**Результаты работы по третьему этапу (01.01.2019-31.12.2019) проекта “Разработка прототипа инновационной программы селекции и использование генетического разнообразия сортов коллекции Всероссийского института генетических ресурсов им. Н.И.Вавилова для селекционного улучшения сои в целях обеспечения продовольственной̆ безопасности России”**

*Соглашение с Минобрнауки о предоставлении субсидии 14.575.21.0136 от  26.09.2017 на период 2017 - 2019 гг., уникальный идентификатор проекта: RFMEFI57517X0136*

Работы третьего этапа были посвящены анализу данных и обобщению результатов. Критичным для данного этапа было успешное выполнение полногеномного поиска ассоциаций, который связывает естественную вариабельность сортов (ОНП) с фенотипическими признаками. В связи с этим были выполнены картирование прочтений Illumina на рефересный геном сои и поиск ОНП и инсерций/делеций у сортов сои. Поиск ОНП выполнен с возможностью перевыравнивать инсерции/делеции и производить повторную проверку качества нуклеотидов в прочтениях. При поиске ОНП идентификация вариантов проводилась для всех генотипов одновременно, фильтрация ОНП выполнена с использованием стандартных вариантов жесткой фильтрации. Далее был проведен полногеномный поиск ассоциаций (GWAS) между однонуклеотидными полиморфизмами и фенотипическими признаками, а также между условиями культивирования и однонуклеотидными полиморфизмами и получены данные о гаплоблоках и их размерах Были определены последовательности и oхарактеризoваны полиморфизмы целевых локусов, генов и аллельных вариантов, ассоциированных с признаками интереса, у сортов/линий сои, которые в дальнейшем могут выступать в качестве молекулярных маркеров селекционного процесса.

Для оценки геномной ценности селекционного материала сои и проведения селекции сои по комплексу хозяйственно-ценных признаков были созданы две панели из 128 и 28 маркеров для маркерной селекции сортов, представляющие собой набор из молекулярных маркеров на основе ОНП генома сои. Эти панели разработаны на основании анализа нуклеотидного разнообразия сортов, используемых в отечественной селекции, они отражают аллельное разнообразие в таких сортах и это делает их применение в отечественной маркерной селекции предпочтительным.

Для отбора сортов/линий сои по признакам интереса были разработана методика генотипирования образцов сои с использованием созданных молекулярных маркеров, в соответствии с которой последовательности ДНК‒маркеров определяют при помощи секвенирования ампликонов. Согласно этой методике было выполнено тестирование панели маркеров путем генотипирования на 96 образцах сои, предоставленных индустриальным партнером. Pезультат тестирования показал, что полученные молекулярные маркеры дают воспроизводимый результат по определению ОНП сои в 128 локусах, отобранных по результатам проекта.

С применением методов генотипирования путем секвенирования, полногеномного поиска ассоциаций и оценки селекционой ценности генотипов было создано 10 новых селекционных линий сои с улучшенными хозяйственно-ценными признаками ‒ высокой продуктивностью, высоким содержанием белка и масла в семенах.

Маркерная селекция имеет ограниченное применение в случае признаков, контролируемых большим числом генов, и для таких признаков лучше работают методы геномной селекции. Исходя из полигенной природы анализируемых признаков, был проведены сравнительный анализ, тестирование и оптимизация существующих программ геномной селекции и оригинальной программы для расчета по математической модели взаимодействий фенотип/генотип по признакам количество белка, масла и урожайность. Модель генотип/ фенотип была применена для предсказания селекционной ценности генотипов, результаты предсказаний загружены в Базу данных селекционной ценности генотипов для отбора по признакам содержание белка, масла в семенах и урожайность.

Существующие предикативные модели геномной селекции не учитывают, что фенотипические признаки могут быть коррелированы, что может проводить к неправильным оценкам. Таким образом, несмотря на известность ряда методов геномной селекции растений, существует потребность в новых способах оценки селекционного материала по генотипу, которые бы учитывали корреляционные зависимости между признаками. Такая задача в рамках проекта решена путем создания новой предикативной модели, которая описывает связь между совокупностью нескольких фенотипических признаков и ОНП и может быть использована не только для оценки селекционных линий, но и для моделирования сортов и подбора желательных родительских пар при скрещиваниях. Также была разработана Методика оценки селекционного материала по генотипу для подбора желательных родительских пар, идентификации желательных генотипов из популяций. Поданa заявкa на получение патентa 2019139622 “Способ оценки селекционного материала по генотипу для подбора желательных родительских пар, идентификации желательных генотипов из популяций и оценки селекционных линий».

В рамках изучения влияния изменения климата на хозяйственно-ценные признаки сои был исследован диапазон изменчивости агроклиматических требований сортов внутри культуры, определены неблагоприятные факторы и выполнен расчет рисков и проведена оценка экономической эффективности возделывания сои в разных областях РФ, а также суперкомпьютерное имитационное моделирование изменения зон возделывания сои и расчет на суперкомпьютере изменчивости агроклиматических требований сортов внутри культуры. Осуществлены повторный посев и фенотипирование сортов сои по продуктивности и содержанию белка и масла.

Анализ содержания патентов показал, что Программа для ЭВМ для расчета взаимодействий фенотип/генотип по математической модели, базы данных Электронный атлас областей соеводства с учетом изменений климатических условий и База данных селекционной ценности генотипов для отбора по признакам количество белка, масла и урожайность, а также панель маркеров для маркерной селекции по хозяйственно полезным признакам и способ оценки селекционного материала по генотипу для подбора желательных родительских пар, идентификации желательных генотипов из популяций и оценки селекционных линий отвечают всем сформулированным в ходе исследований требованиям, в том числе к программному обеспечению, по теме работ, имеют важные конкурентные преимущества над известными аналогами, обладают патентной чистотой и позволяют получить приоритетные результаты в области исследований. Для разработки баз данных в ходе выполнения Технического задания использована Postgresql v10.0 и выше. Полученные результаты патентных исследований совпадают с требованиями по патентноспособности.

**Результаты интеллектуальной деятельности за третий этап**

Программа для расчета взаимодействий фенотип/генотип по математической модели. Свидетельство о государственной регистрации № 2019662319 от 20.09.2019.

Электронный атлас областей соеводства с учетом изменений климатических условий. Свидетельство о государственной регистрации № 2019621657 от 24.09.2019.

База данных селекционной ценности генотипов для отбора по признакам количество белка, масла и урожайность. Свидетельство о государственной регистрации № 2019622365 от 13.12.2019

Заявка на патент на изобретение № 2019139622 от 05.12.2019 «Способ оценки селекционного материала по генотипу для подбора желательных родительских пар, идентификации желательных генотипов из популяций и оценки селекционных линий».

Заявка на патент на изобретение № 2019139627 от 05.12.2019 «Mаркеры для маркерной селекции по хозяйственно-полезным признакам».

**Публикации за третий этап**

О.Д. Таратухин, Л.Ю. Новикова, И.В. Сеферова, Герасимова Т.В., С.В. Нуждин, М.Г. Самсонова, К.Н. Козлов. Нейросетевая модель прогнозирования фенологии скороспелых сортов сои по климатическим факторам. БИОФИЗИКА, 2020, том 65, № 1, с. 125–137

М.В. Грецова, М.Г. Самсонова. Генетическая идентификация сортов сои с использованием однонуклеотидных полиморфизмов Биофизика, 2020, T. 65, № 1, стр. 206-208

# A.A. Канапин, А.Б. Соколкова, А.А. Самсонова, А.В. Щегольков, С.В. Болдырев, А.Ф. Аюпова, Ф.E. Хайтович, С.В. Нуждин, М.Г. Самсонова. Генетические варианты, ассоциированные с продуктивностью и содержанием белка и масла у сои Биофизика, 2020, T. 65, № 2, стр. 280-289