Важность генетических исследований для понимания структуры, функции и адаптации лесных биогеоценозов и для менеджмента сибирских бореальных лесов

Константин Валерьевич Крутовский

Внс лаборатории популяционной генетики Института общей генетики им. Н. И. Вавилова РАН

Научный руководитель Научно-образовательного центра геномных исследований и лаборатории лесной геномики Сибирского федерального университета

Профессор генетики отделения лесной генетики и селекции Гёттингенского университета, Германия

Адьюнкт профессор отделения по изучению и управлению экосистем Техасского агро-механический университета, США

Биогеоценоз

- **Биогеоценоз** (от греч. βίος — жизнь γη — земля + κοινός — общий) — природная саморегулирующаяся устойчивая система, включающая сообщество живых организмов и тесно связанную с ним совокупность абиотических факторов среды (животные, растения) и неорганических факторов (вода, почва) в пределах одной территории, связанных между собой круговоротом веществ и потоком энергии. Примеры: сосновый лес, горная долина. Учение о биогеоценозе разработано Владимиром Сукачёвым в 1942 году.
- В зарубежной литературе — малоупотребимо и вместо этого обычно используется понятие **природная экосистема,** т.е. система, состоящая из взаимосвязанных между собой сообществ организмов разных видов и среды их обитания (термин был введен в 1935 г. английским ботаником А. Тенсли).
- **Экосистема** — более широкое понятие, относящееся к любой подобной системе. **Биогеоценоз**, в свою очередь — класс экосистем, экосистема, занимающая определенный участок суши и включающая основные компоненты среды — почву, растительный покров и атмосферу. (Например, большинство искусственных и морских экосистем не являются биогеоценозами).
- Таким образом, каждый биогеоценоз — это экосистема, но не каждая экосистема — биогеоценоз.

Биогеоценоз

- Определение биогеоценоза по В.Н. Сукачеву (Основы лесной биогеоценологии, М. : Наука, 1964: стр. 23) считается классическим – «... **это совокупность на известном протяжении земной поверхности однородных природных явлений (атмосферы, горной породы, растительности, животного мира и мира микроорганизмов, почвы и гидрологических условий), имеющая особую специфику взаимодействий этих слагающих ее компонентов и определенный тип обмена веществ и энергией: между собой и с другими явлениями природы и представляющая собой внутреннее противоречивое единство, находящееся в постоянном движении и развитии** …".

- В этом определении отражаются все сути биогеоценоза, черты и особенности, присущие только ему:
  - ✓ биогеоценоз должен быть однородным по всем параметрам: живого и неживого вещества: растительности, животному миру, почвенному населению, рельефу, почвообразующей породе, свойствам почвы, глубине и режимам грунтовых вод;
  - ✓ каждому биогеоценозу присуще наличие особого, только ему присущего типа обмена веществ и энергии;
  - ✓ всем компонентам биогеоценоза свойственно единство жизни и ее среды, т.е. особенности и закономерности жизнедеятельности биогеоценоза определяются средой его обитания, таким образом, биогеоценоз представляет собой географическое понятие.

- Каждый конкретный биогеоценоз также должен:
  - быть однородным по своей истории;
  - быть достаточно долговременным сложившимся образованием;
  - ясно отличаться по растительности от соседних биогеоценозов и эти отличия должны быть закономерными и экологически объяснимыми.

Биогеоценоз

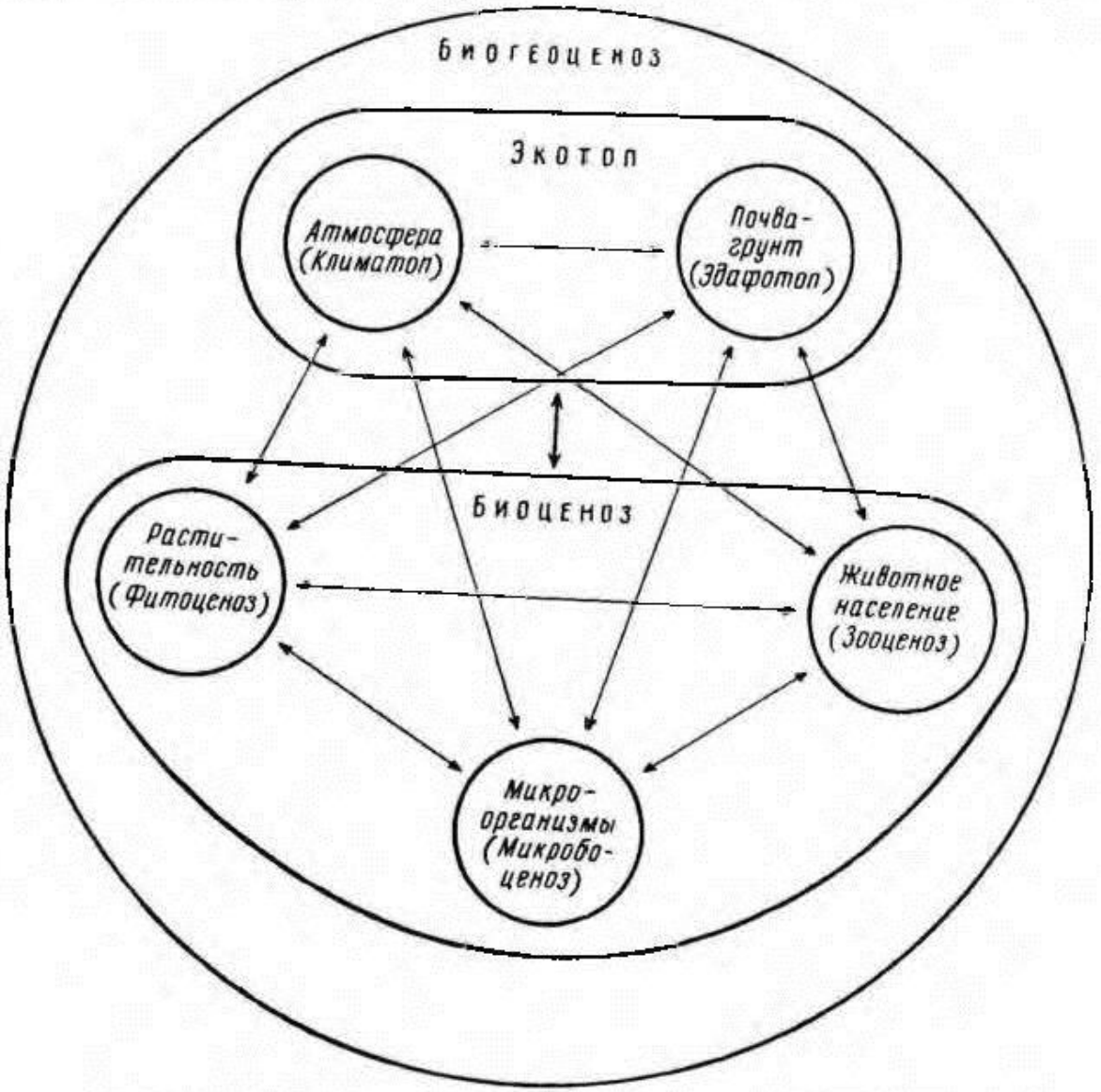


Рис. 1. Схема взаимодействий компонентов биогеоценоза

В.Н. Сукачев и Н. В. Дылис. Основы лесной биогеоценологии, М. : Наука, 1964: стр. 24, рис. 1.

Важность лесных биогеоценозов

- Самые сложные и важные природные биогеоценозы – лесные.
- Лес представляет собой наиболее мощную «пленку жизни». Лесам принадлежит доминирующая роль в сложении растительного покрова Земли. Они покрывают почти третью часть суши планеты – 3,9 млрд. га. Из остальной части почти половину занимают пустыни, полупустыни и тундры около 3,8 млрд. га, а более 1 млрд. га приходится на бросовые, застроенные и другие непродуктивные земли.
- Масса органического вещества, сосредоточенного в лесах, составляет 10^{17}–10^{18} т, что в 5–10 раз превышает массу всей травянистой растительности.
- Именно поэтому особое значение придавалось и придается биогеоценологическим исследованиям лесных систем и термин «биогеоценоз» был предложен академиком В.Н. Сукачевым в конце 30-х гг. 20 в. применительно к лесным экосистемам. Но оно правомерно по отношению к любой природной экосистеме в любом географическом районе Земли.

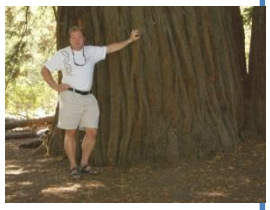
Важность лесных биогеоценозов

- Лес – один из богатейших и важнейших мировых природных ресурсов.
- Леса Российской Федерации (РФ) являются одними из важнейших национальных природных ресурсов и её гордостью. Общая площадь лесов составляет свыше 809 млн. га, или 20-25% от общей площади лесов мира (ФАО: Глобальная оценка лесных ресурсов 2011; http://www.fao.org/docrep/013/i2000e/i2000e.pdf).
- Большую часть из них преимущественно составляют естественные леса.
- Леса России играют исключительно важную роль в сохранении биоразнообразия умеренных и бореальных областей Евразии.
- Леса РФ по преимуществу бореальные: их площадь занимает 88% площади лесопокрытых земель.
- В РФ сосредоточено около 63% (760 млн. га) площади всех бореальных лесов мира (1214 млн. га по оценкам ФАО).
- **Бореальные леса РФ играют важнейшую роль в регуляции климата всей планеты, и их изучение сейчас, в условиях глобального изменения климата, актуально как никогда (катастрофическое сокращение лесов!)**

Важность и уникальность хвойных биогеоценозов

- **Хвойные виды** семейства сосновых (*Pinaceae*: роды, встречающиеся в РФ – *Pinus*, *Picea*, *Abies*, *Larix*) являются **основными бореальными лесообразующими породами РФ**
- **Уникальная репродуктивная система** (гаплоидный мегагаметофит, имеющий гаметическое происхождение и идентичный яйцеклетке гаплотип)

- **Им принадлежат практически почти все биологические рекорды**:
  - продолжительность жизни до 5 тыс. лет (*Pinus longaeva*)
  - размеры особей (высота *Sequoia sempervirens* свыше 110 м и масса *Sequoiadendron giganteum* свыше 1400 т)
  - размер генома (в 4-9 раз больше генома человека)



- **Наиболее экономически выгодные и источник уникальных продуктов для**:
  - строительной, бумажной и мебельной пром-сти
  - получения альтернативной энергии и биотоплива (спирт на основе целлюлозы)
  - медицины (например, противоопухолевые препараты на основе таксола)
  - авиационной и ракетной техники (высококачественное реактивное топливо на основе альфа-пиренов)
  - борьбы с вредителями (экологически безопасные пестициды и репелленты на основе терпенов и др. вторичных метаболитов)

Важность популяционной генетики для изучения лесных биогеоценозов

- Ключевыми (системообразующими) организмами лесных биогеоценозов – являются **древесные виды**, которые состоят из **популяций**, очень часто с высокой **внутрипопуляционной** и **межпопуляционной изменчивостью**
- **Популяцио́нная гене́тика**, или **генетика популяций** позволяет изучать эту изменчивость:
  - раздел генетики, изучающий распределение частот аллелей и их изменение под влиянием движущих сил эволюции: мутагенеза, естественного отбора, дрейфа генов и миграция;
  - изучает и принимает во внимание субпопуляционные структуры, пространственную и временную структуру популяции;
  - пытается объяснить процессы адаптации и видообразования и является одной из основных составляющих синтетической теории эволюции (= синтез эволюционного учения Дарвина и генетики, приведший к формированию математической основы популяционной генетики);
  - широко использует молекулярно-генетические маркёры

Важность изучения популяций

- **Популя́ция** (от лат. *populatio* — население) — это совокупность организмов одного вида, длительное время обитающих на одной территории (занимающих определённый ареал). Этот термин используется в различных разделах биологии, экологии, демографии, медицине и психометрике.
- **В экологии и эволюционной теории популяция** — совокупность особей одного вида, обладающая общим генофондом, способная к более-менее устойчивому самовоспроизводству (как половому, посредством панмиксии в идеальном случае, так и бесполому), относительно обособленная (географически или репродуктивно) от других групп, с представителями которых (при половой репродукции) потенциально возможен генетический обмен.
- **С точки зрения популяционной генетики, популяция** — это группа особей, в пределах которой вероятность скрещивания во много раз превосходит вероятность скрещивания с представителями других подобных групп. Обычно говорят о популяциях как о группах в составе вида или подвида.
- **Популяция** — элементарная единица эволюции.
- Термин **популяция** введён Вильгельм Иогансеном в 1903 г., однако уже Чарлз Дарвин объяснял эволюцию видов изменчивостью и конкуренцией групп особей.

Важно понимать основы генетики

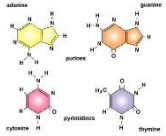
- **Гене́тика** (от греч. γενητως — порождающий, происходящий от кого-то) — наука о закономерностях наследственности и изменчивости.

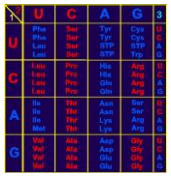
- **В основе всей наблюдаемой нами фенотипической изм**енчивости (морфологической, физиологической и биохимической) живых организмов в биогеоценозе лежит **изменчивость на уровне ДНК** (носителя генетической информации) и **влияние среды во взаимодействии с генотипом**.

- Чтобы понять фенотипическую изменчивость и как она влияет на эволюцию организмов и их адаптацию к определенной среде надо знать связь между этой фенотипической изменчивостью и генетической.

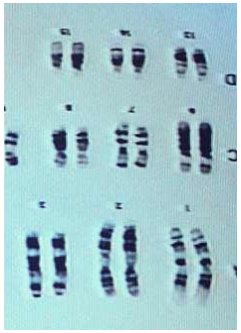
- Большинство организмов в биогеоценозе представлены в той или иной степени генетически различающимися популяциями; т.е. генетическая изменчивость вида определённым образом организована и структурирована в пространсте, образуя популяционно-генетическую структуру вида.

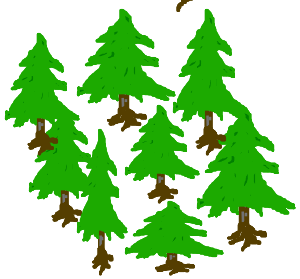
Аналогия между генетической и языковой передачей информации





Нуклеотид	Буква алфавита
↓	
Триплет	Слово
↓	
Ген	Строчка в тексте/Предложение
↓	
Хромосома	Глава
↓	
Геном	
↓	
Генофонд	



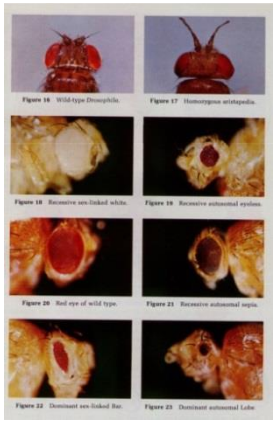


Важность использования генетических маркёров

Вся история генетики связана с развитием и использованием генетических маркёров:

- начиная с **Грегора Менделя** (Gregor Mendel, 1822-1884) - морфологические маркёры семян гороха (цвет, форма) – законы наследования моногенных признаков

- и **Томаса Моргана** (Thomas Morgan, 1866-1945) - морфологические маркёры дрозофилы – генетическое сцепление, генетические карты, хромосомная теория и т.д.



- и заканчивая функциональными маркёрами (ESTP, SNPs и т.д.)

Важность молекулярно-генетических маркёров

Молекулярно-генетические маркёр или ДНК-маркёр:

- признак, выявляемый методами молекулярной биологии и отражающий изменчивость на уровне нуклеотидной последовательности ДНК для определенного гена или для любого другого участка хромосомы при сравнении различных генотипов особей, пород, сортов, линий, популяций;
- изучаются на уровне белков, ДНК и РНК;
- эффективно используются для решения многих задач генетики, селекции, сохранения биоразнообразия, изучения механизмов эволюции, картирования хромосом, а также для семеноводства и племенного дела;
- условно можно подразделить на следующие три типа:
  - «селективно-нейтральные» маркеры некодирующих участков
  - «функциональные» маркеры участков структурных генов, кодирующих аминокислотные последовательности белков (электрофоретические варианты белков) и регуляторных генов
  - «анонимные» маркеры участков различных последовательностей ДНК, отношение которых к структурным генам, как правило, неизвестно.

Задачи, решаемые с помощью генетических маркёров

- наследование и картирование генетических признаков
- построение генетических и физических карт
- популяционно-генетическая структура, уровни изменчивости и генетическая дифференциация
- система скрещивания
- поток генов, обмен генами, интрогрессивная гибридизация
- судебно-медицинская экспертиза, генетическая идентификация
- молекулярная филогения и систематика
- **ассоциация генетической изменчивости с фенотипической**
- Marker-aided selection и **геномная селекция**

Из всей изменчивости нас большего всего интересует адаптивная генетическая изменчивость - Почему?

- понять и использовать адаптивный потенциал
- предсказать эффект изменения климата
- противостоять этим эффектам путём селекции более устойчивых деревьеа и для целенаправленного и обоснованного переноса и интродукции растительного материала (assisted migration)
- понять **эволюционный ответ** и молекулярные механизмы генетической адаптации

Важность геномики

- **эволюционный ответ** это **генетическая адаптация** через генетическое изменение, которое увеличивает приспособленность растений и животных к их среде обитания, включая биотические и абиотические взаимодействия

- **множественные гены** обычно вовлечены в **генетическую адаптацию**, таким образом их изучение требует применения **геномных методов** и **полногеномных подходов**

Как измерить адаптивную генетическую изменчивость в популяциях?

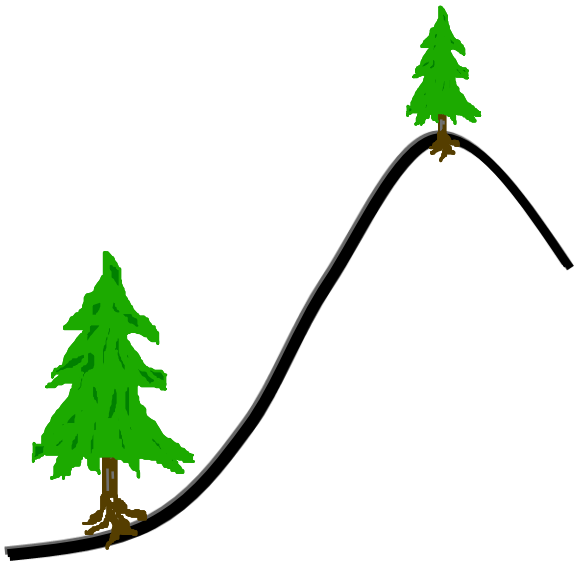
- **Традиционные методы**
  - географические культуры
  - молекулярно-генетические маркеры
  - картирование генов количественных признаков, или Quantitative Trait Locus (QTL)-картирование
- **Современные популяционно-геномные подходы**
  - новый тип **функциональных геномных маркеров**
  - **ассоциативное картирование**
  - обнаружение генов или генетических элементов с необычно высокими или низкими уровнями изменчивости и дифференциации, выпадающими из общего распределения генетической изменчивости и дифференциации (так называемые **гены-аутсайдеры** или **outliers**)
- **Интеграционный подход**

Nature vs. Nurture

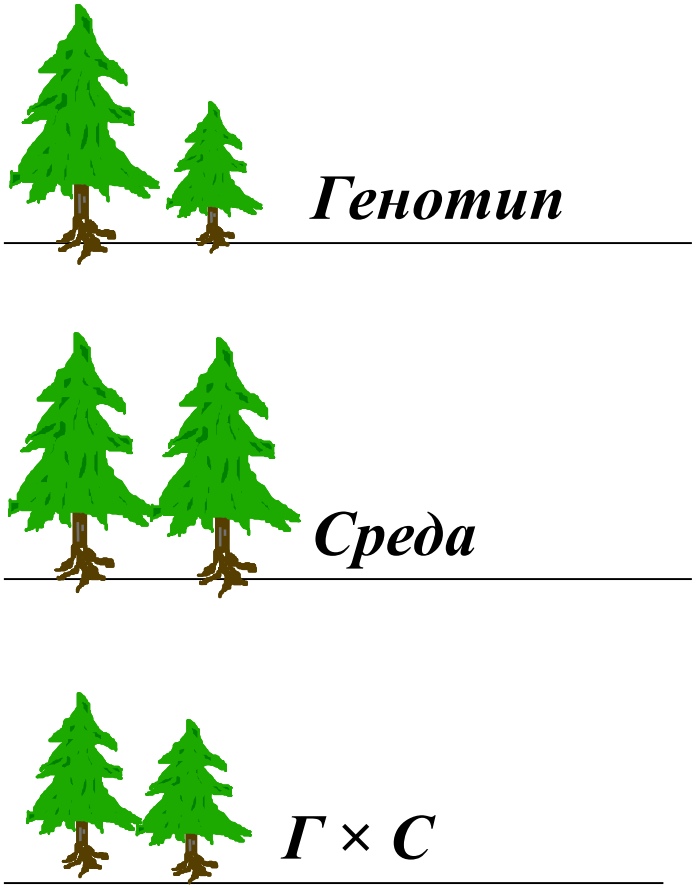
Как их разделить?

Пример:

Высотный градиент



Географическая культура



Фундаментальная проблема генетики – связь фенотипа с генотипом!

Её можно решить для сложных признаков только изучая изменчивость по совокупности всех генов в геноме

Nature (genome) vs. *Nurture* (environment)

$$P = G + E + G \times E$$

Phenotype = *Genome(Genotype)* + *Environment* + *Interaction*

Organisms are different because of the:

- genomic/genetic (*G*) differences among individuals
- different environments (*E*) where individuals are growing
- and interactions between the genotypes and the environments in which they grow (*G*×*E*)

Nature (genome) vs. *Nurture* (environment)

Сколько генотипа в фенотипе?

Mendelian traits vs. Complex traits

- **Mendelian = Qualitative**
  - **single gene responsible for most of the observed phenotypic variance**
- **Complex = Quantitative**
  - **multiple genes with gene × gene, gene × environment interactions contributing to phenotypic variance**

Single vs. Multiple Genes in Population

$$P^n = G^n + E^n + G^n \times E^n$$

n – multiple phenotypes, genes and environments

Great Ferma Theorem : $Z^n = X^n + Y^n$

does not have integer solutions *X*, *Y*, *Z* for $n > 2$

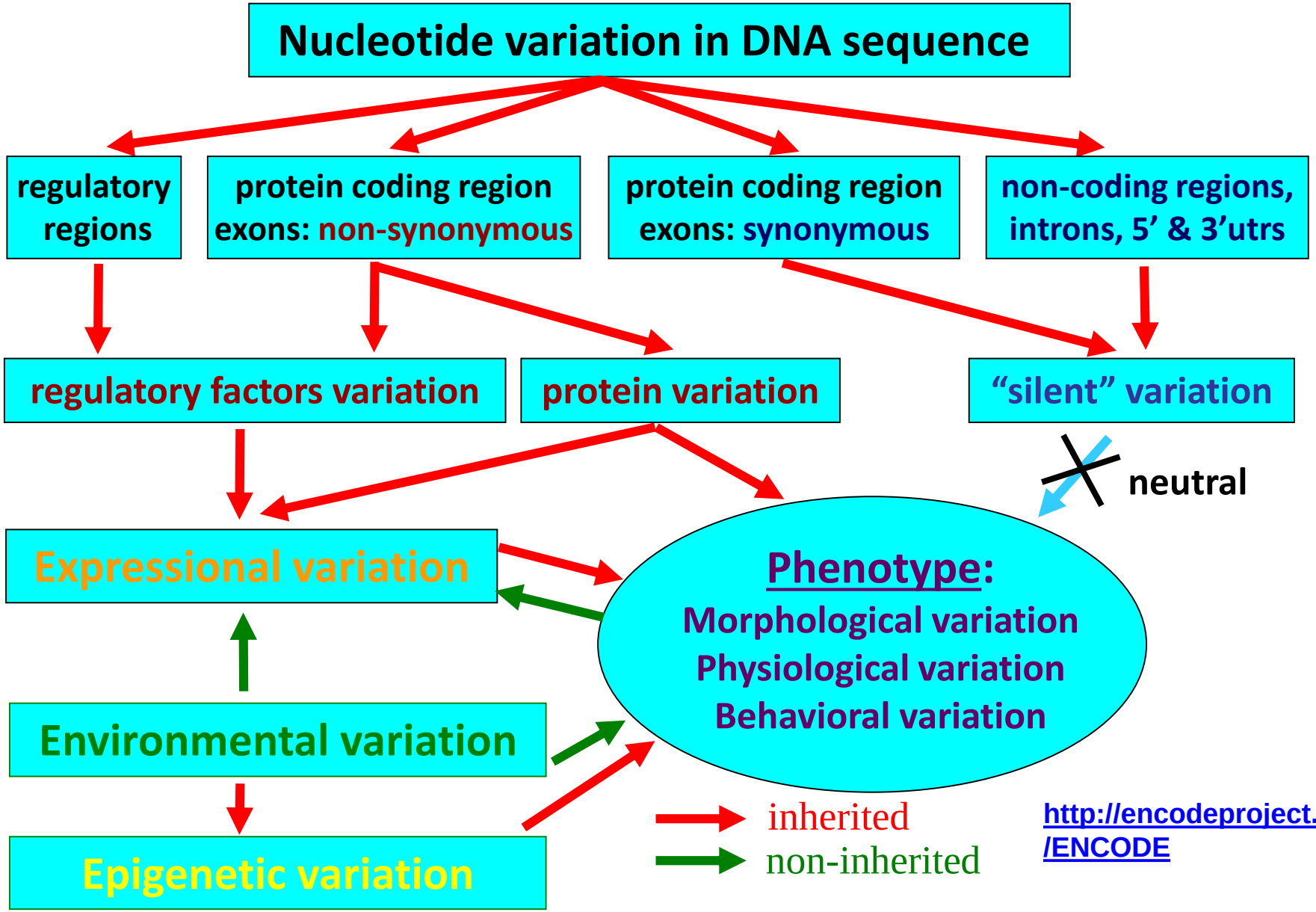
Andrew Wiles, 1994

My Theorem : $P^z = G^x + E^y$

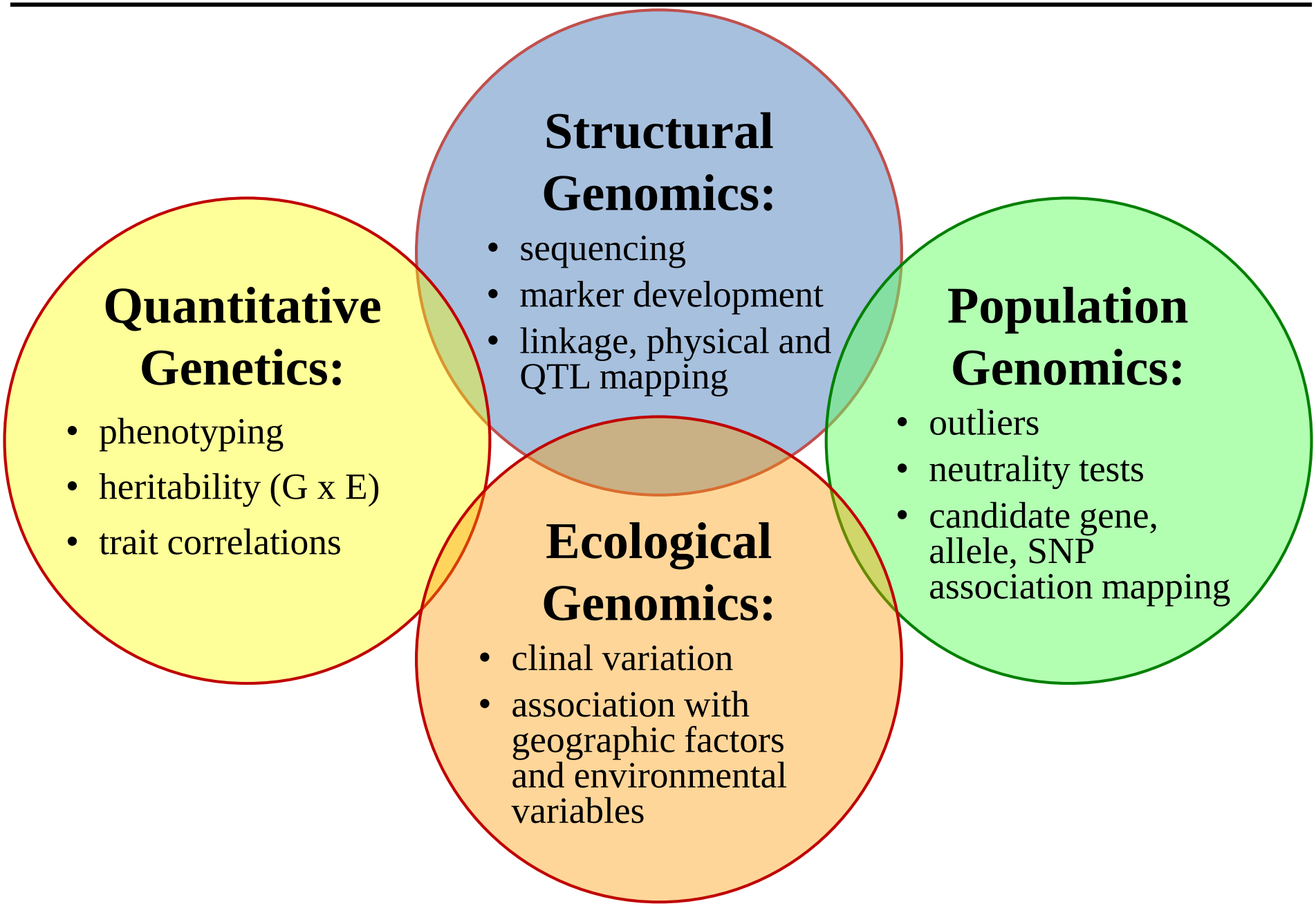
My Great Theorem: $P^z = G^x + E^y + G^x \times E^y$

Геномика – единственное решение!

Expression of genetic variation



Linking Genotype to Phenotype & Environment



Как связать сложную фенотипическую изменчивость с генетической?

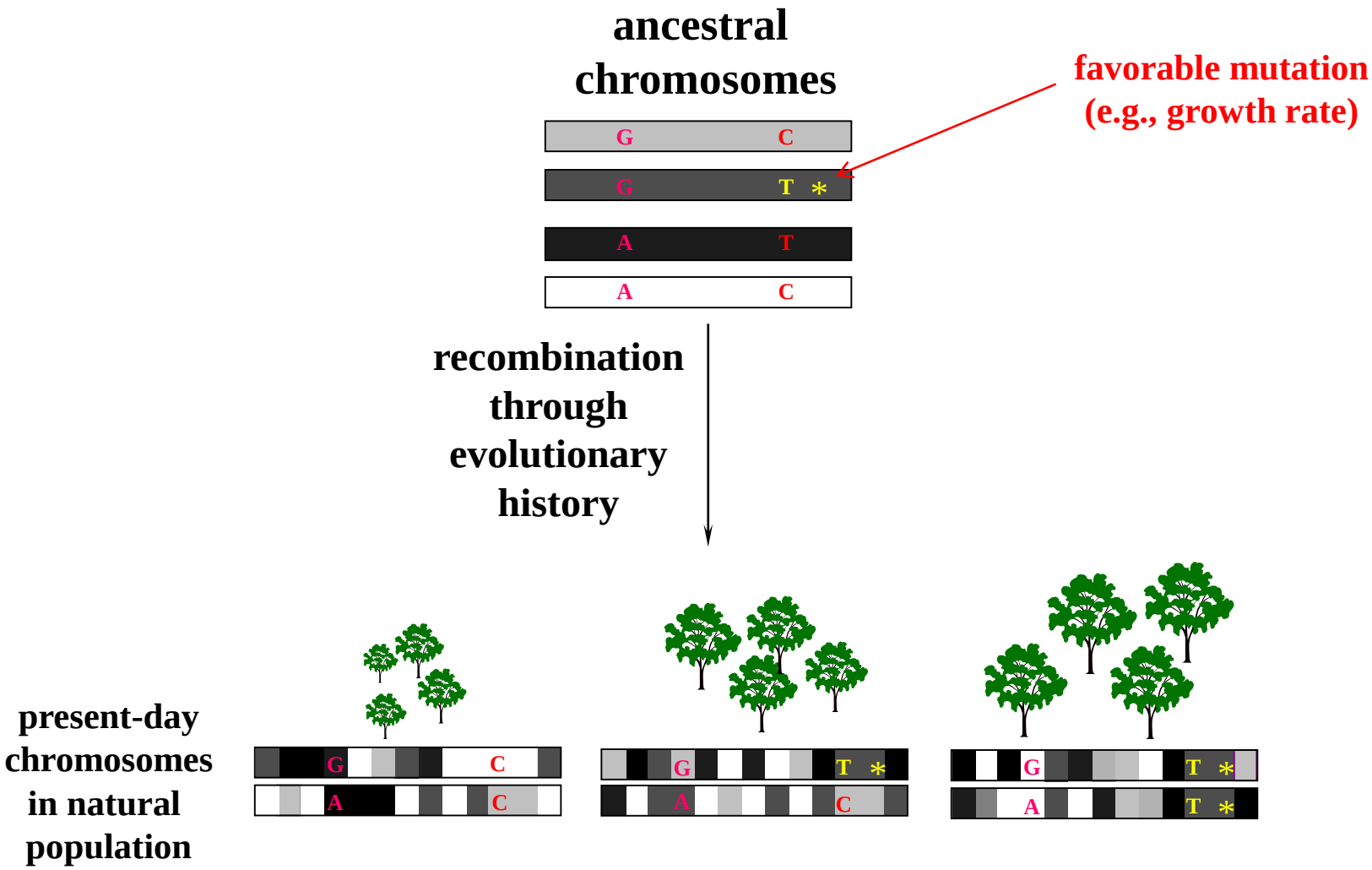
Современные подходы в популяционной, ландшафтной и экологической геномике:

- изучение связи генетической изменчивости по большому числу маркёров с фенотипической и средовой изменчивостью (т.н., **ассоциативное картирование**)
- обнаружение генов-«аутсайдеров» (outliers) путём сравнение параметров изменчивости и дифференциации по селективно-нейтральным и функциональным геномным маркерам

Крутовский К.В. 2006. От популяционной генетики к популяционной геномике лесных древесных видов: интегрированный популяционно-геномный подход. *Генетика* Т. 42. №10. С. 1304–1318.

González-Martínez S.C., Krutovsky K.V., Neale D.B. 2006. Forest tree population genomics and adaptive evolution. *New Phytologist* **170**(2): 227-238.

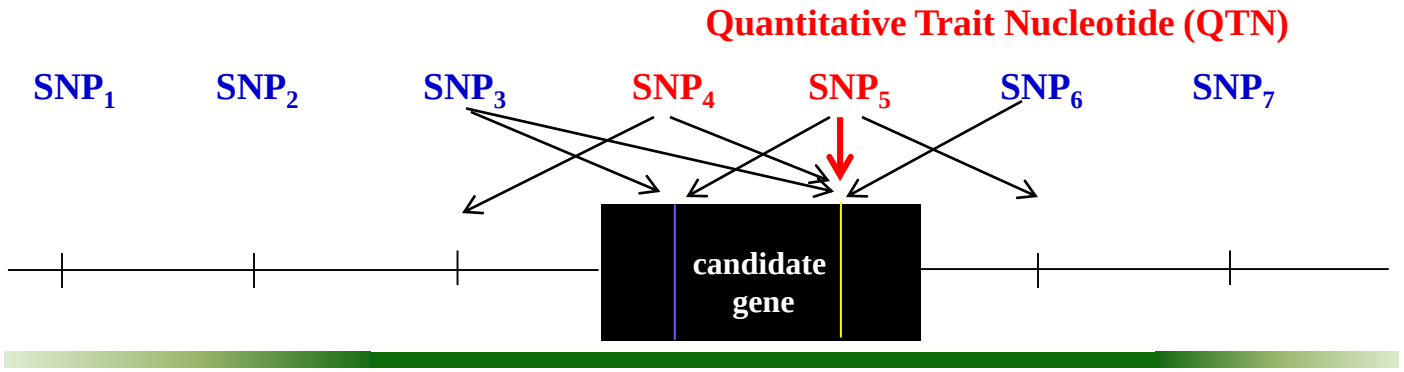
Ассоциативное картирование



Полно-геномное ассоциативное картирование с использованием случайных маркёров (например, «снипов» - SNPs – single nucleotide polymorphisms)

vs.

Избирательного ассоциативного картирования, основанного на фунциональных маркёрах в генах-кандидатах



Неравновесие по сцеплению

Неравновесие по сцеплению (Linkage Disequilibrium - LD) – это неслучайная ассоциация аллелей сцепленных локусов

Association Mapping Components

- **Phenotypes**
  - trait values
- **Numerous Molecular Markers**
  - SNPs
    - ✓ SNP genotyping assays based on preselected SNPs
    - ✓ SNP genotyping by sequencing
- **Statistical Models**
  - Linear model: phenotype as response and genotype as predictor

Современные методы получения маркёров и генотипирования путём прямого секвенирования

bar-coded DNA or mRNA library pools

individual genomic or mRNA (tissue-specific or total)

next generation high-throughput massively parallel DNA sequencing (NGS)

1) RAD complexity reduced, or 2) target-enriched genomic DNA bar-coded pools

individual genomic DNA

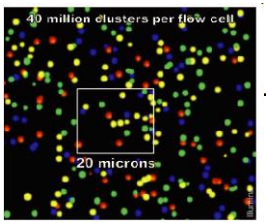
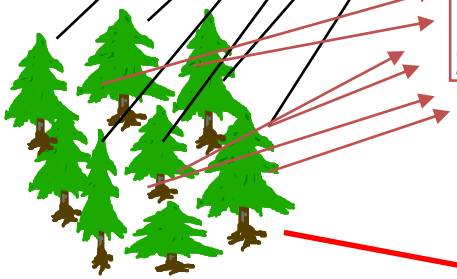
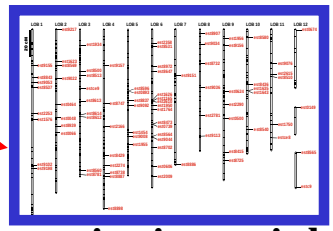


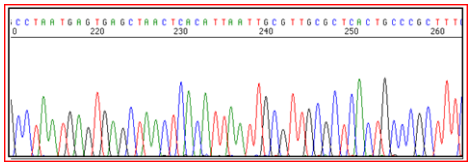
image analysis





genotyping and phenotyping in mapping or natural population

quantitative trait loci (QTL), candidate gene or association mapping



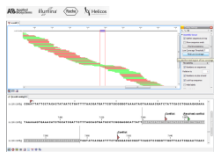
DNA chromatograms

genotyping by sequencing (GBS)

high-throughput SNP genotyping

high-density SNP marker development

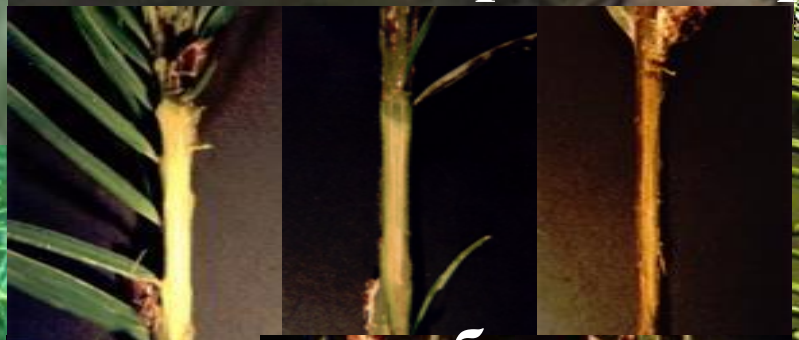
sequence processing and analysis

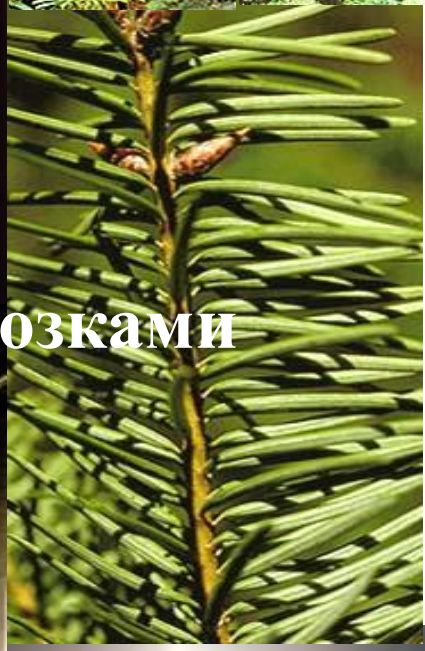


Геномика адаптивных признаков хвойных

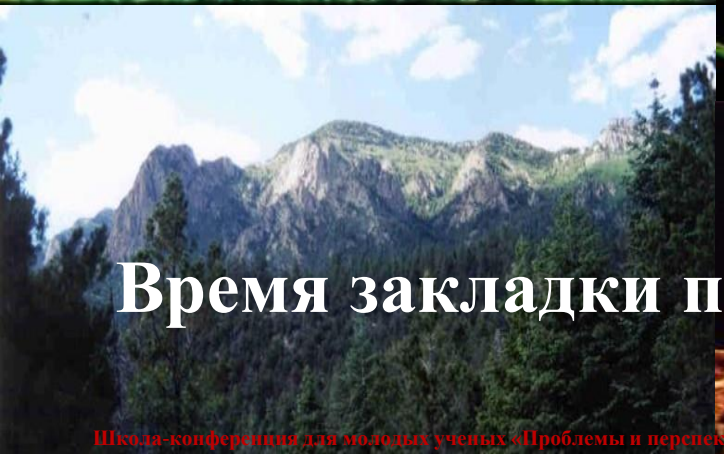
Время раскрытия почки

Вторичное прорастание

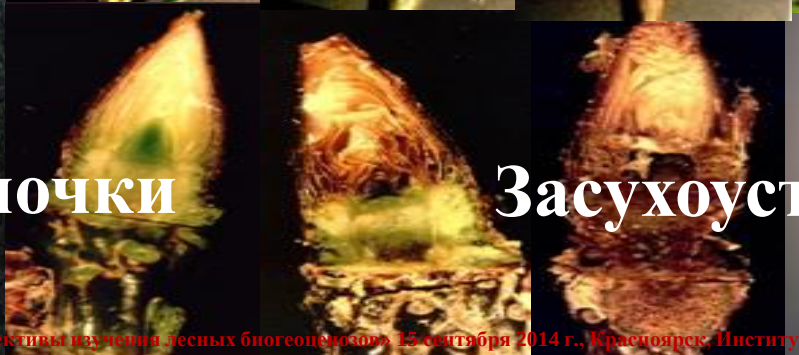




Повреждение почек, хвои и стебля заморозками



Время закладки почки



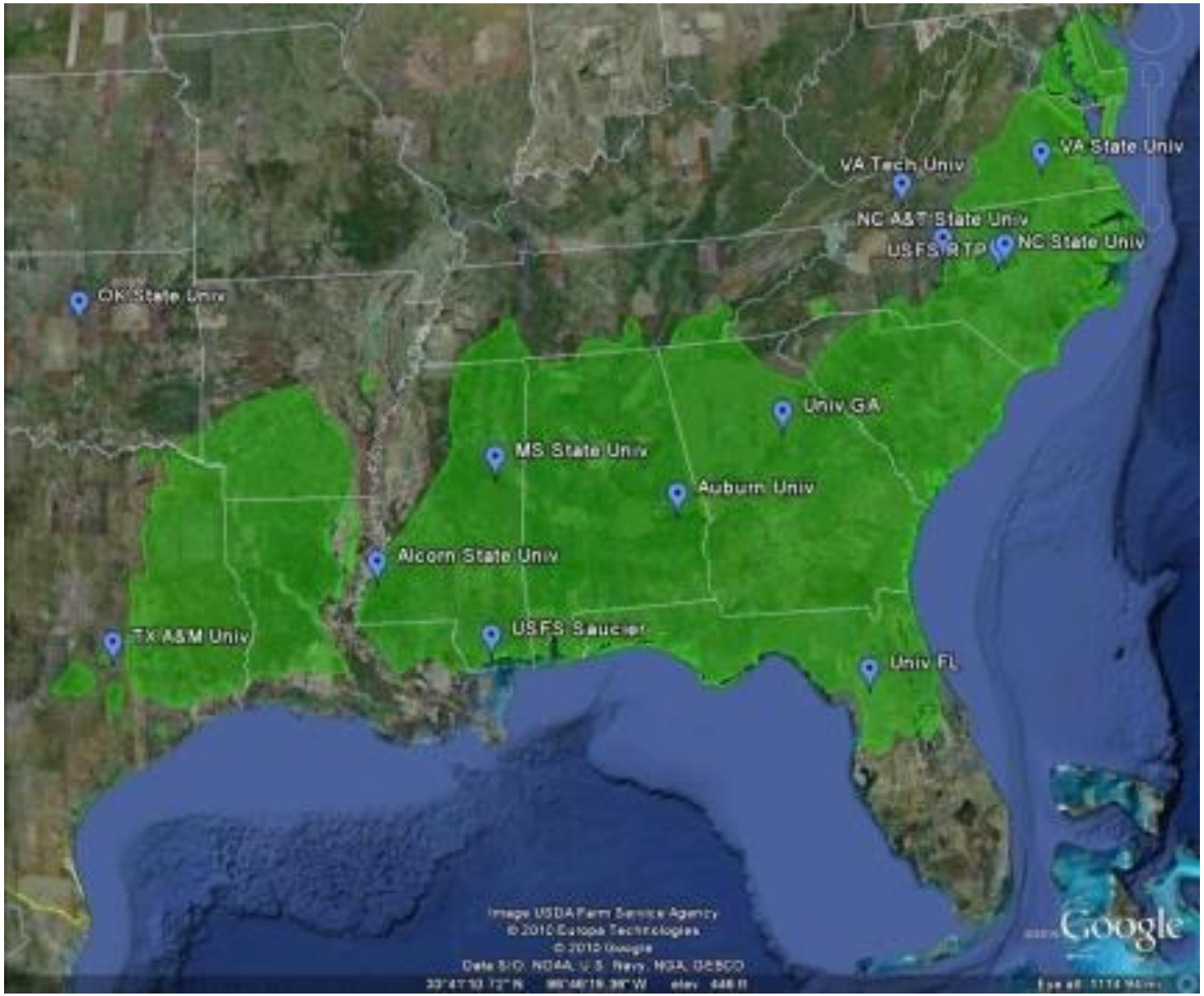
Засухоустойчивость

Проекты по популяционной, ландшафтной и экологической геномике ладанной сосны (*Pinus taeda* - Loblolly pine)

- USDA NRICGP Plant Genome Program / National Institute of Food and Agriculture (NIFA) Agriculture and Food Research Initiative (AFRI) Competitive Grants Program, Applied Plant Genetics Coordinated Agricultural Project (CAP), #CA-D-PLS-2038-CG, PI: D. B. Neale (dendrome.ucdavis.edu/ctgn/people), 2004-2011, $6,000,000; **“Conifer Translational Genomics Network” (CTGN)**

- USDA NIFA AFRI Competitive Grants Program, CAP, Climate Change Program 1: Regional Approaches to Climate Change, Program Area Code – A3101, #2011-68002-30185, PI: Timothy Martin (www.pinemap.org/about/team-members), 5 years, 2011-2016, $19,976,825; **“Integrating research, education and extension for enhancing southern pine climate change mitigation and adaptation” (PINEMAP)**

Ареал ладанной сосны



Conifer Translational Genomics Network (CTGN)

- CTGN много институтский Координируемый с.-х. проект (Coordinated Agricultural Project - CAP) финансируемый USDA и US Forest Service

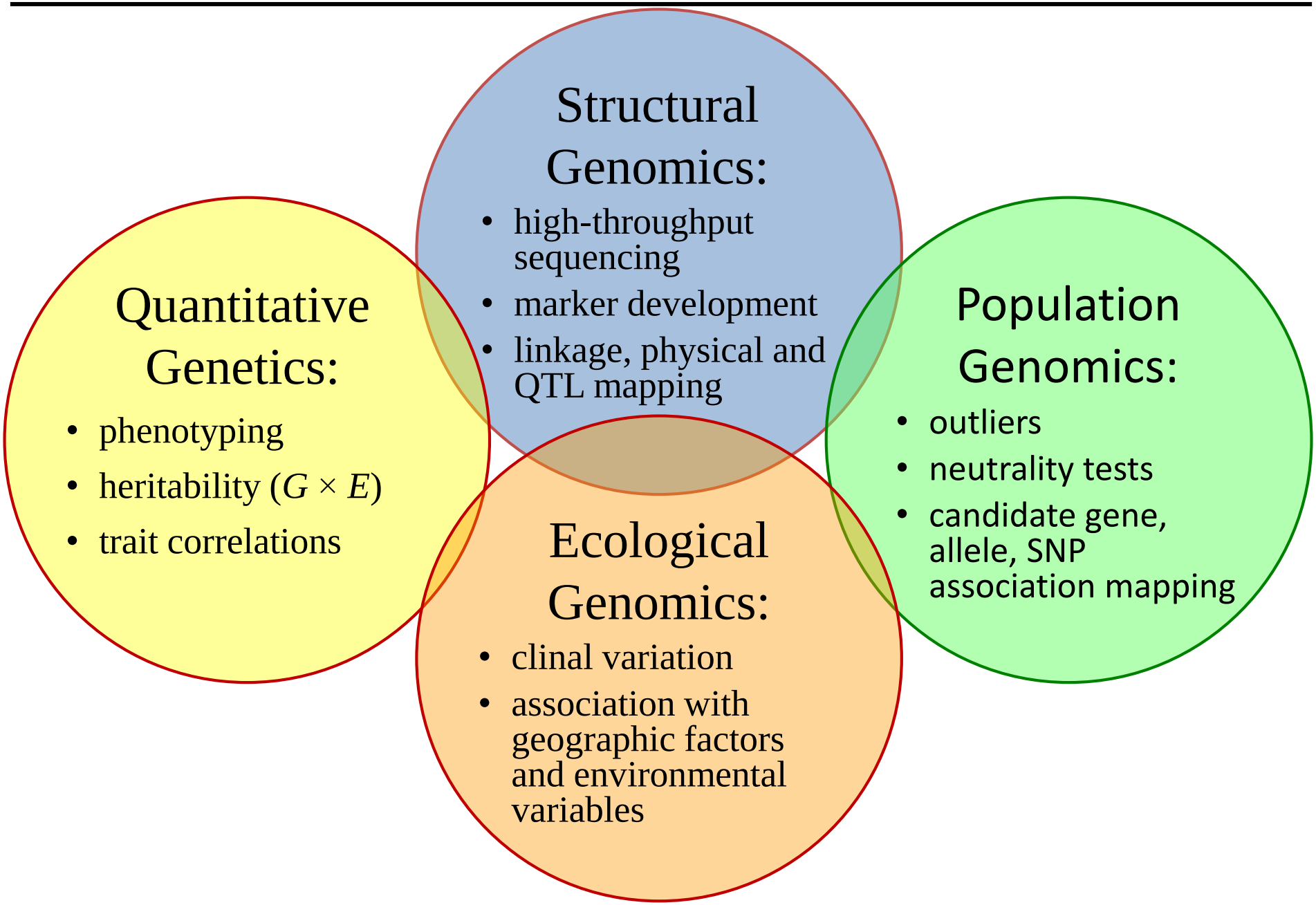
USDA United States Department of Agriculture
National Institute of Food and Agriculture

- Основная задача: применение экспериментальных геномных ресурсов и средств в практической лесной селекции

www.pinegenome.org/ctgn

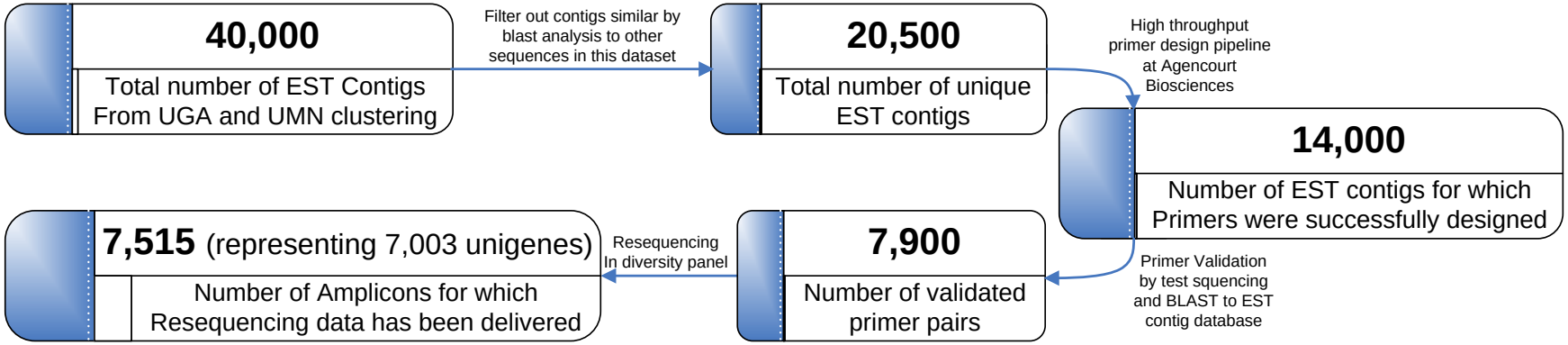
- Подходы:
  - масштабное генотипирование элитных селекционных популяций ладанной сосны и дуглассии
  - обнаружение и подтверждение ассоциаций между генетическим маркёрами и фенотипическими признаками
  - моделирование, разработка и внедрение оптимальных методов включения маркёров в селекционные программы геномной селекции

CTGN: Linking Genotype to Phenotype & Environment



Секвенирование и генотипирование SNP

NSF Allele Discovery of Economic Pine Traits 2 (ADEPT2) resequencing & SNP discovery project, $6 mln



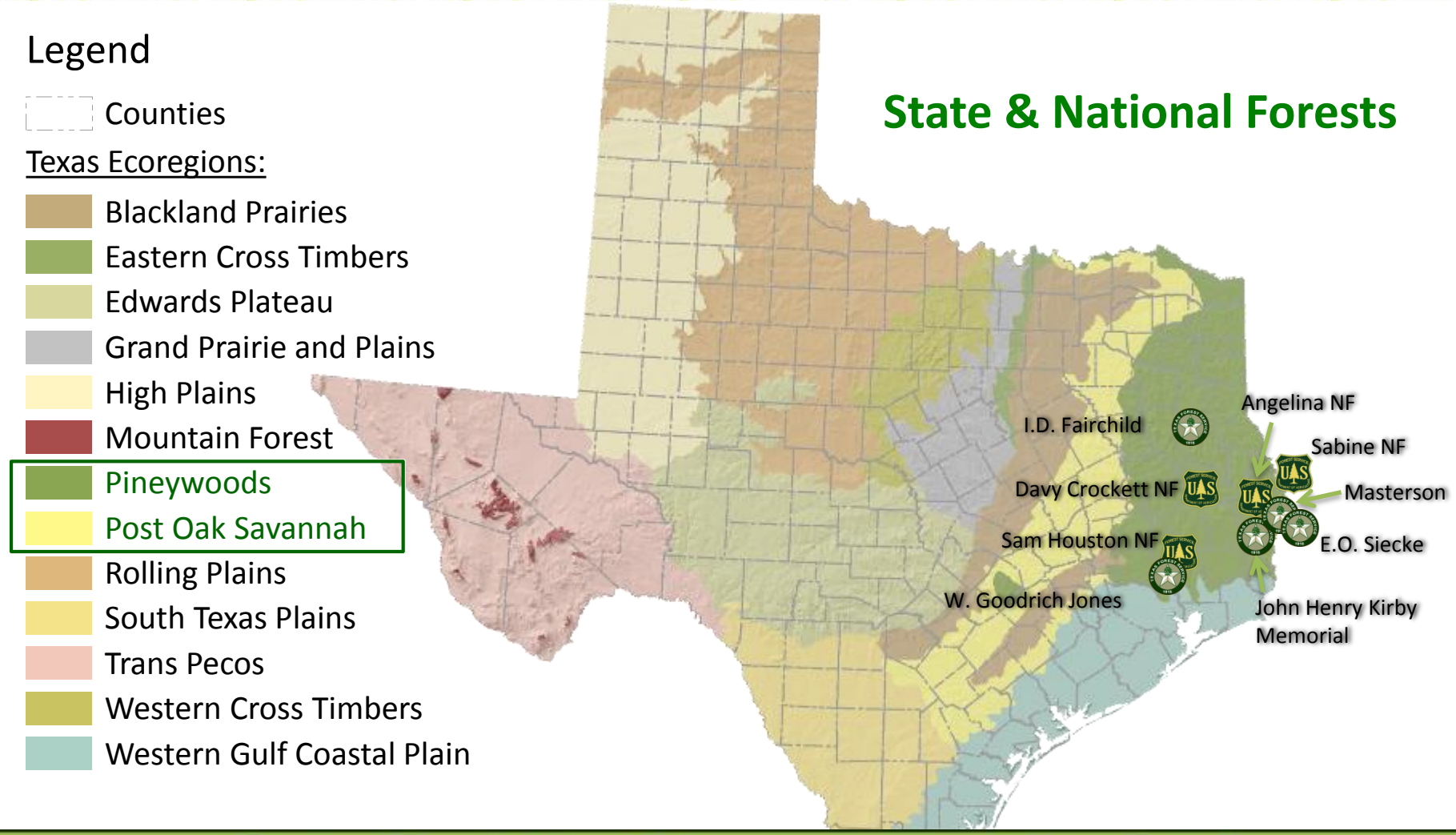
- ~23,000 SNPs обнаружено в ~7,900 частично амплифицированных и секвенированных уникальных генах в небольшой популяционной выборке и ~7000 из них были использованы для создания Illumina Infinium платформы для последующего генотипирования в разных CTGN популяциях

USDA Conifer Translational Genomics Network (CTGN) project , $6 mln

5,379 SNPs были генотипированы в >4500 trees из нескольких селекционных популяций для ассоциативного картирования, используя Illumina Infinium платформу;

4,264 SNPs были полиморфны в популяциях Восточного Техаса

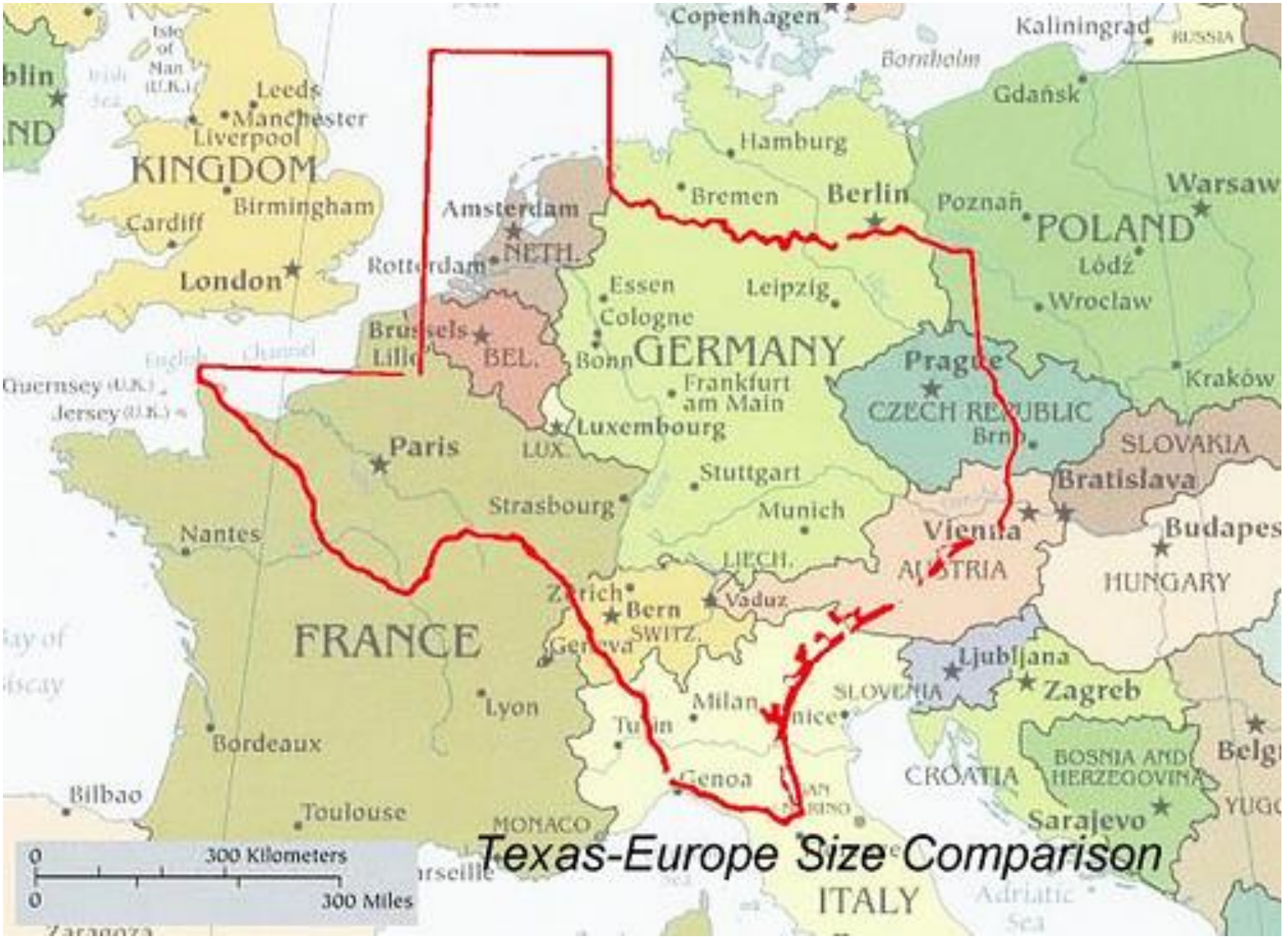
Экозоны Техаса



Texas has **60 million acres** of forestland
- more than any other state in the lower 48 United States

http://txforestservice.tamu.edu/main/popup.aspx?id=7002

EVERYTHING IS BIGGER AND BETTER IN TEXAS — EVEN THE FORESTS


Copenhagen
Kaliningrad
RUSSIA
Bornholm
Gdańsk
Leeds
Manchester
Liverpool
KINGDOM
Birmingham
Cardiff
London
Hamburg
Bremen
Berlin
Poznań
Warsaw
POLAND
Łódź
Wrocław
Amsterdam
NETH.
Rotterdam
Essen
Cologne
Leipzig
Brussels
Lille
BEL.
Bonn
GERMANY
Frankfurt am Main
Prague
CZECH REPUBLIC
Kraków
Luxembourg
LUX.
Paris
Stuttgart
Strasbourg
Munich
SLOVAKIA
Bratislava
Vienna
AUSTRIA
Budapest
HUNGARY
Nantes
LIECH.
Zurich
Bern
Vaduz
SWITZ.
FRANCE
Geneva
Ljubljana
Zagreb
SLOVENIA
Lyon
Milan
Venice
Turin
Bordeaux
Genoa
CROATIA
BOSNIA AND HERZEGOVINA
Sarajevo
Bilbao
Toulouse
MONACO
ITALY
Adriatic Sea
0
300 Kilometers
0
300 Miles
Texas-Europe Size Comparison

Популяции ладанной сосны в Восточном Техасе и изученные параметры

14 populations of the 1st generation selection

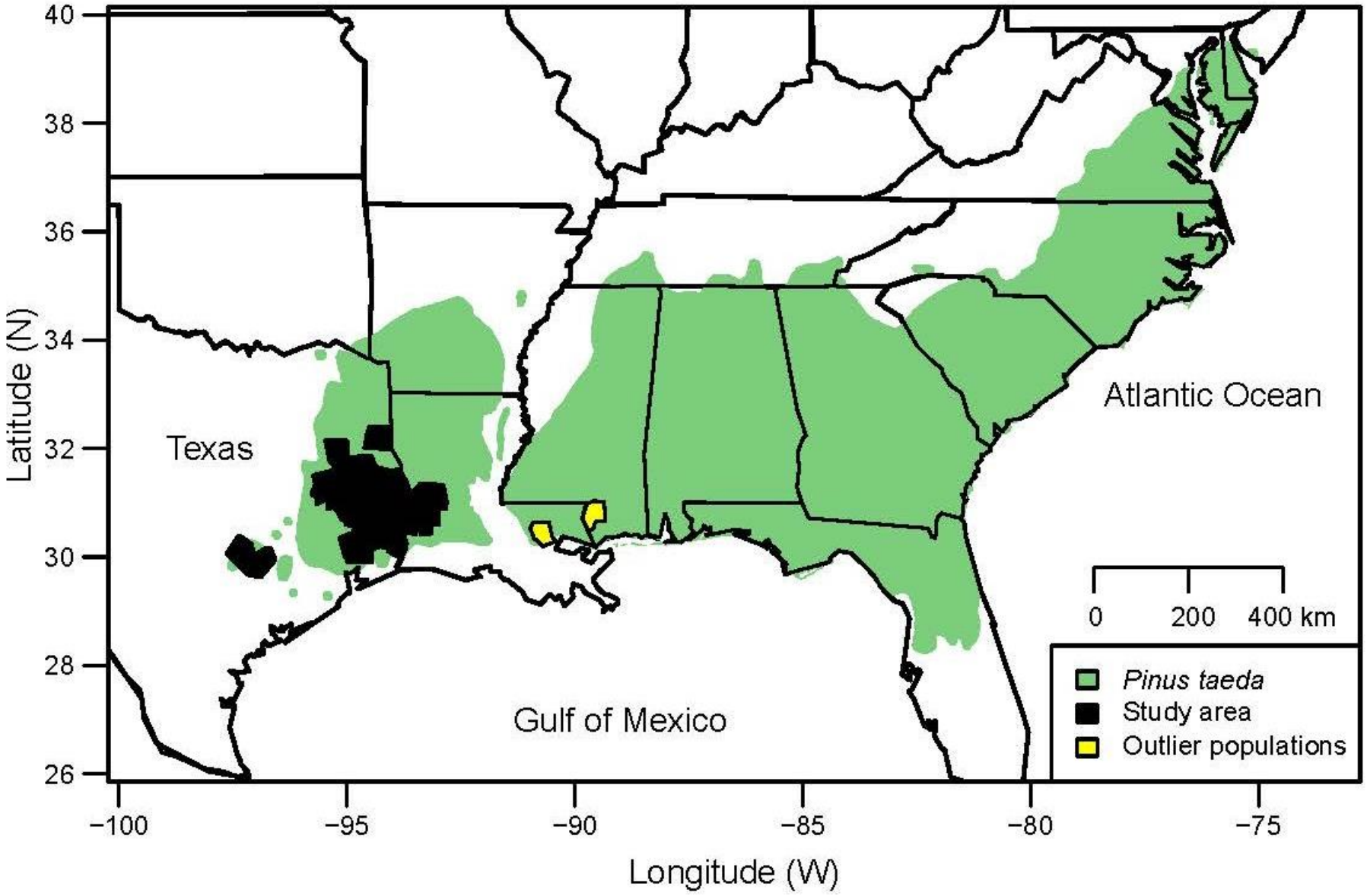
8 populations of the 2nd generation selection

30 breeding groups

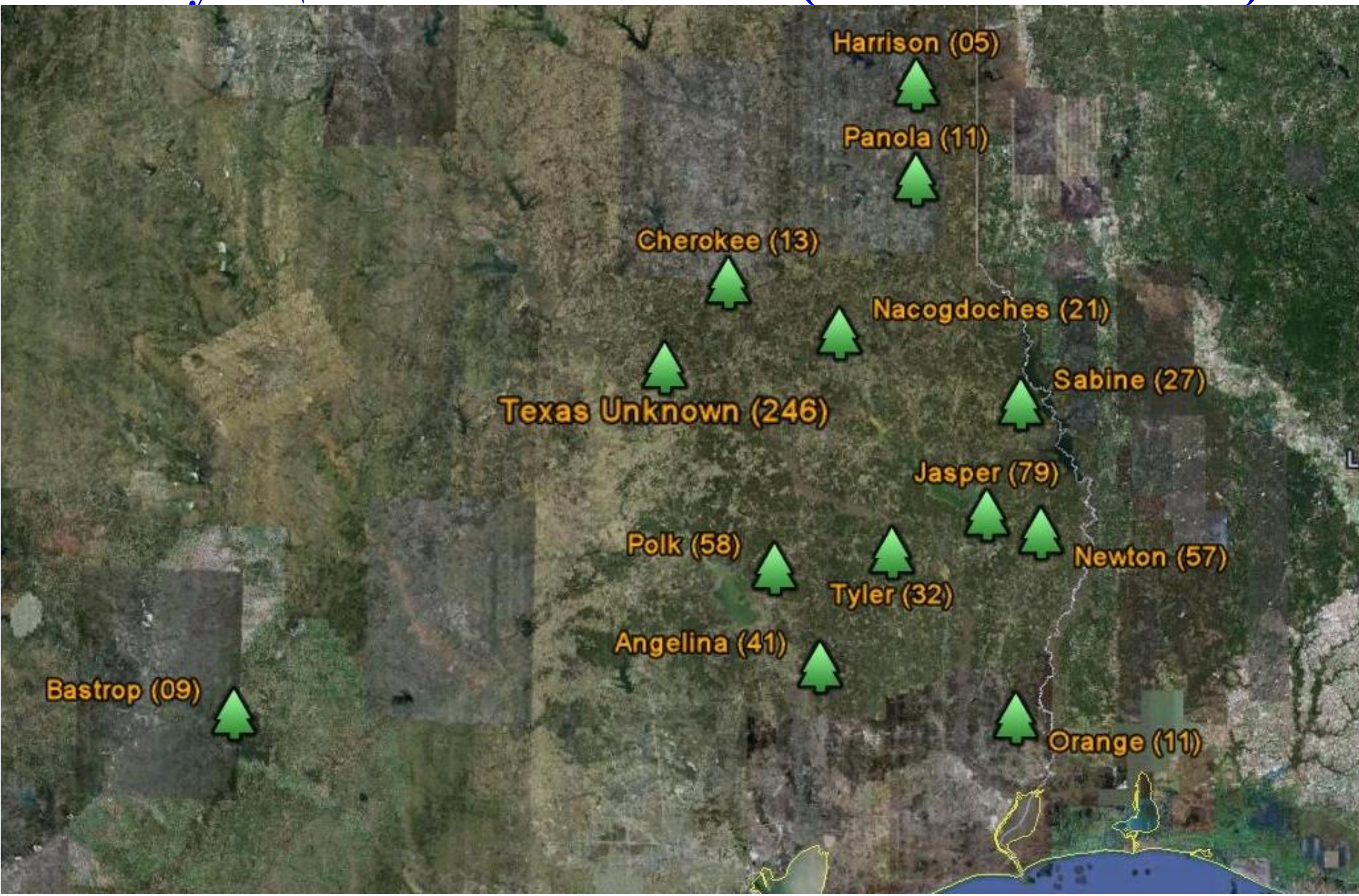
4,264 SNPs / >1,700 trees + Phenotypes: *DBH, height, volume, etc.* Environmental variables: *t, prec., arid.*

- Genetic diversity: H_e, F_{IS}, etc.
- Genetic differentiation index: F_{ST}
- Selection detection via search for F_{ST} outliers
- Population structure inference using STRUCTURE & ΔK
- Association mapping & functional annotation of significant SNPs
- Genome wide LD

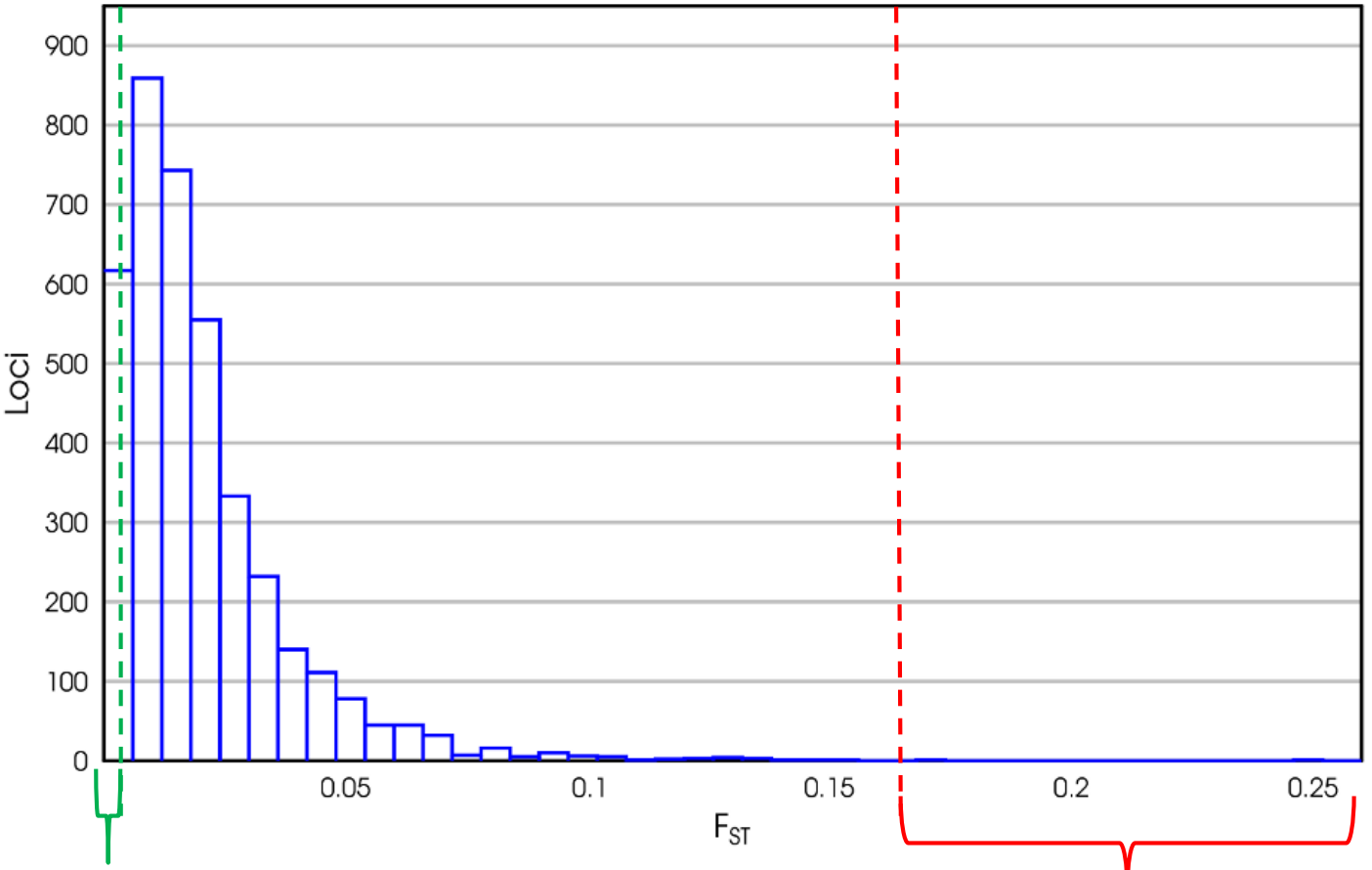
Изученные популяции ладанной сосны в Восточном Техасе



Географическое происхождение селекционных популяций 1-ого поколения (Восточный Техас)



Генетическая дифференциация (F_{ST}) 4,264 SNPs в популяциях 1-ого поколения селекции

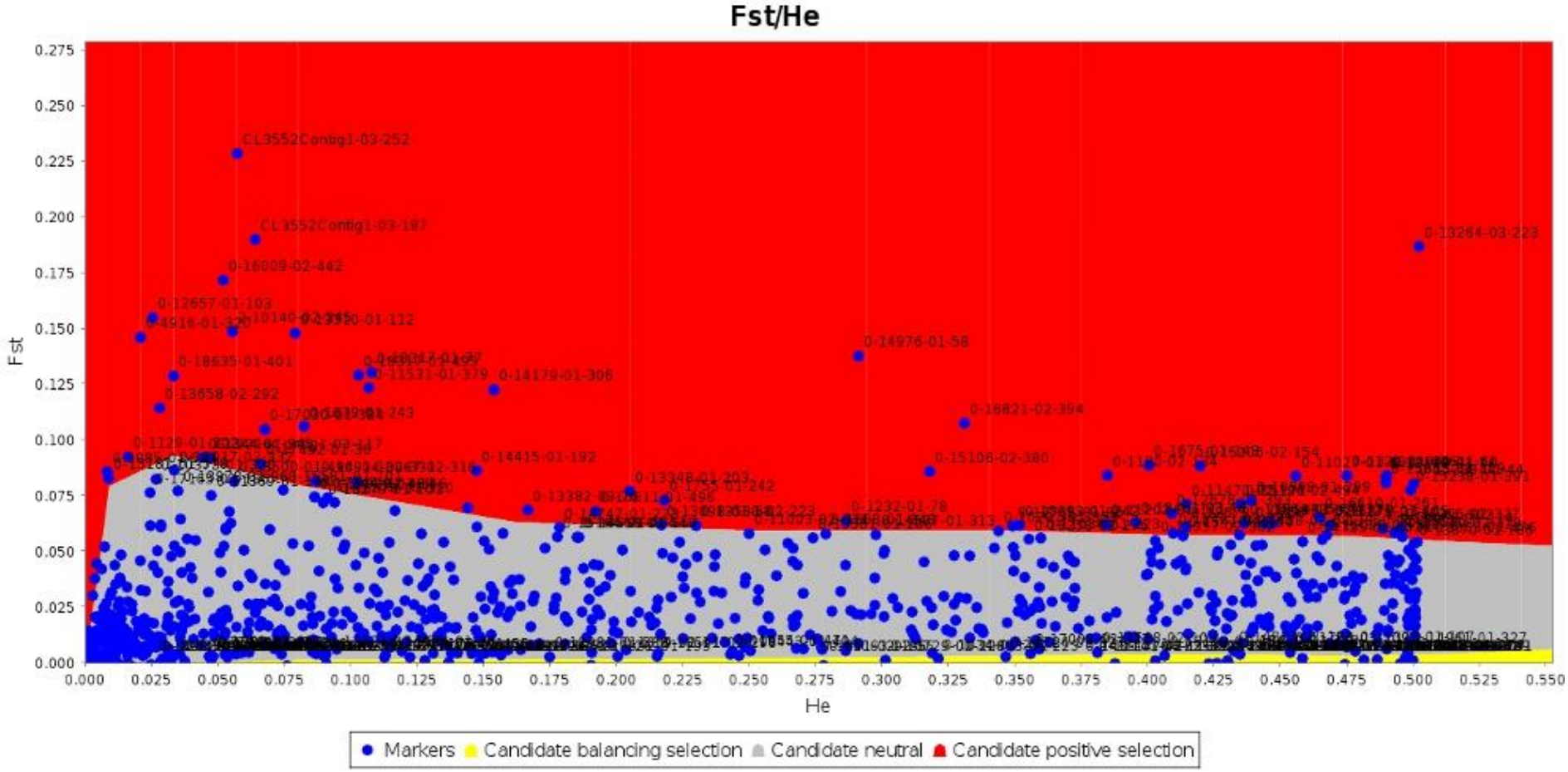


Очень низкая дифференциация; балансирующий или стабилизирующий отбор?

Очень высокая дифференциация; позитивный или дивергентный отбор?

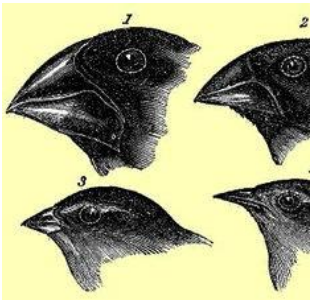
Поиск «outlier» SNPs, находящихся под действием отбора и селективно-нейтральных маркёров с помощью коалесцентного моделирования

Распределение F_{ST} и гетерозиготности (H_e) для 4,264 SNPs в популяциях 1-ого поколения селекции, 95K циклов моделирования для обнаружения порога



Beaumont *et al.* (1996) Royal Soc. B; Antao *et al.* (2008) Bioinformatics

F_{ST} «outliers»



Примеры генов-кандидатов, изменчивость которых вероятно находится под влиянием отбора (after correction for false positive):

Балансирующий отбор (244):

Arabinofuranodisase, Glycoprotease protein, Dehydrin, Lypoxygenase, Cytokinin oxidase, Transmembrane transporters, Endoglucanase, MYB transcription factor, Pinus taeda Heme oxygenase I, etc.

Dirigent protein, Homeobox-leucine zipper, Cytochrome p450, Gras transcription factor, Gigantia protein, Ethylene-forming enzymes Histone H4, Reductases, etc.

Направленный отбор (74):

Geranyldiphosphate, Disease resistance proteins, Arabinogalactan-like proteins, Pinus Expansin, Pinus alpha-Xyloxidase, etc.

Potassium/proton antiporter, Laccase 90DProtein kinases, Histone H3.2, etc.

Chhatre V, Byram T, Neale DB, Wegrzyn JL, Krutovsky KV (2013) Genetic structure and association mapping of adaptive and selective traits in the East Texas loblolly pine (*Pinus taeda* L.) breeding populations. *Tree Genetics and Genomes* 9: 1161-1178

Ассоциативное картирование SNPs и фенотипической изменчивости в адаптивных и селективных признаках (таких как скорость роста, плотность древесины, устойчивость к заболеваниям и засухе, и т.д.)

Значимые ассоциации были обнаружены, например, для:

- Arabinofuranosidase
- Xylosidase
- Protein kinases
- Chloroplast proteins
- Metallothionein-like protein
- Chlorophyll binding protein
- Glucuronase 4-epimerate
- Clavata-like receptor
- RNA polymerases
- Decarboxylases
- Sodium simporter family protein
- Acyl CoA synthetase
- Tubulin beta-chains
- NBS disease resist. protein - P. taeda Transmembrane protein
- Universal stress protein
- Cyclin-D like protein
- cdc2 protein kinases
- Synaptotagmins etc.

Chhatre *et al.* 2013 *Tree Genetics and Genomes* 9: 1161-1178

Клинальная изменчивость и ассоциация с изменчивостью средовых факторов: Логистическая регрессия (LR)

Number of SNPs genotyped in 463 trees from 27 populations in East Texas:

- total = **5379**
- polymorphic = **4264**
- used for LR = **3667**

Clinal variation - significant correlation with:

- Latitude = **210**
- Longitude = **293**
- both latitude and longitude = **34**

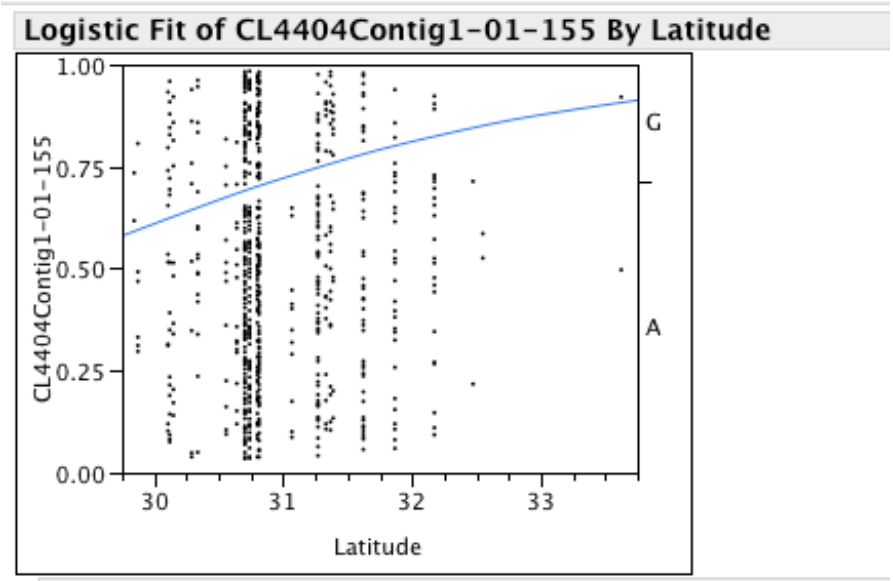
Environmental variables- significant correlation with:

- Monthly mean total annual temperature above 5˚C or Growing Degree Days (MEAN_annGDD5)* = **245**
- Mean Annual Precipitation (MEAN_annP)** = **268**
- Mean Annual Temperature (MEAN_MAT)*** = **259**
- Aridity index = *in progress*

Chhatre *et al*. 2014 *Molecular Ecology* (in prep.)

Клинальная изменчивость: Логистическая регрессия (LR)

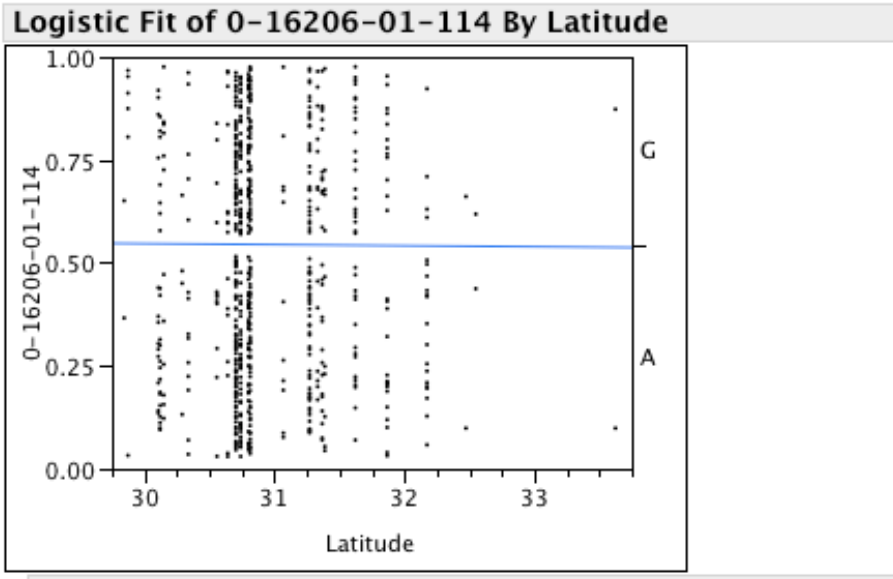
Significant association of the A/G alleles of the CL4404Contig1-01-155 SNP with Latitude:



Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	4.42624	1	8.85247	0.0029*

Insignificant association of the A/G alleles of the 0-16206-01-114 SNP with Latitude:

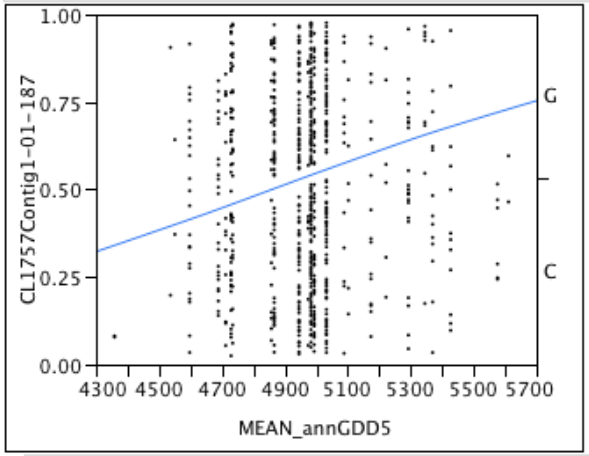


Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	0.00239	1	0.004776	0.9449

Ассоциация с изменчивостью средовых факторов: Логистическая регрессия (LR)

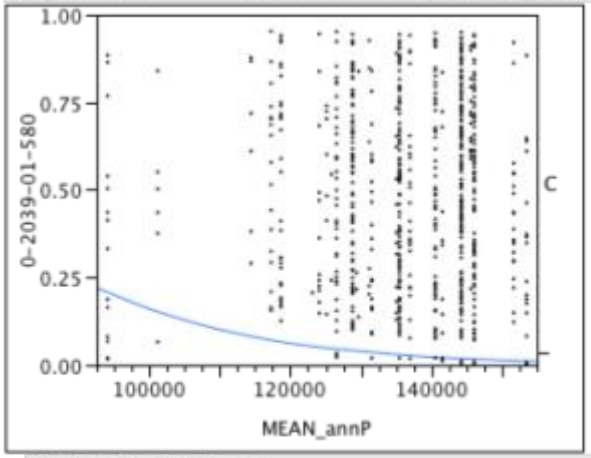
Logistic Fit of CL1757Contig1-01-187 By MEAN_annGDD5



Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	5.44054	1	10.88108	0.0010*

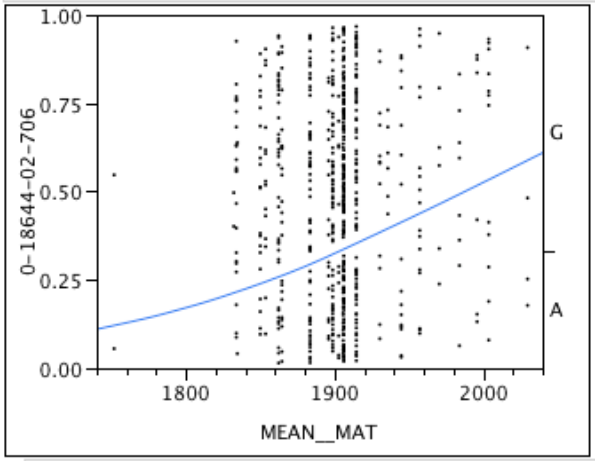
Logistic Fit of 0-2039-01-580 By MEAN_annP



Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	6.76393	1	13.52786	0.0002*

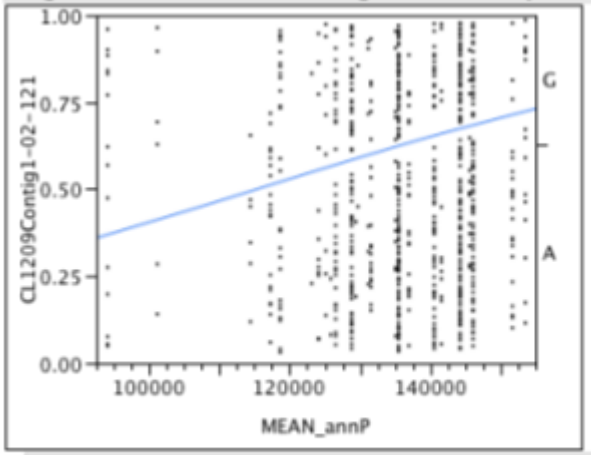
Logistic Fit of 0-18644-02-706 By MEAN__MAT



Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	6.40695	1	12.81391	0.0003*

Logistic Fit of CL1209Contig1-02-121 By MEAN_annP



Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	7.23560	1	14.47121	0.0001*

CTGN: Результаты изучения популяционной, ландшафтной и экологической геномики ладанной сосны (*Pinus taeda* L.)

- SNPs from >5000 genes were genotyped in >4500 trees sampled from numerous natural and breeding populations covering the full-range of the species
- Significant associations were found between adaptive trait phenotypes, geographic and environmental variables (temperature, growing degree-days, precipitation and aridity) and a diverse sets of genes including abiotic stress response genes ranging from trans-membrane proteins to proteins involved in sugar metabolism and transcription factors
- Numerous genes under selection were found (outliers)
- Multiple allele candidates for local adaptation were discovered

Eckert *et al*. 2010 *Genetics* 185: 969–982
Eckert *et al*. 2010 *Molecular Ecology* 19: 3789–3805
Chhatre *et al*. 2013 *Tree Genetics and Genomes* 9: 1161-1178
Chhatre *et al*. 2014 *Molecular Ecology* (in prep.)

USDA NIFA Climate Change Program 1: Regional Approaches to Climate Change

PI: Timothy Martin, 2011-2016, $19,976,825; "Integrating research, education and extension for enhancing southern pine climate change mitigation and adaptation".

ESSM Team of Scientists

Drs. Tom Byram, Carol Loopstra and Kostya Krutovsky will be the genetics team from Texas A&M University.

Among this project's main objectives are the study of loblolly pine's genetic adaptation to potential climate change. The goal is to use this knowledge to develop a new seed deployment tool that will help mitigate the detrimental effects of warmer and drier climate in the southeastern United States. Association and population genetics analysis will be used to characterize important adaptation and mitigation traits to support future breeding efforts. The genetics program will support development of growth and yield models, stand-level biophysical carbon balance modeling, multi-scale policy and economic analysis of market and non-market forest benefits and services, and an education program to deliver state-of-the-art forest management solutions. Texas A&M will assist collaborators at sister organizations in meeting these objectives through a local genetics team.

$20 million grant to study effects of climate change

http://www.pinemap.org

Six scientists from Ecosystem Science and Management will be part of a $20 million grant to study effects of climate change on agricultural and forest production

On Friday, Feb. 18, the USDA National Institute of Food and Agriculture (NIFA) awarded three Coordinated Agriculture Projects (CAP) representing a major scientific investment in studying the effects of climate change on agriculture and forest production. NIFA Director Roger Beachy made the announcement at the annual meeting of the American Association for the Advancement of Science in Washington, D.C.

"Climate change has already had an impact on agriculture production. Going forward agriculture producers need sound scientific information to plan and make decisions to ensure their economic viability," Beachy said. "These projects ensure we have the best available tools to accurately measure the effects of climate change on agriculture, develop effective methods to sustain productivity in a changing environment and pass these resources on to the farmers and industry professionals who can put the research into practice."

Institute of Food and Agriculture announced the award of a five-year, $20 million grant, to fund research, outreach and education to develop and transfer better management methods for southern pine, notably loblolly pine. They will study climate change mitigation and adaptation as it relates to southern pines, particularly loblolly pine, which comprises 80 percent of the planted forestland in the Southeast. It's widely used for lumber, pulp and paper production, and has great potential for biofuel production.

NIFA made the awards through its Agriculture and Food Research Initiative funding opportunity. AFRI's Climate Change challenge area is focused on reducing greenhouse gas emissions and increasing carbon sequestration in agricultural and forest production systems and preparing the nation's agriculture and forests to adapt to changing climates.

Two-thirds of all the drinking water in the U.S. comes from forested watersheds.

PINEMAP проект: Генетика и селекция

Основные цели:

- Обнаружение адаптивной генетической изменчивости через ассоциативное картирование и генотипирование секвенированием (genotyping-by-sequencing - GBS) и использование полученных данных для лесохозяйствования и геномной селекции ладанной сосны

Полногеномное GBS пока невозможно и необходимо избирательное секвенирование

USDA NIFA
United States Department of Agriculture
National Institute of Food and Agriculture

www.pinemap.org

PINEMAP

PINEMAP проект: избирательное секвенирование ладанной сосны (*Pinus taeda* L.) для генотипирования

1. Restriction-enzyme-based double digest Restriction site-Associated DNA Sequencing (RAD-Seq) using Illumina Hiseq2000 (R. Whetten, F. Isik: the North Carolina State University Cooperative Tree Improvement Program will genotype the Plantation Selection Seed Source Study (PSSSS) association mapping population planted across the region from four different coastal provenances).

2. Single Digest Restriction site-Associated DNA Sequencing (RAD-Seq) using Illumina Hiseq2000 (G. Peter: the Cooperative Forest Genetics Research Program at the University of Florida will genotype the Comparing Clonal Lines ON Experimental Sites (CCLONES) association mapping population).

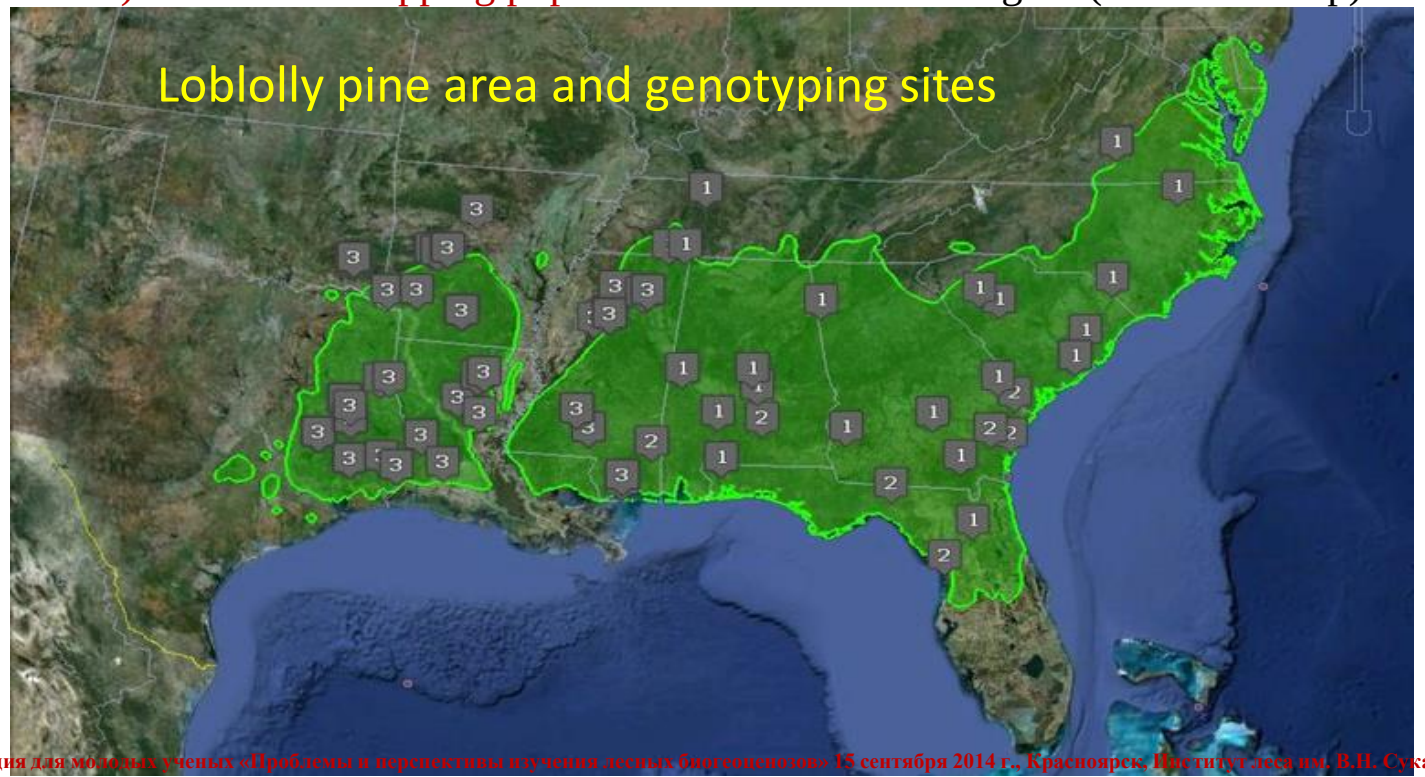
3. In solution hybridization-capture based Agilent SureSelect Target Enrichment followed by Illumina Hiseq2000 sequencing (K. V. Krutovsky, C. Loopstra, T. Byram: the Western Gulf Forest Tree Improvement Program at Texas A&M University and Texas Forest Service will genotype the Allele Discovery of Economic Pine Traits2 (ADEPT2) association mapping population for the western region; Dr. Jason Holiday: Virginia Tech).

PINEMAP проект: Генетика и селекция

1. Ross Whetten, Fikret Isik, Steve McKeand: the North Carolina State University Cooperative Tree Improvement Program will genotype the Plantation Selection Seed Source Study (PSSSS) association mapping population planted across the region from four different coastal provenances ("1" in map).

2. Gary Peter, John Davis: the Cooperative Forest Genetics Research Program at the University of Florida will genotype the Comparing Clonal Lines ON Experimental Sites (CCLONES) association mapping population ("2" in map).

3. Kostya Krutovsky, Carol Loopstra, Tom Byram: the Western Gulf Forest Tree Improvement Program at Texas A&M University and Texas Forest Service will genotype the Allele Discovery of Economic Pine Traits2 (ADEPT2) association mapping population for the western region ("3" in the map)



PINEMAP

Обогащение экзомной ДНК с помощью Agilent's SureSelect Target Enrichment System для последующего генотипирования через секвенирование (genotyping by sequencing - GBS) с помощью NGS

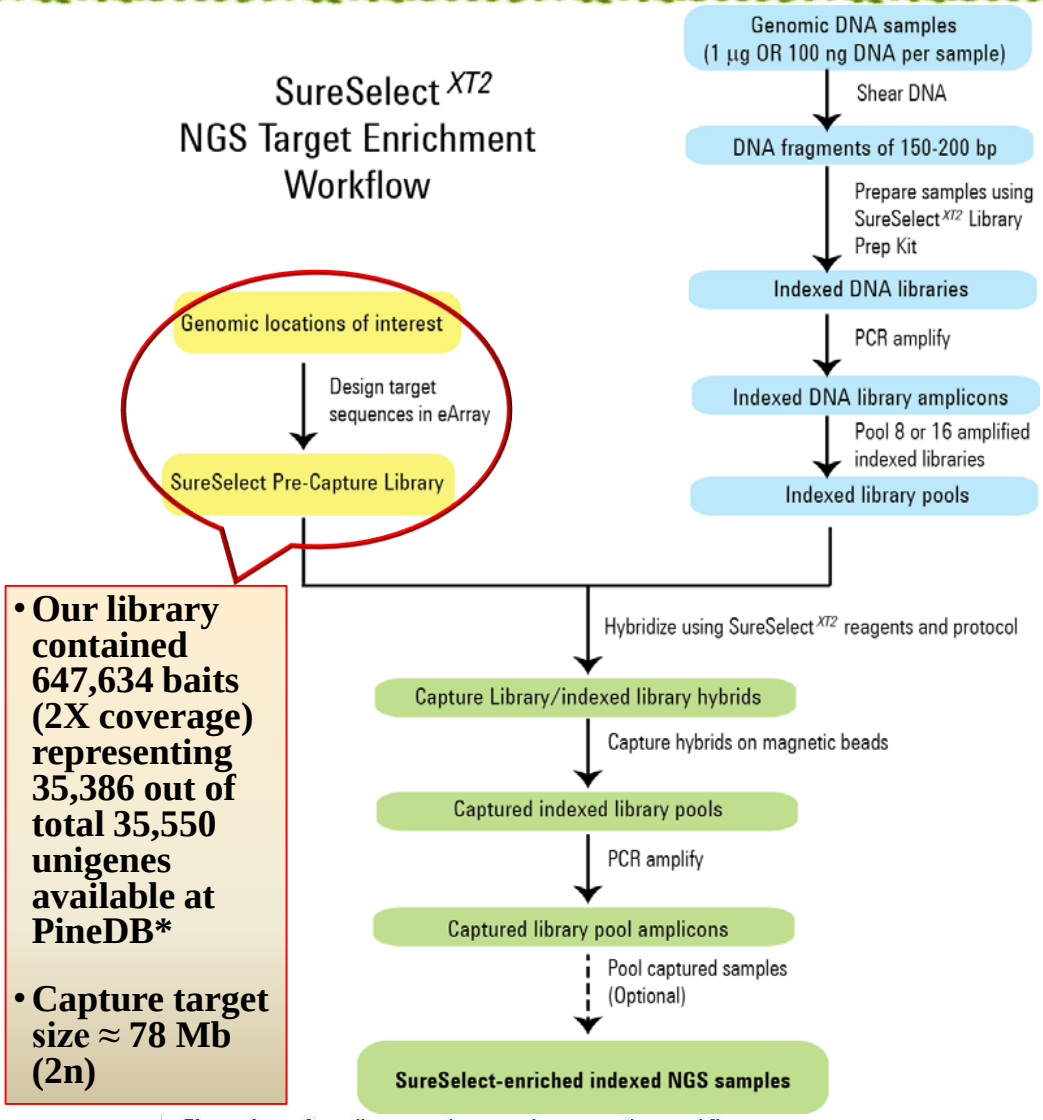


Figure 1 Overall sequencing sample preparation workflow.

- **Our library contained 647,634 baits (2X coverage) representing 35,386 out of total 35,550 unigenes available at PineDB***
- **Capture target size ≈ 78 Mb (2n)**

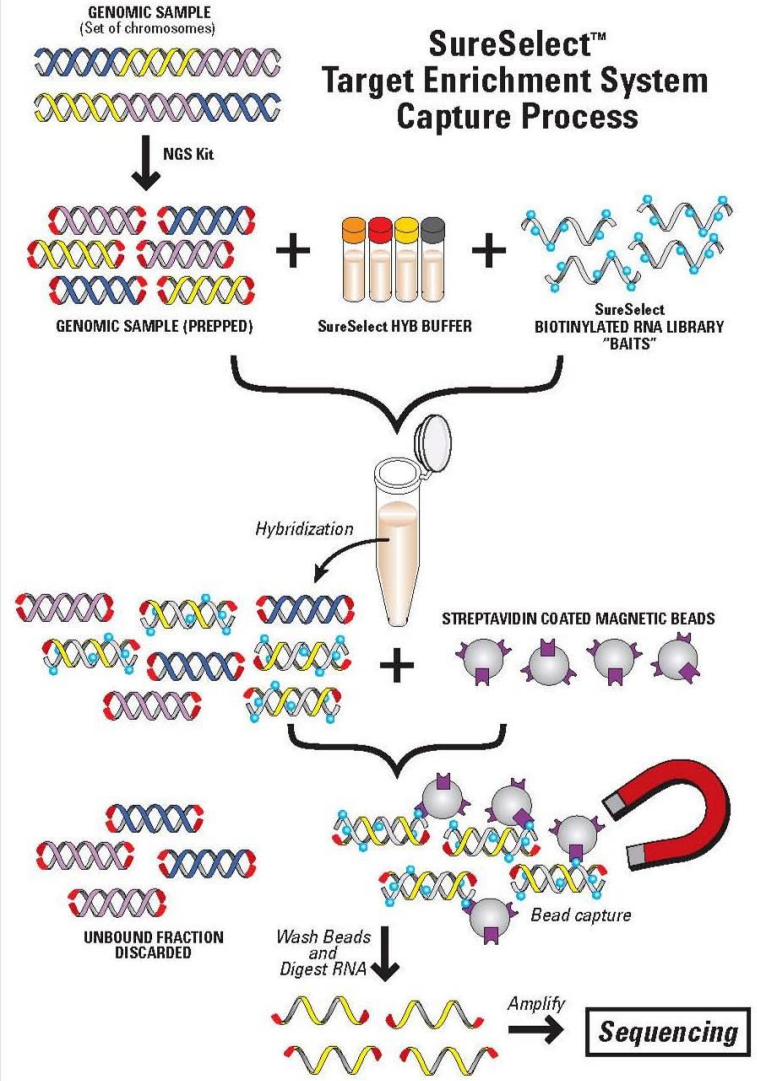
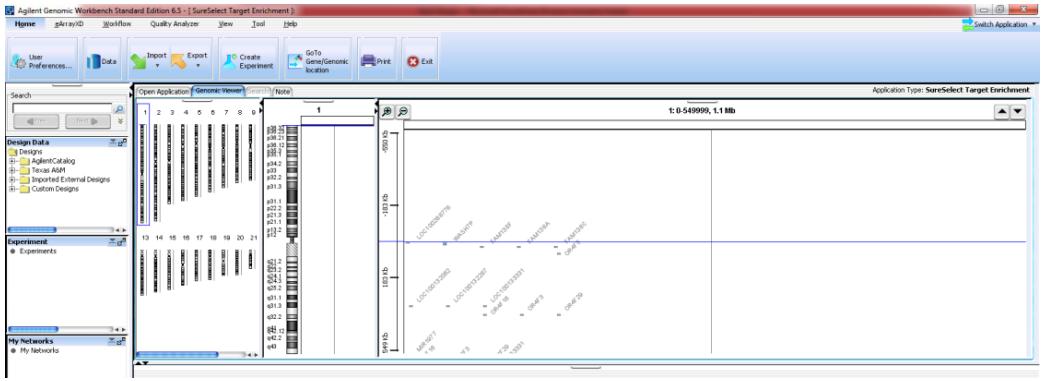


Figure 2 SureSelect Target Enrichment System Capture Process

* http://bioinfolab.muohio.edu/txid3352v1

http://www.chem.agilent.com/library/usermanuals/Public/G9630-90000.pdf

Обогащение экзомной ДНК 35,386 генов в ладанной сосне с помощью Agilent's SureSelect Target Enrichment System для последующего «штрих-кодирования» индивидуальных образцов и секвенирования в популяционных пулах



647,634 олигонуклеотидных проб длиной в 120 bp каждая были разработаны с помощью Agilent Genomic Workbench software на основе 35,386 юнигенов (сгенерированных Dr. Chun Liang, Miami University; PineDB v.1 bioinfolab.muohio.edu/txid3352v1), синтезированы и использованы для гибридизации с 78 Mb экзона для обогащения ДНК-х библиотек экзонами и последующего секвекенирования на Illumina HiSeq2000

SNP маркёры, обнаруженные с помощью SAMtools и только в уникально картированных прочтениях (reads) Illumina HiSeq2000

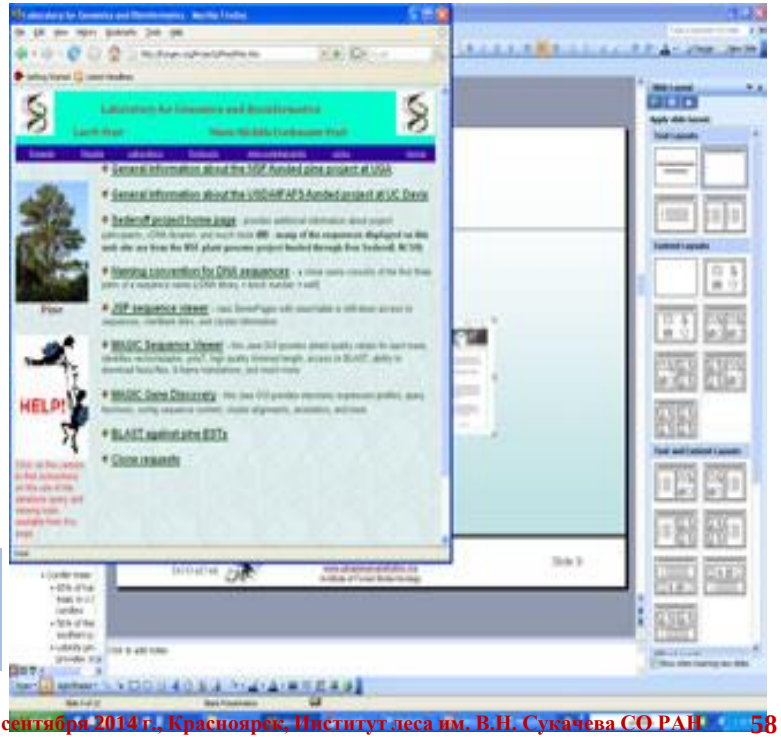
Образцов на пул	Обнаруженных SNP	Уровень разнообразия SNP, π
2	1,905,814	0.0026
4	1,870,997	0.0061
8	1,816,724	0.0050

Mapping criteria for SAMtools and Freebayes: minimum read depth = 10, at least 30% of total uniquely mapped reads contain an alternate allele

USDA NIFA Climate Change Program 1: Regional Approaches to Climate Change Project: “Integrating research, education and extension for enhancing southern pine climate change mitigation and adaptation” http://www.pinemap.org

- **2,8 млн SNPs уже генотипировано в почти 40,000 генах ладанной сосны** в моей лаборатории в Texas A&M University в этом проекте путём прямого секвенирования геномной ДНК, обогащённой экзомными районами с помощью гибридизации тотальной ДНК с 600 млн олигонуклетидных проб, представляющих почти полный транскриптом (~40 тыс. экспрессируемых генов) ладанной сосны
- более чем **400 деревьях со всего ареала, профенотипированных по большому числу адаптивных и селекционно-ценных признаков**, а также изученных по большому числу средовых факторов **будут генотипированы по всем обнаруженным SNPs** для обнаружения аллелей и гаплотипов связанных с изменчивостью адаптивных признаков, а также с устойчивостью к средовым факторам
- фактически, это означает **переход от отдельных маркёров к полному генотированию через секвенирование!**
- **эра маркёров заканчивается – наступает эра полногеномного секвенирования!**
- **популяционная геномика вместе с молекулярной экологией (экогеномикой) позволят:**
  - обнаружить гены и аллели ответственные за адаптацию
  - связать генотипы с адаптивными фенотипами и средой

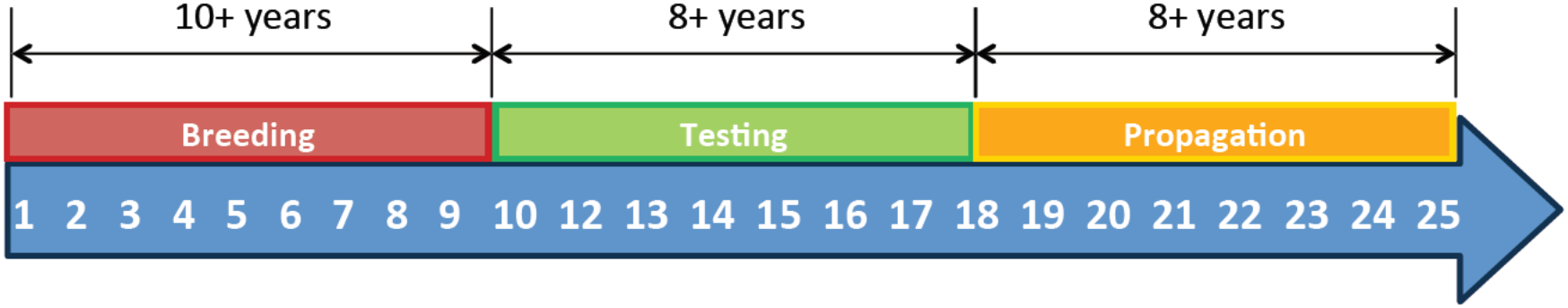
Геномные базы данных хвойных



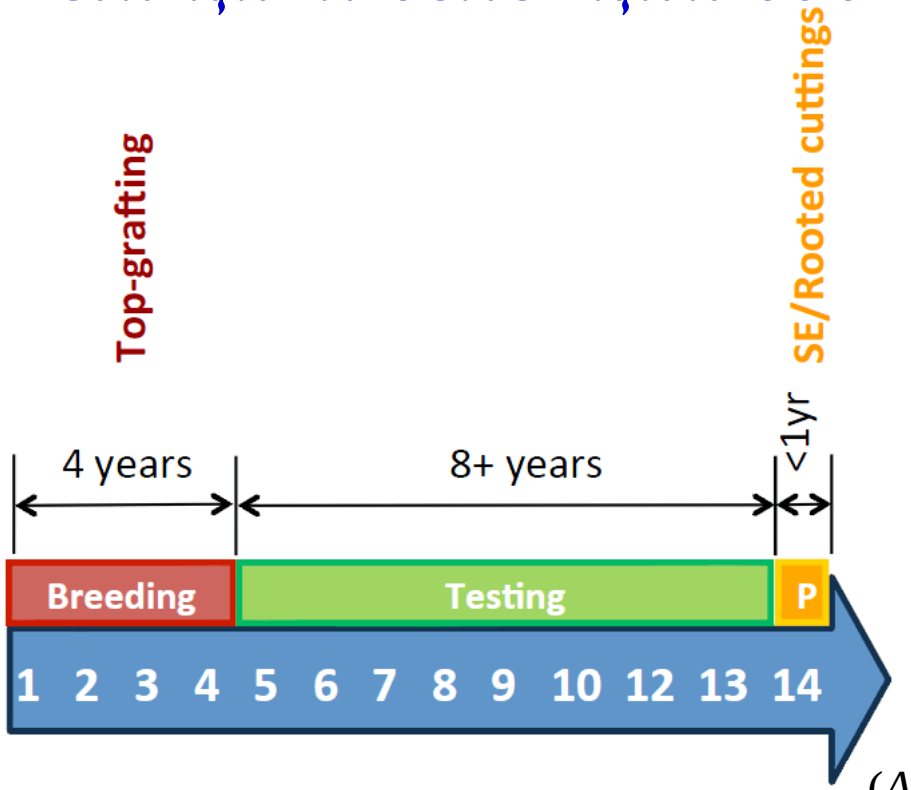
http://dendrome.ucdavis.edu

Геномная селекция генетически-улучшенных пород животных и растений, в том числе древесных, устойчивых также к экстремальным факторам среды

Традиционный цикл селекции сосны > 26 лет

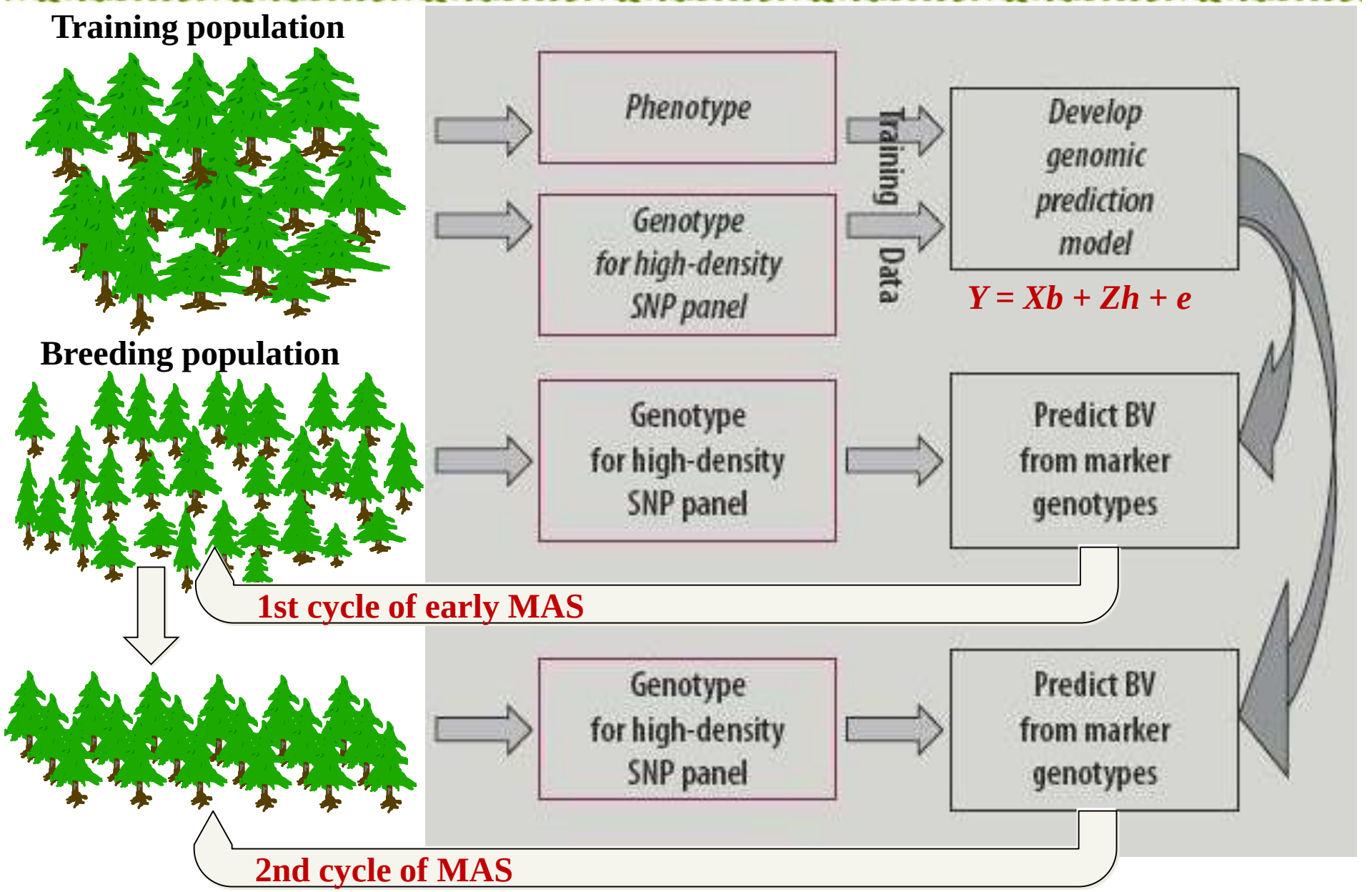


Ускоренный цикл селекции сосны > 13 лет

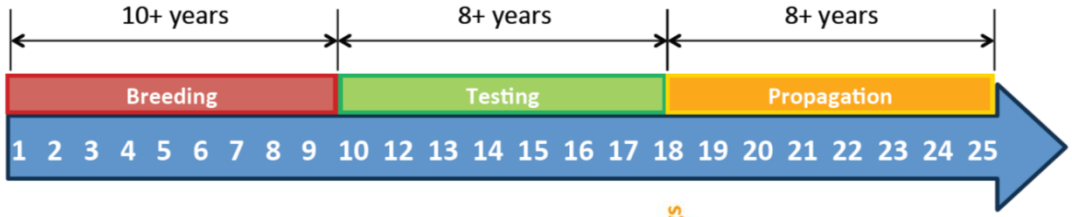


(*Adapted from Matias Kirst*)

Геномная селекция

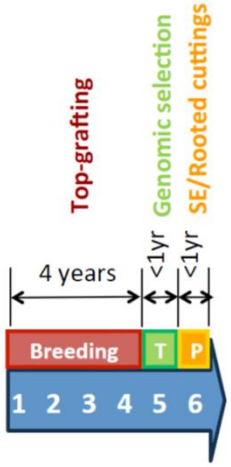


Традиционный цикл селекции сосны

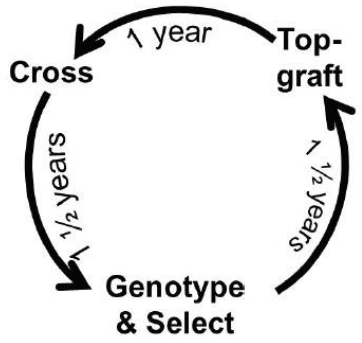


Ускоренный цикл селекции сосны

Геномная селекция



Genomic Selection Guided Crosses



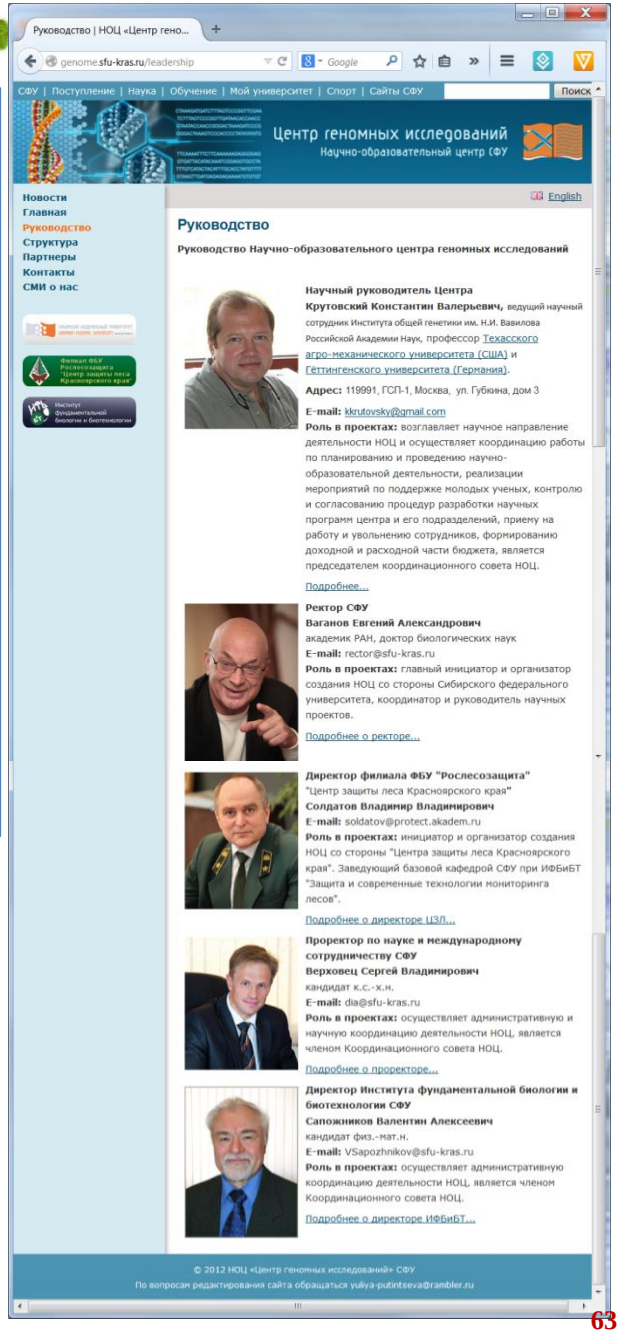
"Surgical" breeding

(*Adapted from Matias Kirst*)

Современные методы секвенирования ДНК

- Выпуск с 2005 г. коммерческих высоко-производительных секвенаторов ДНК на основе новых технологий массивного параллельного секвенирования компаниями Roche и Illumina

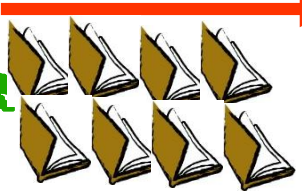
- Создание в 2012 г. Научно-образовательного центра геномных исследований Сибирского федерального университета при поддержке Отдела генетики и селекции Центра защиты леса Красноярского края и Лаборатории лесной генетики и селекции Института леса им. В. Н. Сукачева СОРАН (http://genome.sfu-kras.ru) оборудованного самым производительным секвенатором Illumina HiSeq 2000 и суперкомпьютером



Проблема полногеномного секвенирования с помощью новых секвенаторов:

возможность секвенировать только короткими фрагментами 100-150 но

Выделение тотальной геномной ДНК



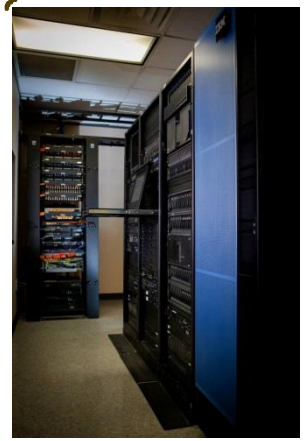
Длинные фрагменты ДНК

Дополнительное фрагментирование ДНК

Короткие фрагменты ДНК

Приготовление библиотек для секвенирования и кластеризация (cBot)

Секвенирование на HiSeq 2000 и генерация сотен миллионов коротких прочтений («ридов»)



Биоинформатическая обработка миллионов «ридов» и сборка их в контиги и скаффолды на IBM x3950 X6 суперкомпьютере

Заключение

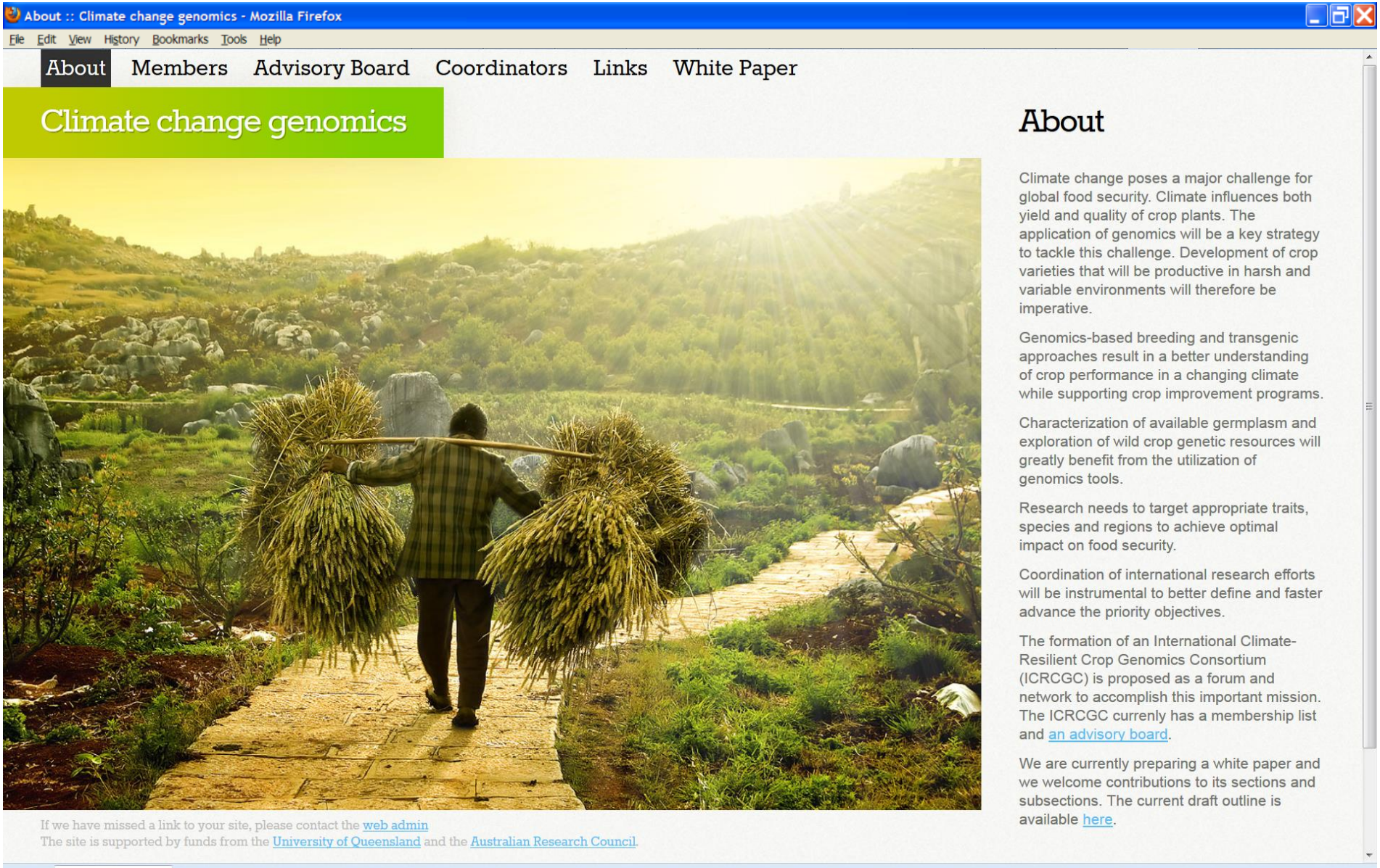
Основной вклад геномики в изучение лесных популяций:

- разделение популяционной структуры на селективно-нейтральную структуру, вызванную генетическим дрейфом и адаптивную структуру, вызванную естественным отбором
- выявление генов, аллелей и участков генома, ответственных за локальную адаптацию, и, следовательно, важных для сохранения
- более точное установление эффективного размера популяций
- более точное установление демографических событий в истории популяций, вызванных флюктуацией размера, в том числе «bottlenecks» – в отличие от влияния отбора (важно для принятия правильных природоохранных мероприятий)

Заключение

- **Полногеномное или целевое** (target) **секвенирование стали реальностью и наиболее информационным способом генотипирования** (genotyping by sequencing - GBS)
- Полногеномное ассоциативное картирование - **эффективный популяционно-геномный подход к обнаружению связи генотипа и фенотипа**
- Геномная селекция должна помочь получению более устойчивых деревьев
- **Хвойные – идеальная система для популяционно-геномного анализа:**
  - гаплоидный мегагаметофит - идеальная система для прямого секвенирования, обнаружения SNP и анализа гаплотипов
  - высокий уровень фенотипической и нуклеотидной изменчивости
  - слабое внутригенное LD
  - возможность клонирования и оценки одного и того же генотопа в разных средах
  - мощная селекционно-производственная инфраструктура

The International Climate-Resilient Crop Genomics Consortium (ICRCGC) http://www.climatechangegenomics.org



The International Climate-Resilient Crop Genomics Consortium (ICRCGC) http://www.climatechangegenomics.org/members.php

About :: Climate change genomics - Mozilla Firefox

File Edit View History Bookmarks Tools Help

About Members Advisory Board Coordinators Links White Paper

Climate change genomics

Genomics of Climate Resilient Crops

1. Assessment of effects climate change on agriculture with examples from case studies on major crop plants
2. Work done so far on genetics and breeding for climate-resilience traits (CRTs)
3. Rationale for using genomics resources and allied gene pools (AGPs) including wild crop relatives (WCRs) for accelerated breeding for adaptation

Genomics in major crops with the following examples: briefs on classical genetics and traditional breeding for CRTs and genetic mapping and molecular breeding of/for CRTs – information available from genome drafts – structural and functional genomics resources focusing CRTs – libraries, transctipomics, proteomics, metabolomics – utilization of AGPs – requirement of WGS and genotyping by sequencing of AGPs.

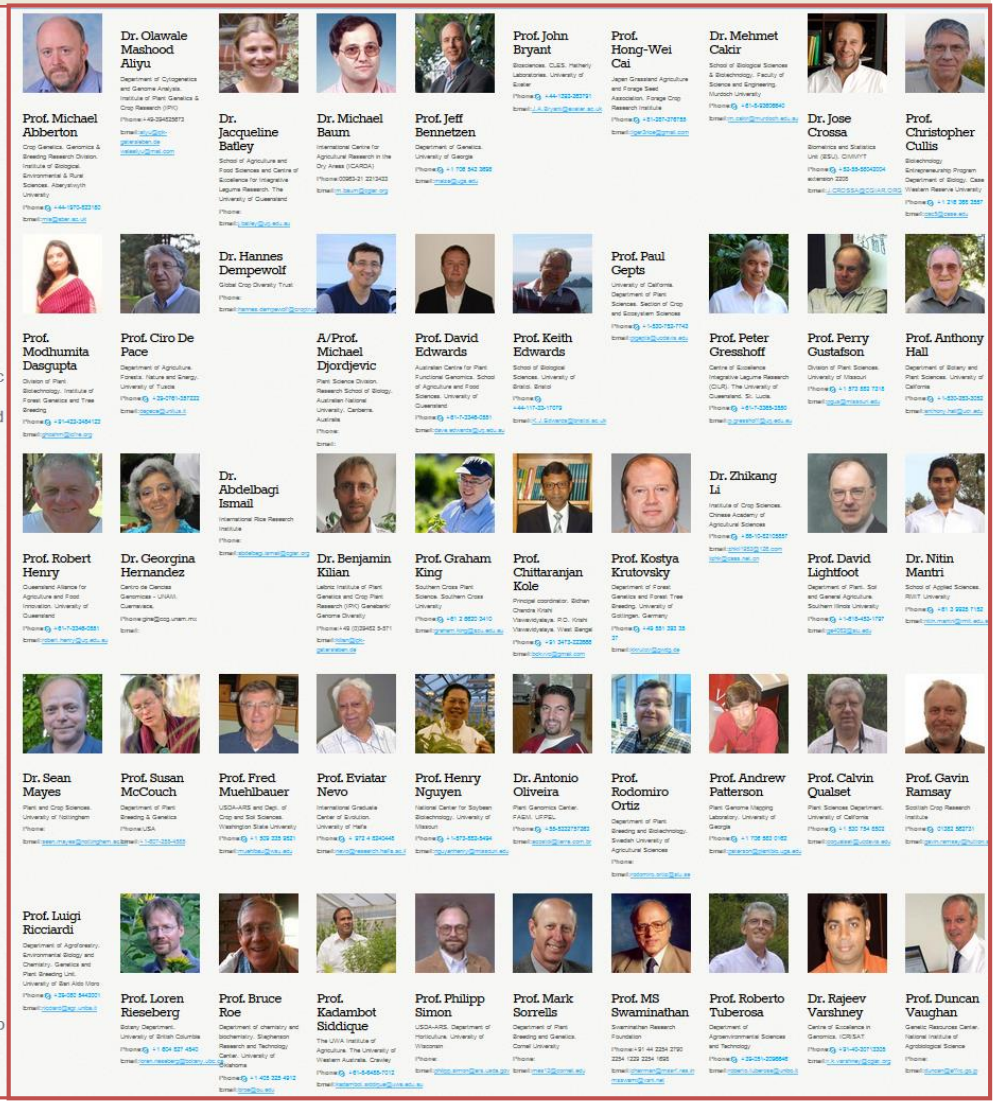
1. Introduction
2. Cereals: Rice, Maize, Wheat, Sorghum, pearl Millet
3. Oilseeds and Pulses: Soybean, Brassicas
4. Pulses: Pigeonpea, Cowpea, Common Bean
5. Vegetables: Tomato, Cucumber, Melon, Water Melon
6. Fruit Crops: Apple, Peach, Grapes, Papaya, Apricot, Almond, Cherry, Plum, Strawberry, blueberry
7. Forest Trees: Eucalyptus, Poplar, Oak, Chestnut, Pine
8. Industrial Crops: Cotton, Cocoa, Sugarcane

Proposed strategies for improvement in CRTs: focus on advanced tools and AGPs

1. Introduction
2. Early and late maturity
3. Drought adaptation
4. Cold tolerance
5. Heat stress tolerance
6. Flooding and submergence tolerance
7. Salinity tolerance
8. Disease resistance
9. Insect resistance
10. Higher nutrient and water use efficiency
11. CO2 sequestration
12. Greenhouse gas emission

Potential for genomic characterization of wild and collected germplasm to enhance global germplasm exchange and use in crop improvement – socio-political and regulatory issues

Education on genomics for plant breeders and plant breeding for genomicists



Благодарность за финансирование

- United States Department of Agriculture
- National Science Foundation, USA
- Western Gulf Forest Tree Improvement Program, USA
- Texas Forest Service and Industry Partners, USA
- Conifer Translational Genomics Network
- Genetics Graduate Program, Texas A&M University
- Siberian Federal University
- Ministry of Science and Education, Russian Federation
- Russian Foundation for Basic Research
- Russian Government

Благодарность соавторам и коллегам

V. N. Sukachev Institute of Forest
Институт Леса им. В.Н. Сукачёва

Александр Александрович Онучин
Леонид Иосифович Милютин
Наталья Викторовна Орешкова
Ираида Николаевна Третьякова
Игорь Николаевич Павлов
Елена Николаевна Муратова

Genome Research and Education Center
Siberian Federal University

Евгений Александрович Ваганов
Сергей Васильевич Первухин
Валентин Алексеевич Сапожников
Сергей Владимирович Верховец
Юлия Андреевна Путинцева

Krasnoyarsk Center for Forest Protection

Владимир Владимирович Солдатов
Елена Алексеевна Шилкина
Ирина Владимировна Чубугина
Алексей Александрович Ибе
Ксения Олеговна Дейч

Институт общей генетики им. Н.И. Вавилова РАН

Николай Каземирович Янковский
Дмитрий Владиславович Политов

Dr. Carol Loopstra

Texas A&M University ESSM

Dr. Tom Byram

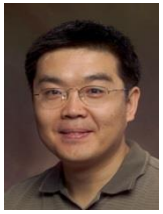
TEXAS FOREST SERVICE
The Texas A&M University System

Dr. David Neale
UCDAVIS
DEPARTMENT OF PLANT SCIENCES
College of Agricultural and Environmental Sciences

THE UNIVERSITY OF CALIFORNIA DAVIS

Dr. Jill Wegrzyn
UCONN
Biotechnology•Bioservices Center

UNIVERSITY OF CONNECTICUT 1881



Dr. Chang Liang
Department of Botany

MIAMI UNIVERSITY
OXFORD OHIO

SIFG
Southern Institute of Forest Genetics

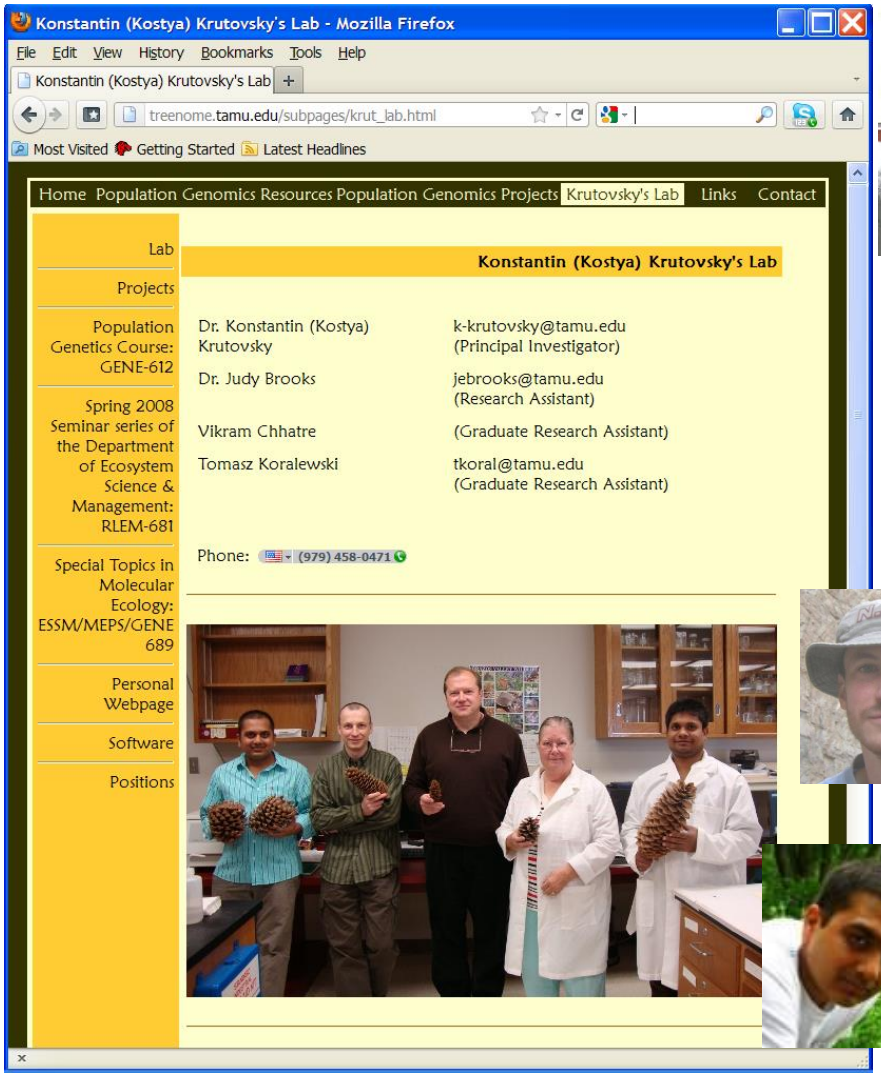
Dr. Nurul Islam-Faridi
Dr. Dr. Dana Nelson
Dr. Craig Echt
Sedley Josserand

FOREST SERVICE US DEPARTMENT OF AGRICULTURE

Благодарность сотрудникам

Laboratory of Forest Genomics

TAMU Department of Ecosystem Science and Management



Dr. Konstantin Krutovsky
Adjunct Professor
k-krutovsky@tamu.edu
(979) 458-1417

Dr. Judy Brooks
Research Assistant
jebrooks@tamu.edu

Xixu Cai
Graduate Student
xixucai@tamu.edu

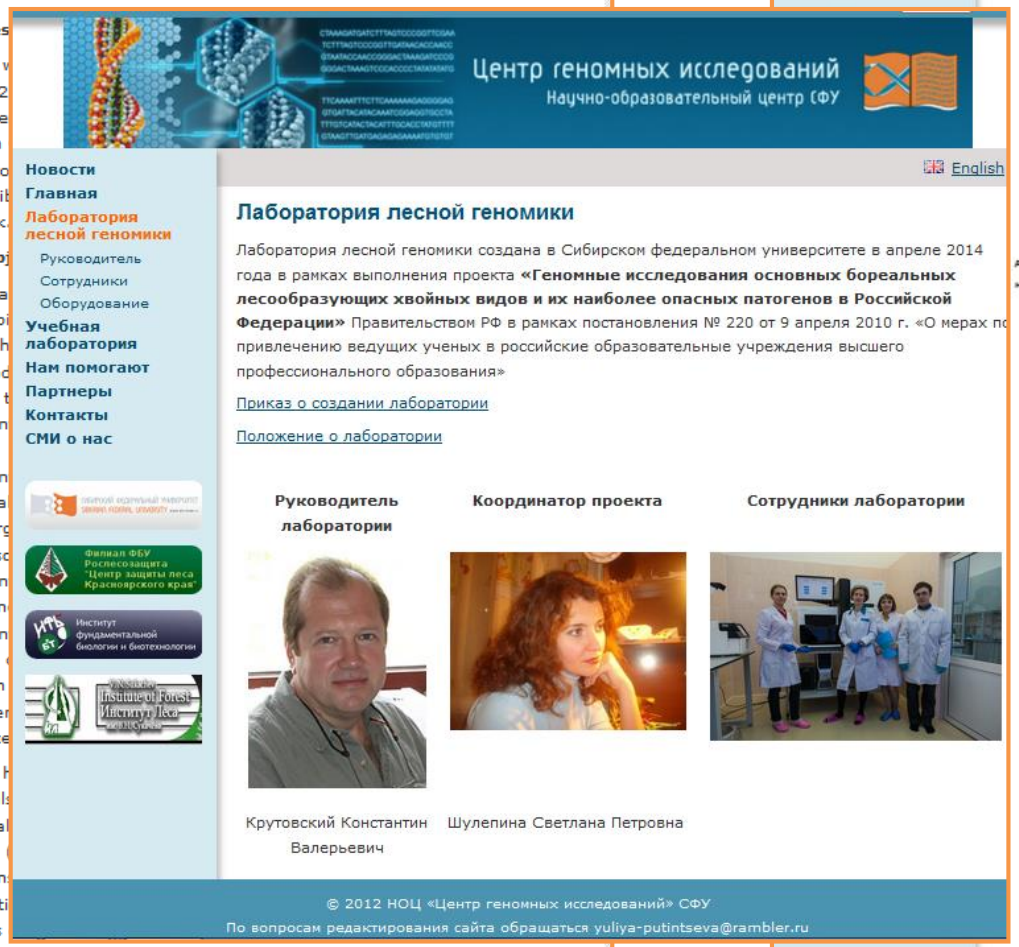
Dr. Tomasz Koralewski
Former Graduate Student, now Research Associate
tkoral@tamu.edu

Mengmeng Lu
Graduate Student
mira0501@tamu.edu

Dr. Vikram Chhatre
Former Graduate Student
crypticlineage@gmail.com
(now Research Associate at the University of Maryland)

Habibul Islam
Graduate Student
milon237@tamu.edu

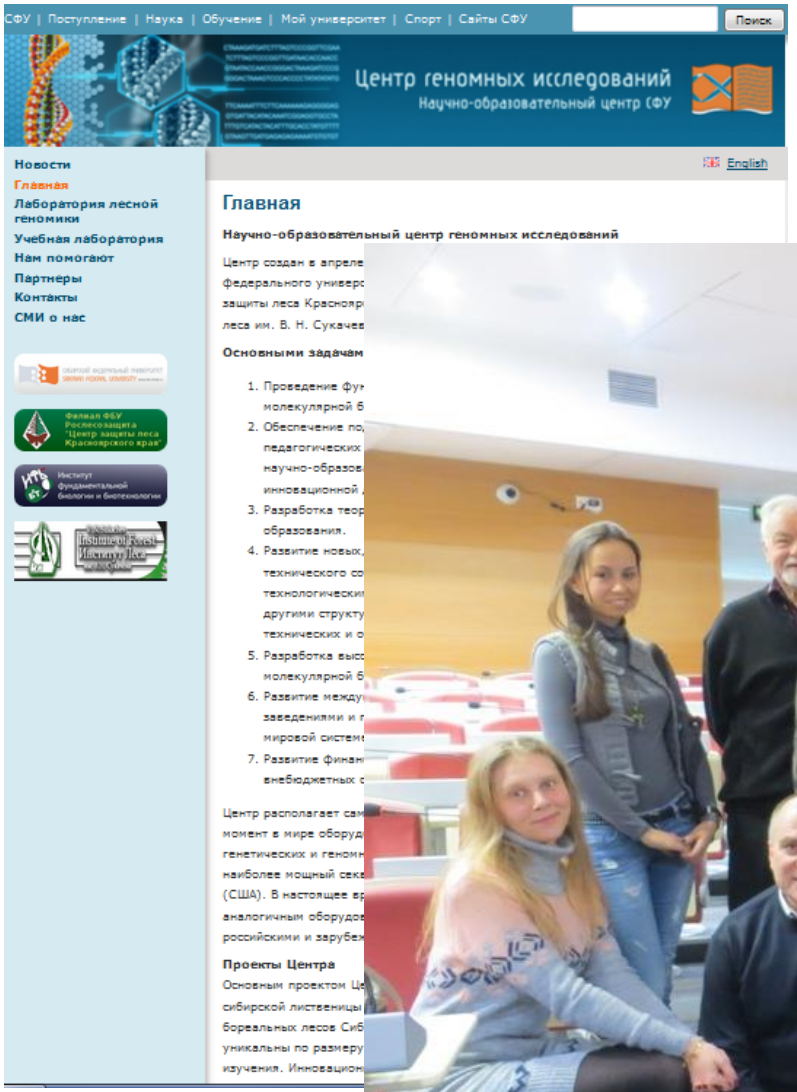
Благодарность сотрудникам НОЦ геномных исследований и Лаборатории лесной геномики СФУ



http://genome.sfu-kras.ru/en/forest_genomics

Благодарность сотрудникам СФУ, Центра лесозащиты, НОЦ геномных исследований и Лаборатории лесной геномики СФУ

http://genome.sfu-kras.ru/en/forest_genomics

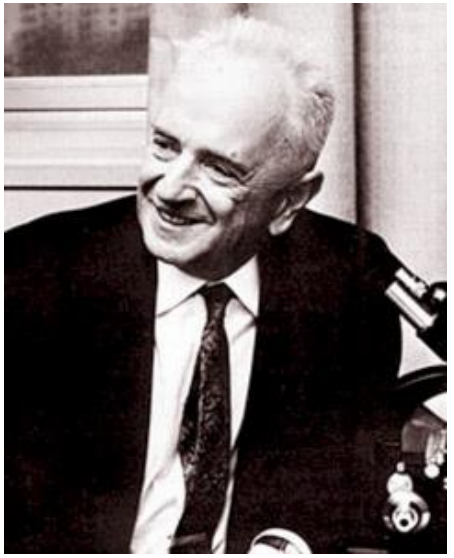


- **20-ый век:**

  Эволюционное учение + Генетика = Синтетическая теория эволюции (Генетическая теория эволюции или Эволюционная генетика)

  ↓

  популяционный уровень мышления



Феодосий Добжанский (1900-1975)

- **21-ый век:**

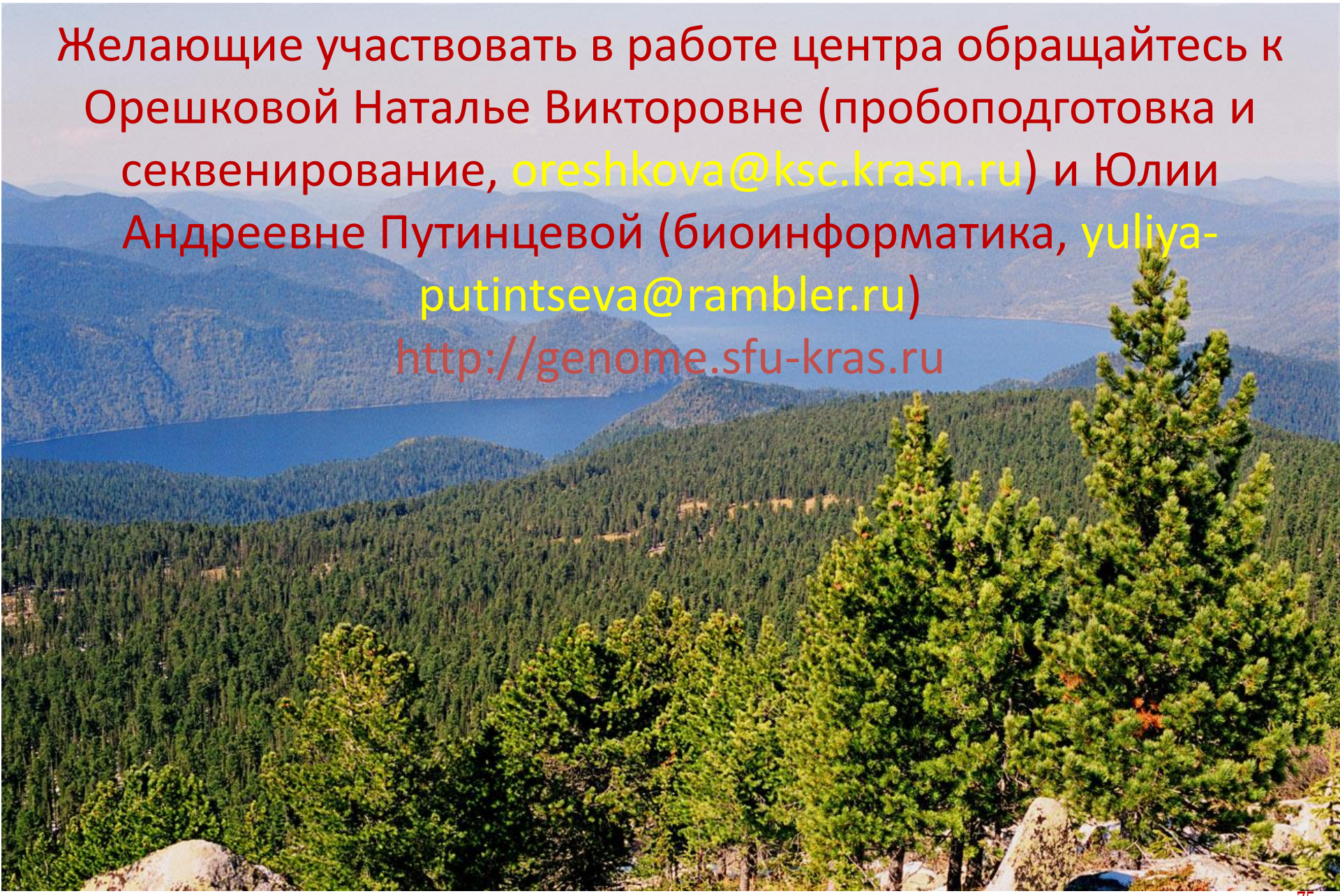
  Молекулярная генетика + Биоинформатика = Геномика

  ↓

  популяционно-геномный уровень мышления

Крутовский К. В. От популяционной генетики к популяционной геномике лесных древесных видов: Интегрированный популяционно-геномный подход // Генетика. 2006. Т. 42. №10. С. 1304–1318.

Спасибо за внимание!



Желающие участвовать в работе центра обращайтесь к Орешковой Наталье Викторовне (пробоподготовка и секвенирование, oreshkova@ksc.krasn.ru) и Юлии Андреевне Путинцевой (биоинформатика, yuliya-putintseva@rambler.ru)

http://genome.sfu-kras.ru