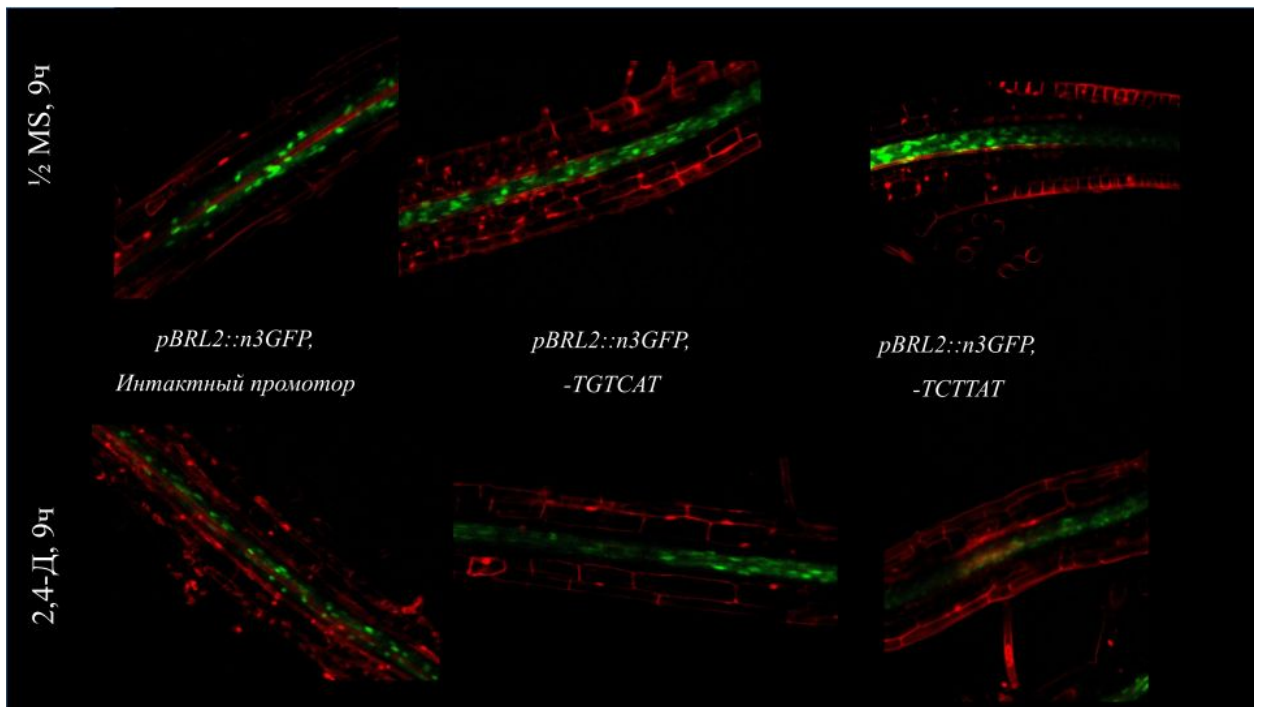
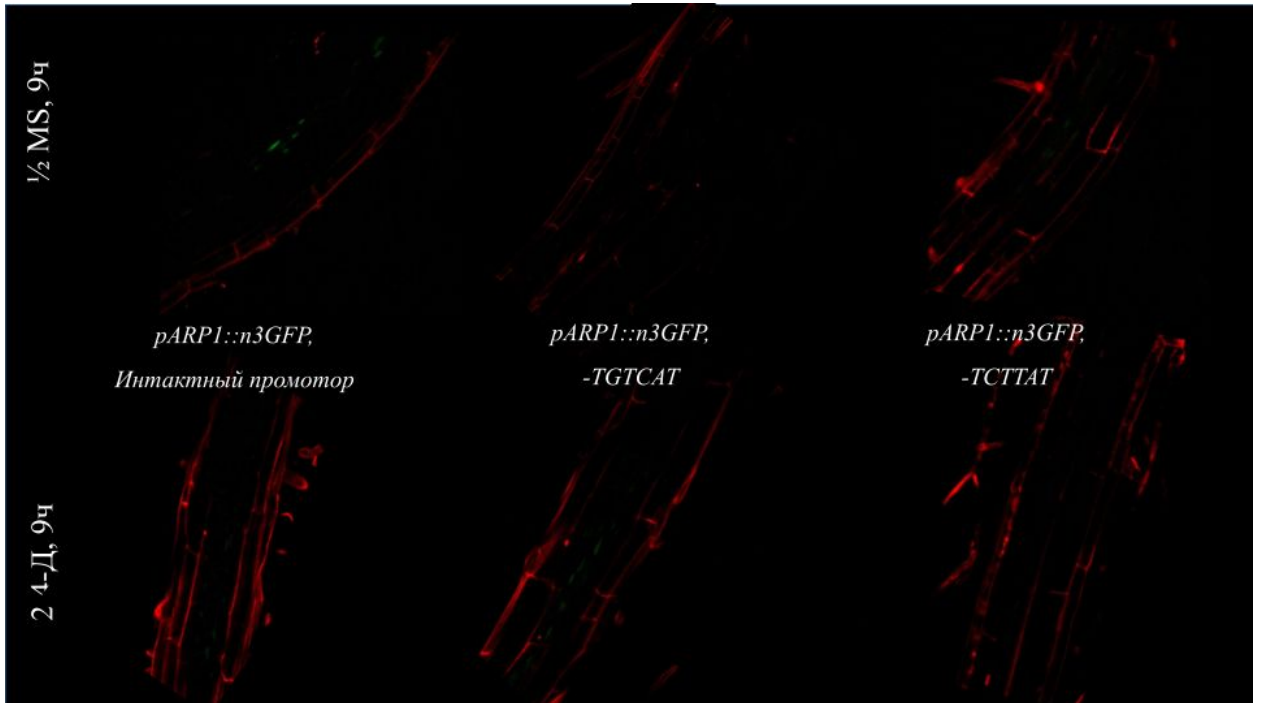
	Название гена	изменение экспрессии		
Предсказание ТАА		корни	семядольные листья	проростки
повышение экспрессии в раннем ответе	ICR1	1.082602533	0.7435527473	0.4214462245
	HAT2	4.444345894	0.4108152953	0.6078686954
	SAUR63	4.962882561	4.075033796	4.375290041
	IAA30	15.14270669	0.1541488764	4.038301202
	IAA10	2.638216345	1.712516569	0.4574653031
	TCP20	1.087231657	1.038643832	0.3411711923
	GH3.3	78.60175573	0.7019245469	14.57148688
	HEC1	0.6362545404	0.9336523063	1.539923294
	ASL16	44.99589155	1.362547076	36.3
	CRF2	1.488308797	2.366095591	0.6313693876
повышение экспрессии в позднем ответе	ABCG8	10.21623688	1.340044827	0.4030703903
	PBP	13.61412902	0.9951117416	3.998189727
	PUCHI	3.617604079	1.206405436	0.2304533084
	EBF2	1.571175401	1.230131391	0.4502560815
	HAM2	2.908378106	1.99193101	0.7070217258
	WRKY23	0.9727804822	0.9140144373	0.8450604675
	bHLH093	1.295645959	1.046193334	0.6446370047
	ATCLV1	1.087486075	1.766313007	1.267519722
	CLE44	5.42483749	1.077192705	2.064243226
	IAA29	23.11419056	1.03709092	1.755680591
понижение экспрессии	AGP22	0.6234192454	0.8869493828	1.111156512
	HMT3	0.8867914645	1.600643892	1.459417578

в раннем ответе	WRKY60	1.233604374	0.8064886989	0.7118802465
	ERANT1	0.783816806	4.5258539	0.5806848451
	ROPGEF4	1.043352009	1.125417976	0.4285137166
	WIP3	0.7903993928	1.155430428	0.5545639453
	SHA1	0.9232686664	5.19651118	1.002305293
	APL	1.329293856	0.9297959452	0.6651829672
	ATHKT1	0.6645761362	1.201433539	0.6169858768
	BT5	1.148922557	2.249385606	0.4848222914
понижение экспрессии в позднем ответе	GASA1	0.740308421	4.336711888	0.6632043101
	CYP710A1	0.9976131785	1.004630798	0.2225373996
	TBL37	0.7562465001	0.997423689	0.1865203702
	ARP1	0.5465816826	0.9635597541	0.1189973231
	EXPB1	0.1193749248	0.965272645	0.08867007402
	MYB40	0.3493768686	1.028419534	0.2544781692
	FRA8	0.9599316749	0.9935497269	0.1160037234
	BRL2	0.7703065434	6.986477137	0.1365138675
	YUC3	0.7840318768	4.839305414	0.7263744807
	PLT6	0.7677432058	206.8681703	0.9454532532

Примечание: зеленым обозначено повышение экспрессии генов, малиновым – понижение.

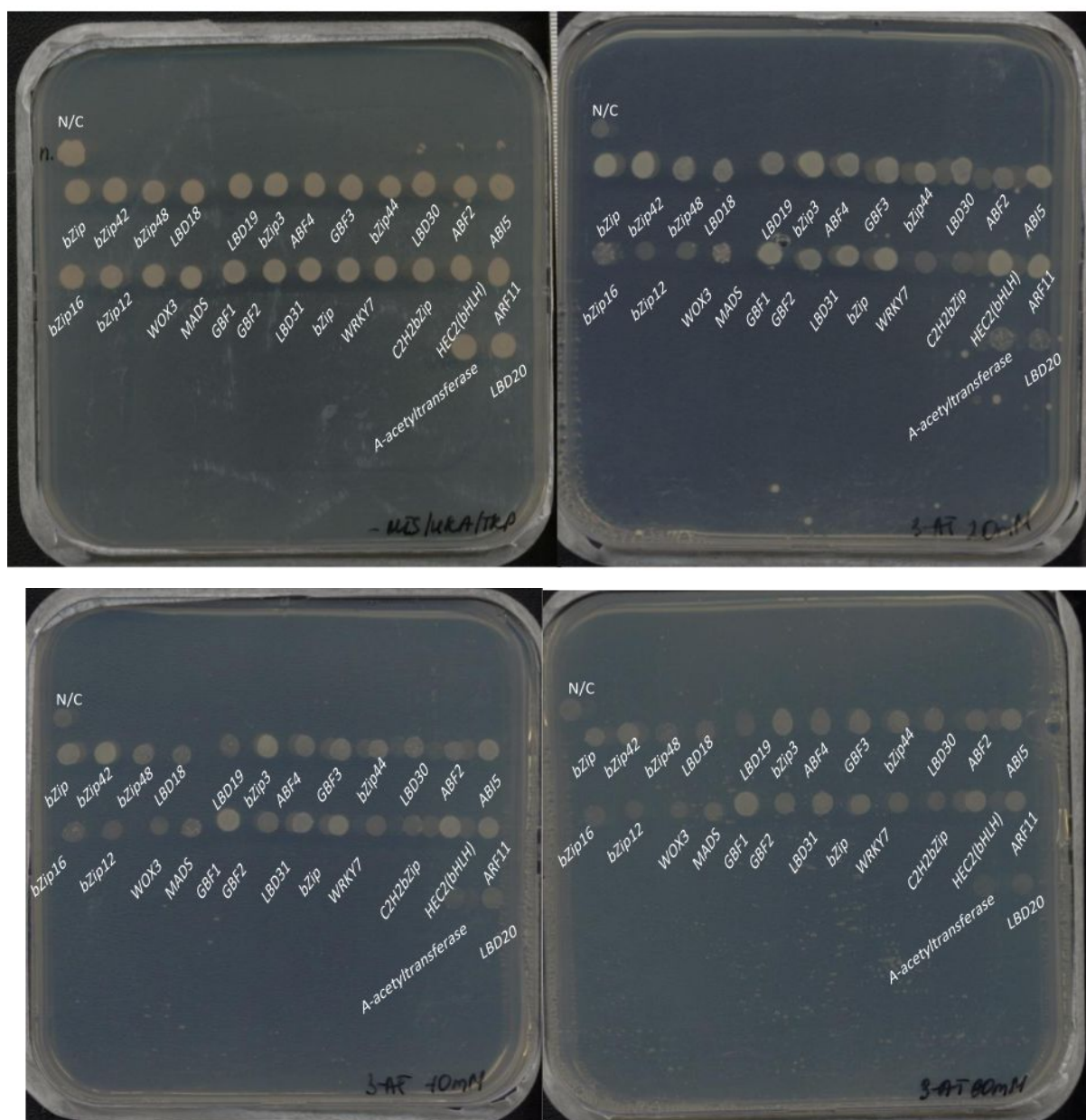
Приложение 5. Экспрессия GFP под интактными и мутированными промоторами исследуемых генов при стандартных условиях и после обработки 2,4-Д.





Приложение 6. Скрининг взаимодействий транскрипционных факторов с регуляторным модулем промотора гена *IAA30* в дрожжевой одногибридной системе на селективной среде.

Примечание: были использованы следующие селективные среды: -HIS/URA/TRP; -HIS/URA/TRP, 20мМ 3-АТ; -HIS/URA/TRP, 40мМ 3-АТ; -HIS/URA/TRP, 60мМ 3-АТ; -HIS/URA/TRP, 100мМ 3-АТ. На чашках Петри соответствующие колонии дрожжей подписаны названием ТФ, с которым проверялось взаимодействие. N/C – отрицательный контроль.





Приложение 7. Список ТФ, взятых в анализ взаимодействий с ТФ ARF с помощью базы данных STRING.

Идентификатор AGI	Семейство ТФ	Название ТФ
AT5G42910	bZIP a,b,c	bZip
AT3G30530	bZIP a,b,c	bZip42
AT2G04038	bZIP a,b,c	bZip48
AT2G45410	LOB a / AS2 b	LBD18
AT2G45420	LOB a / AS2 b	LBD19
AT5G15830	bZIP a,b,c	bZip3
AT3G19290	bZIP a,b,c	ABF4
AT2G46270	bZIP a,b	GBF3
AT1G75390	bZIP a,b,c	bZip44
AT4G00220	LOB a / AS2 b	LBD30
AT1G45249	bZIP a,b,c	ABF2
AT2G36270	bZIP a,b,c	ABI5
AT2G35530	bZIP a,b	bZip16
AT2G41070	bZIP a,b,c	bZip12
AT3G03760	LOB a / AS2 b	LBD20
AT2G28610	HB a,b,c	WOX3
AT5G27810	ND	MADS
AT4G36730	bZIP a,b	GBF1
AT4G01120	bZIP a,b	GBF2
AT4G00210	LOB a / AS2 b	LBD31
AT4G34000	bZIP a,b,c	bZip
AT4G24240	WRKY a,b,c	WRKY7
AT1G02040	C2H2 a,b,c	C2H2bZip
AT3G50330	bHLH a,b	HEC2(bHLH)
AT2G41450	GNAT a	A-acetyltransferase
AT1G59750	ARF a,b,c	ARF1
AT2G28350	ARF a,b,c	ARF10
AT2G46530	ARF a,b,c	ARF11
AT1G34310	ARF a,b,c	ARF12
AT1G34170	ARF a,b,c	ARF13

AT1G35540	ARF a,b,c	ARF14
AT1G35520	ARF a,b,c	ARF15
AT4G30080	ARF a,b,c	ARF16
AT1G77850	ARF a,b,c	ARF17
AT3G61830	ARF a,b,c	ARF18
AT1G19220	ARF a,b,c	ARF19
AT5G62010	ARF a,b,c	ARF2
AT1G35240	ARF a,b,c	ARF20
AT1G34410	ARF a,b,c	ARF21
AT1G34390	ARF a,b,c	ARF22
AT1G43950	ARF a,b,c	ARF23
AT2G33860	ARF a,b,c	ARF3
AT5G60450	ARF a,b,c	ARF4
AT1G19850	ARF a,b,c	ARF5
AT1G30330	ARF a,b,c	ARF6
AT5G20730	ARF a,b,c	ARF7
AT5G37020	ARF a,b,c	ARF8
AT4G23980	ARF a,b,c	ARF9
AT5G62000	ARF a,b,c	