



# Современные методы анализа и визуализации пространственной генетической структуры лесных популяций



## Modern methods of analysis and visualization of spatial genetic structure of forest tree populations

**Политов Д.В., Белоконь Ю.С.**

*Лаборатория популяционной генетики им. акад. Ю.П. Алтухова*

*Институт общей генетики им. Н.И. Вавилова РАН*

**Politov D.V., Belokon Yu.S.**

*Laboratory of Population Genetics named after Acad. Yu.P. Altukhov*

*Vavilov Institute of General Genetics Russian Academy of Sciences*

V КОНФЕРЕНЦИЯ  
НАУЧНЫЕ ОСНОВЫ  
УСТОЙЧИВОГО  
УПРАВЛЕНИЯ ЛЕСАМИ



VIII КОНФЕРЕНЦИЯ  
АЭРОКОСМИЧЕСКИЕ МЕТОДЫ  
И ГЕОИНФОРМАЦИОННЫЕ  
ТЕХНОЛОГИИ  
В ЛЕСОВЕДЕНИИ, ЛЕСНОМ  
ХОЗЯЙСТВЕ И ЭКОЛОГИИ





# Особенности древесных растений с точки зрения популяционной и экологической генетики

- ✓ **долговечность/длинноцикловость** → невозможен или сильно затруднен традиционный гибридологический анализ (результатов скрещивания придется ждать десятилетиями) – особенно эффективно использование МГМ
- ✓ **большой размер и сложность генома** → мало полногеномных данных
- ✓ **экономическое значение** → внимание госструктур и бизнеса, лесосеменное районирование, охрана и рациональная эксплуатация генофондов, нелегальные рубки
- ✓ **экологическое значение** (виды-эдификаторы лесных экосистем) → внимание ученых и общественности
- ✓ у многих видов - **большие непрерывные ареалы** + ветроопыление + анемохория → трудно определять границы популяций и субпопуляций + не работает островная модель
- ✓ ареалы с **гетерогенными условиями** → клинальная изменчивость, локальные адаптации, мозаичность паттернов генетической структуры





# Угрозы



Глобальные климатические сдвиги  
Пожары  
Патогены и вредители  
Загрязнение среды  
Нерациональная эксплуатация человеком

Все это приводит к эрозии лесных генофондов



# Анализ пространственной генетической дифференциации

- Оценка степени генетической дифференциации
- Анализ современной геногеографической картины (распространение аллелей и гаплотипов)
- Реконструкция филогеографических паттернов («внутривидовая филогения»)
- Выявление факторов динамики генофондов



## ТЕОРИЯ

- Интенсивность и механизмы миграции
- Механизмы локальная экогенетической и эпигенетической адаптации, естественный отбор
- Дрейф генов («бутылочное горлышко» и «эффект основателя»)



## ПРАКТИКА

- Генетическая идентификация, паспортизация и сертификация
- Разработка стратегий сохранения и рациональной эксплуатации генофондов
- Лесосеменное районирование
- Меры борьбы с незаконным оборотом продукции
- ЛПК





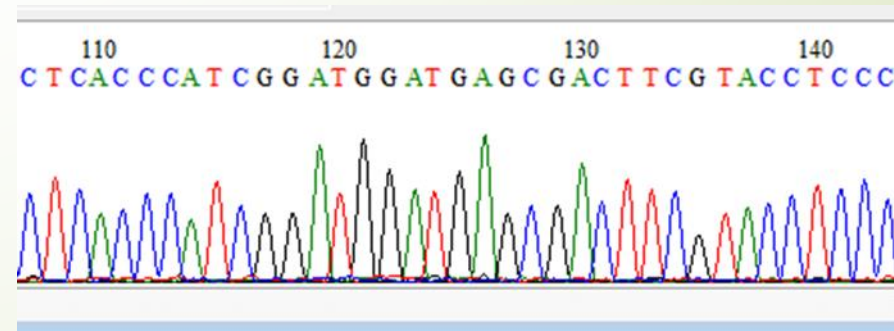
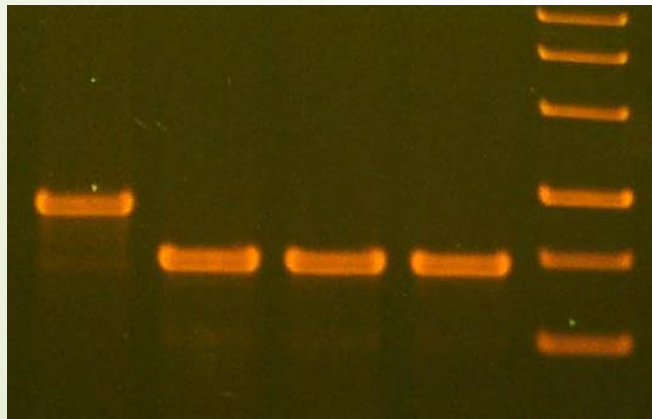
# Роль генетического разнообразия

- ▶ Популяционный уровень организации живой материи - ключевой для функционирования биосферы, поскольку именно на нём происходит передача наследственной информации следующему поколению, реализуются адаптационные и эволюционные изменения.
- ▶ Популяции лесных древесных растений - основой существования широко распространённых биоценозов.
- ▶ Эдификаторами ряда типов лесов Палеарктики, например, бореальных лесов, являются виды древесных растений с огромными ареалами, в пределах которых наблюдается разнообразие условий, к которым виды-лесообразователи должны адаптироваться, и в этой адаптации ключевую роль играет генетическая изменчивость.



# Оценка структуры генетического разнообразия, роль молекулярных маркеров

- ▶ Задача оценки уровней и распределения генетического разнообразия по ареалам древесных растений долгое время представляла методически сложную проблему из-за большой доли модификационной компоненты фенотипической изменчивости.
- ▶ С развитием методов молекулярно-генетического маркирования появилась возможность оценивать непосредственно генетической изменчивости, сначала на базе серологических маркеров, полиморфизма ферментных и неферментных белков.



## Роль генетического разнообразия

- ▶ Классический подход - оценка генетической подразделённости (genetic subdivision) по молекулярно-генетическим маркерам (МГМ) при помощи  $F$ -статистик (или  $G$ -статистик) Райта, т. е. разложение общей изменчивости на её внутри- и межпопуляционную компоненты
- ▶ для этого рассчитываются частные коэффициенты инбридинга, или индексы фиксации (S. Wright, 1973, 1978)
- ▶ Значения индекса фиксации  $F_{ST}$  ( $G_{ST}$ ,  $N_{ST}$ ,  $R_{ST}$ ) варьируют от 0 (полная гомогенность частот аллелей между субпопуляциями) до 1 (фиксация разных аллелей в субпопуляциях, то есть отсутствие общих аллелей).

$$F_{ST} = 1 - \frac{H_S}{H_T} = 1 - \frac{2p_S q_S}{2p_T q_T}$$

# Генетическая подразделенность

Представления о подразделенности видовых генофондов, иерархической организации популяционных систем и о происходящих них процессах стали уже классическими благодаря:

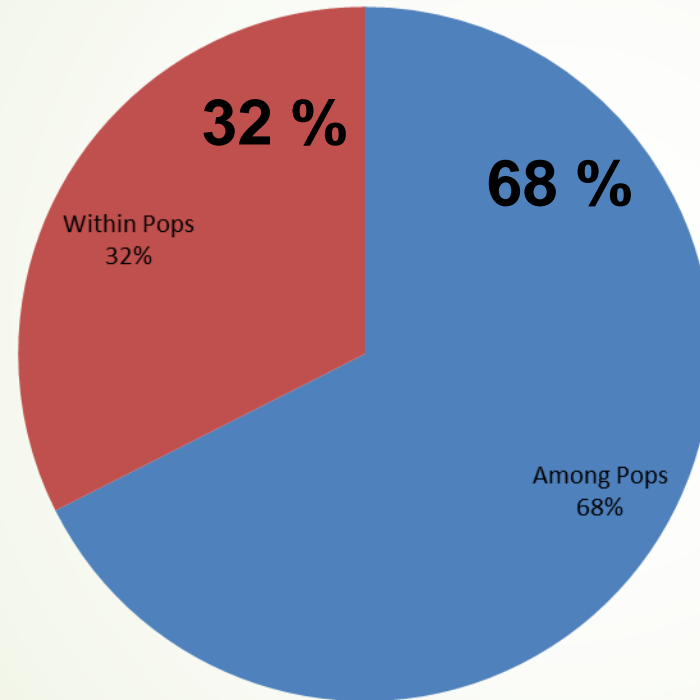
- ▶ теоретическим разработкам (Wright, 1973; Weir & Cockerham, 1984; Алтухов, 2003), AMOVA (Excoffier et al., 1992)
- ▶ результатам математического моделирования
- ▶ эмпирическим данным, полученным главным образом с помощью молекулярных методов.



# Анализ молекулярной дисперсии (AMOVA) у елей комплекса *Picea abies* – *P. obovate* по мтДНК



Percentages of Molecular Variance



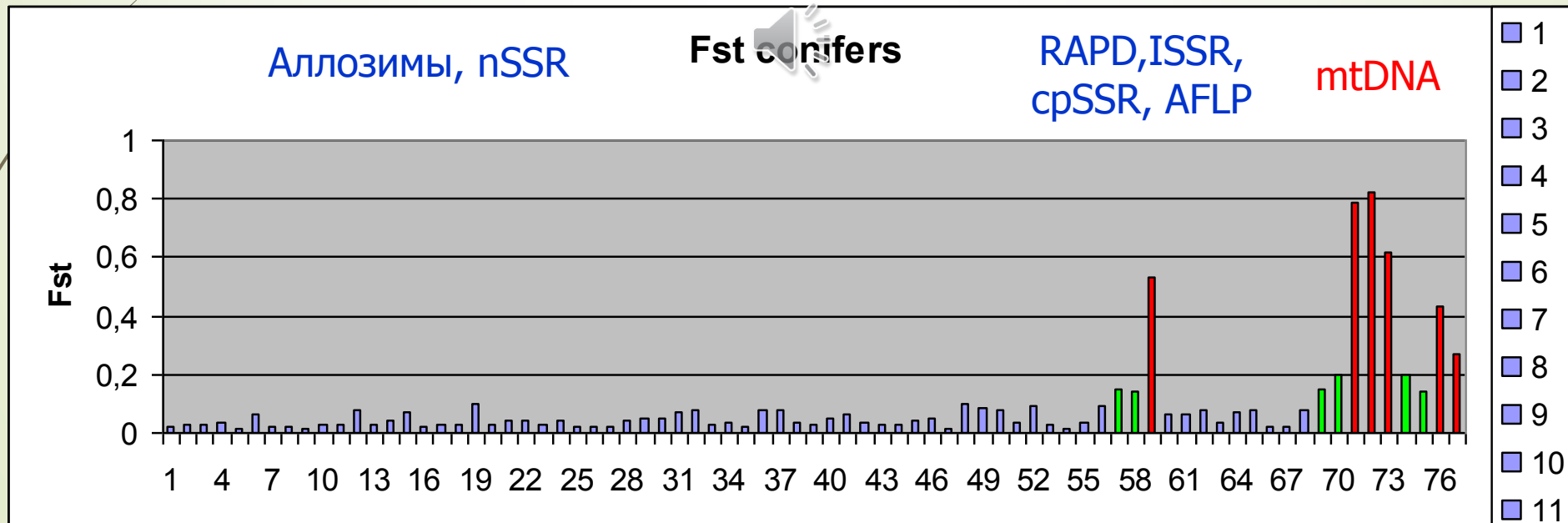
аллозимы: **Fst = 9.9 %**  
(Krutovskii, Bergmann, 1995)

микросателлиты: **Fst = 2.9 %**  
(Tollefsrud et al., 2009)

**Fst = 9.7 %**  
(Мудрик и др., 2008)

# Эмпирические данные

- Диапазон значений  $F_{ST}$  для популяций хвойных России:





# Подходы и параметры:

Оценка уровней и структуры межпопуляционной генетической дифференциации

оценка степени изоляции (ограничения панмиксии)

оценка влияния отбора

Филогеография (внутривидовая филогения)

$F_{ST}$  доля межпопуляционной изменчивости в общей изменчивости (Wright 1978, Nei 1977, Weir & Cockerham 1984), AMOVA

$D_N$  генетическое расстояние Нея (Nei 1972)

UPGMA кластеризация методом невзвешенных парно-групповых средних

NJ кластеризация методом присоединения ближайшего соседа

Байесовская кластеризация

MDS неметрическое многомерное шкалирование

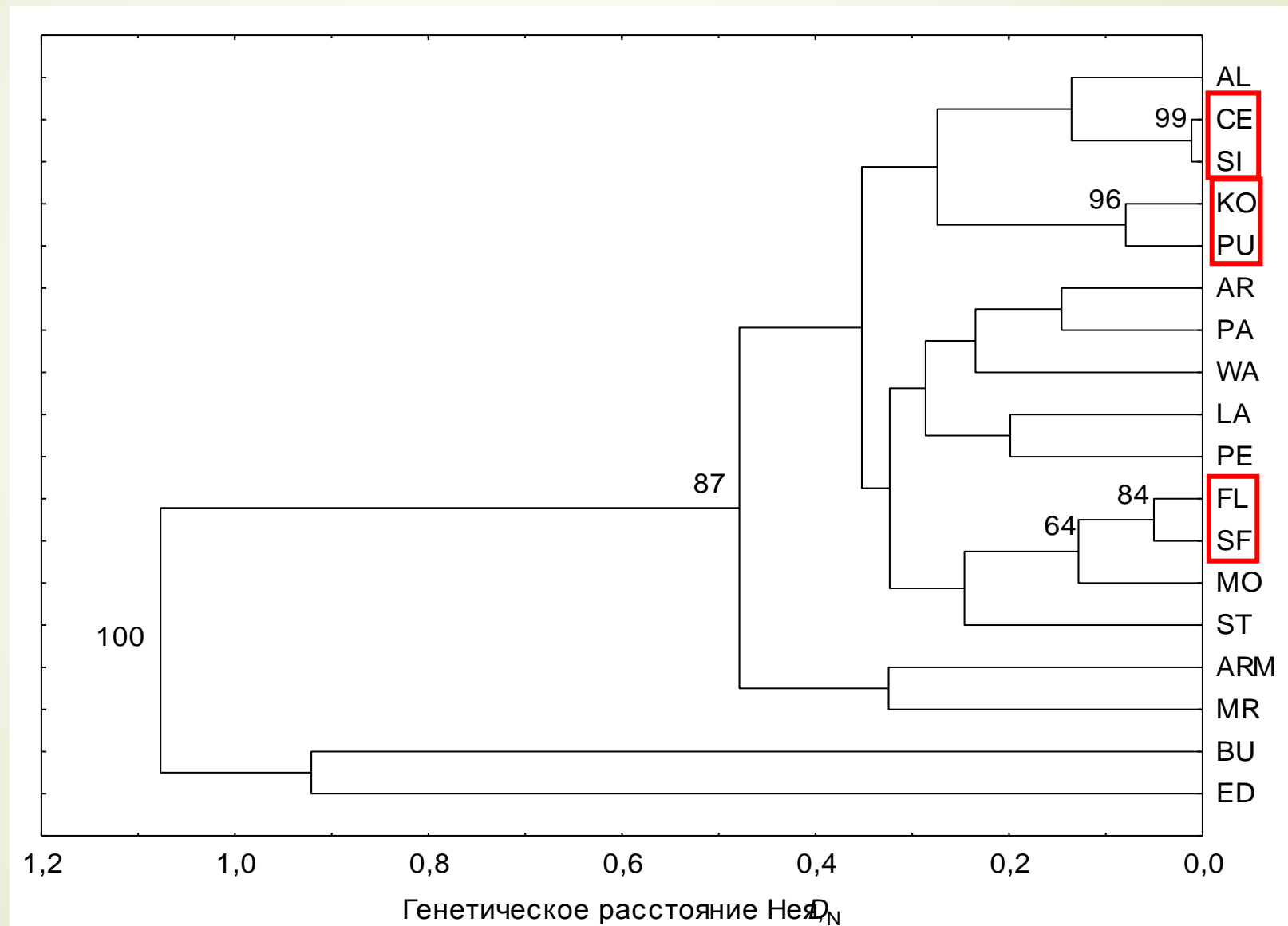
PCA анализ главных компонент

Анализ соответствий

Анализ многолокусных генотипов (*STRUCTURE*, байесовские подходы)

# UPGMA-фенограмма генетической дифференциации видов сосен секции *Strobus* по аллозимам

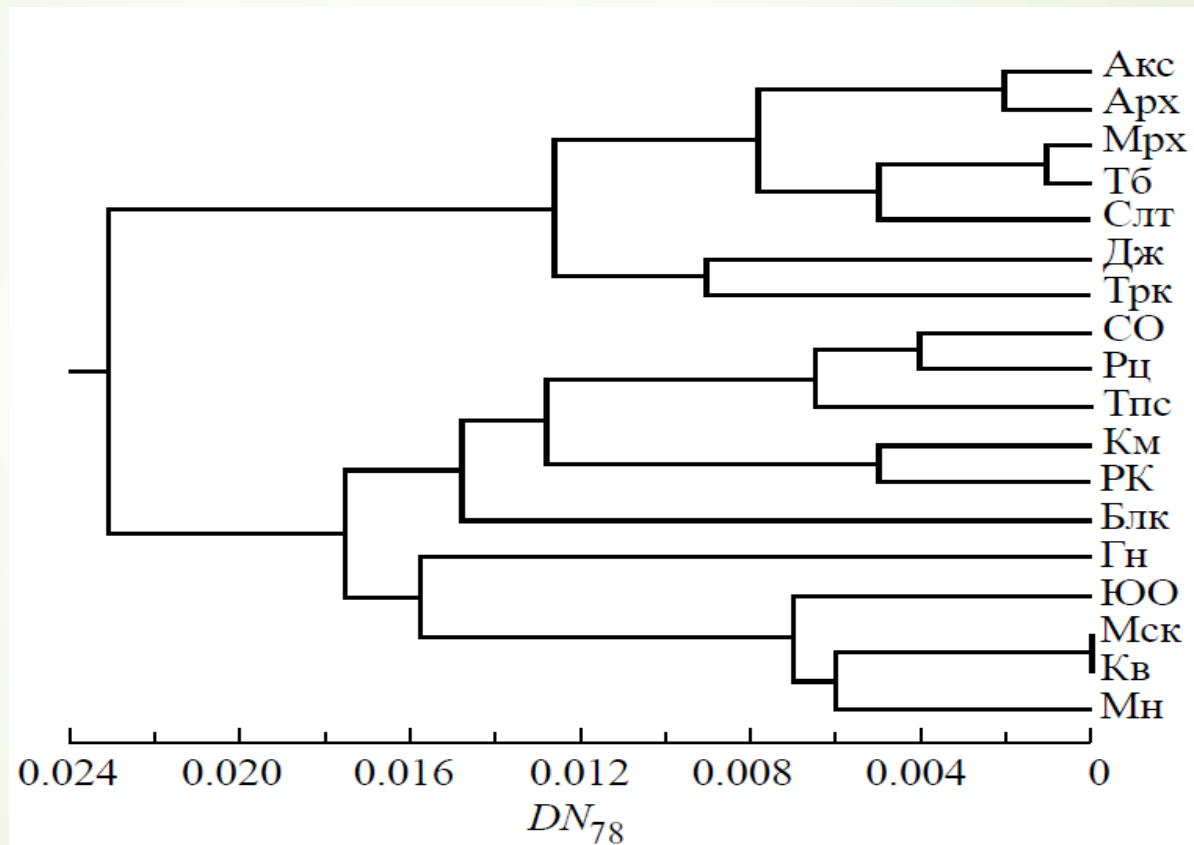
(Белоконь и др., 1998; Politov, Krutovskii, 2004, Belokon, 2007; Politov, 2007)





# Кластерный анализ выборок одного вида

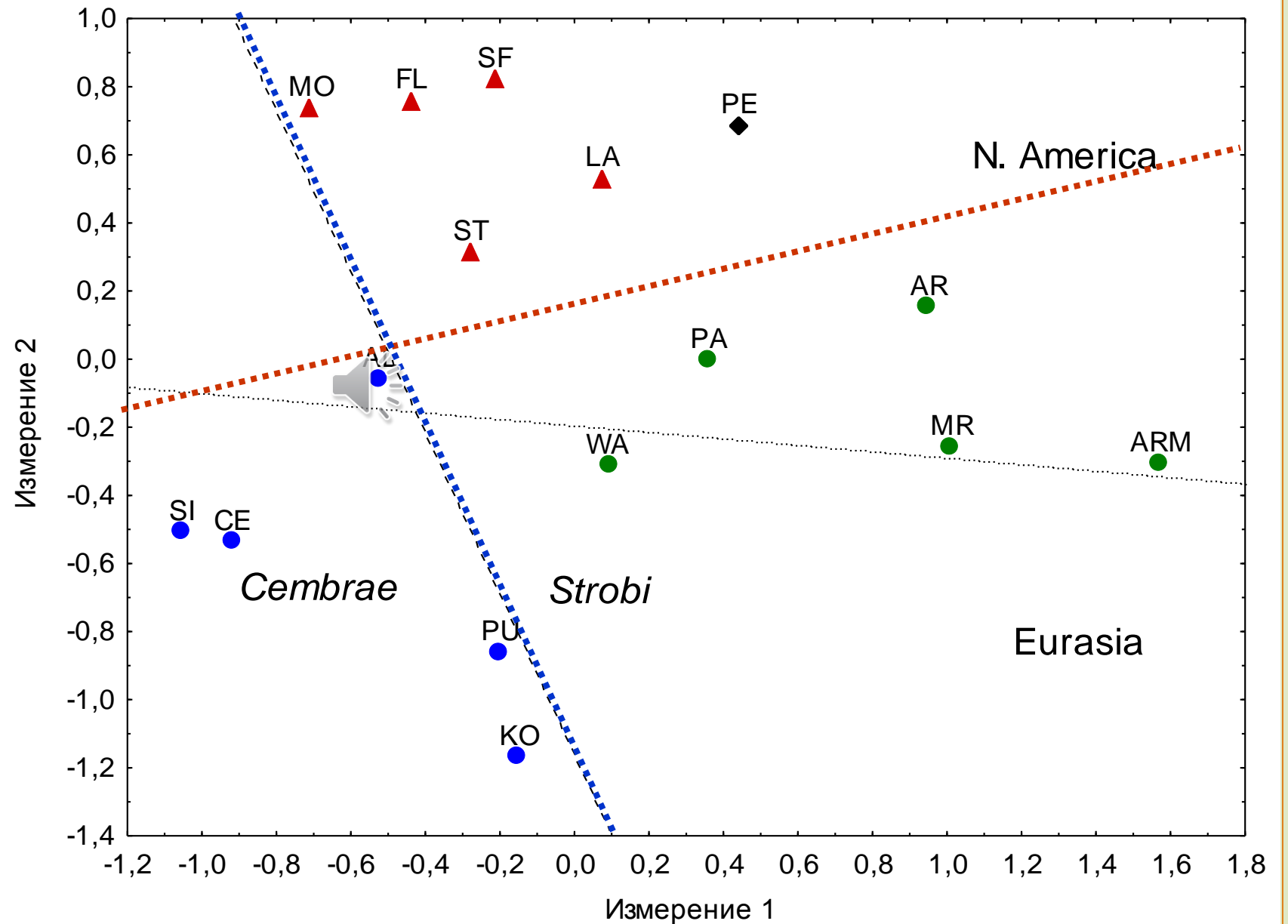
(Сосна обыкновенная, аллозимы, из Петрова и др., 2017)



**Рис. 2.** Дендрограмма генетических дистанций Неи (Nei, 1978;  $DN_{78}$ ) между популяциями *P. sylvestris*. Шифры выборок см. в табл. 1.



Неметрическое многомерное шкалирование по матрице  $D_N$  для видов секции *Strobus* (17 аллозимных локусов)  
(Белоконь и др., 1998; Politov, Krutovskii, 2004, Belokon, 2007; Politov, 2007)

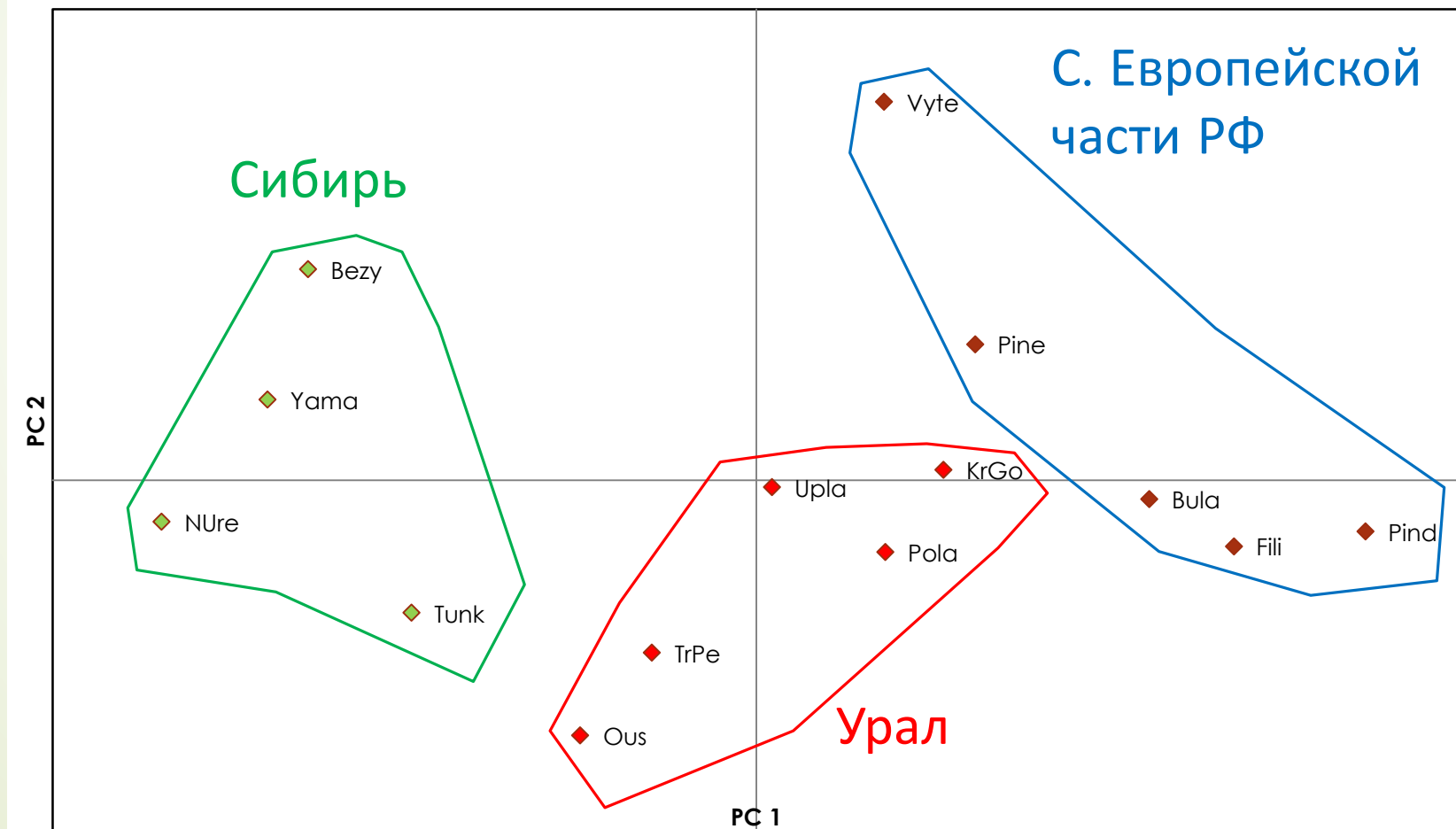




# Многомерный анализ выборок одного вида

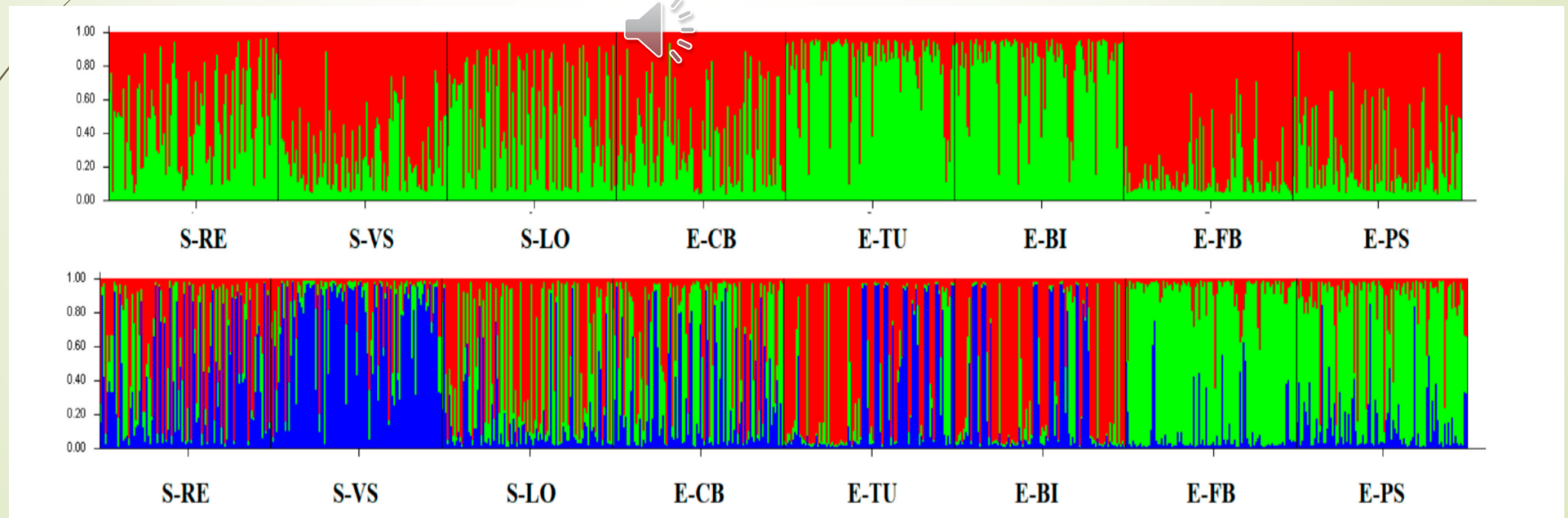
(Ели европейская и сибирская, SNP/CAPS, Grechnevikova et al., 2021)

Главные координаты (PCoA)



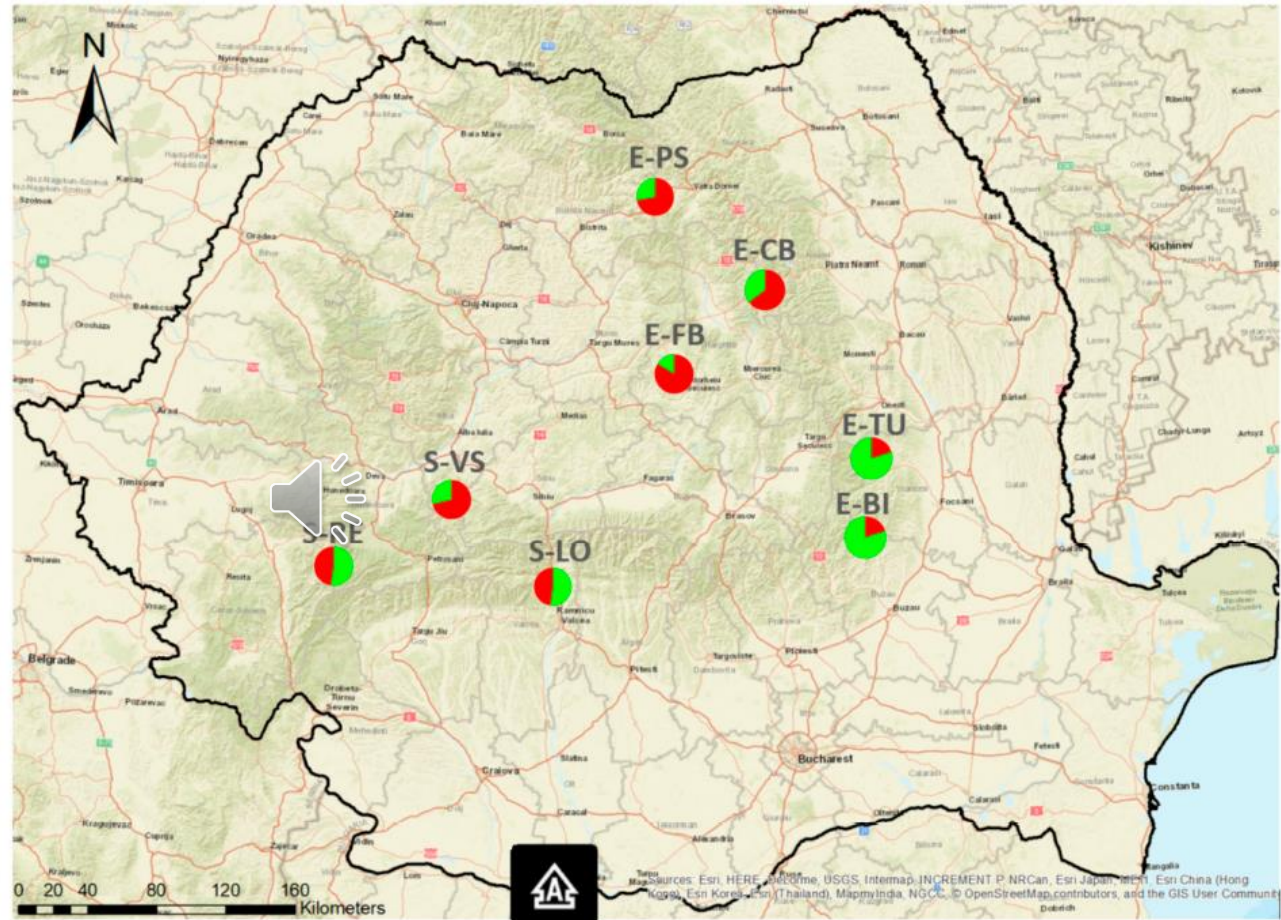
# Анализ с помощью байесовской кластеризации в программах семейства Structure (Pritchard et al., 2000)

Реконструкция доли участия «исходных кластеров» в мультилокусном генотипе каждой особи и в отдельных популяциях (сосна обыкновенная, *Pinus sylvestris*, 8 микросателлитных локусов, from **Sofletea et al., Forests 2020, 11, 1047**)



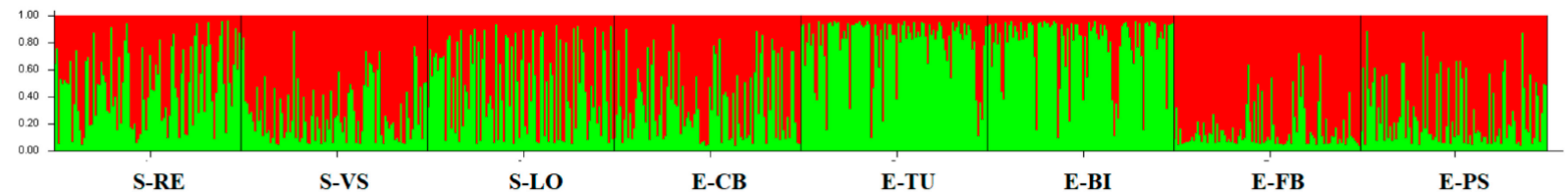
# Анализ с помощью байесовской кластеризации в программах семейства Structure (Pritchard et al., 2000)

Пространственное  
распределение  
частот  
исходных  
кластеров,  $K=2$



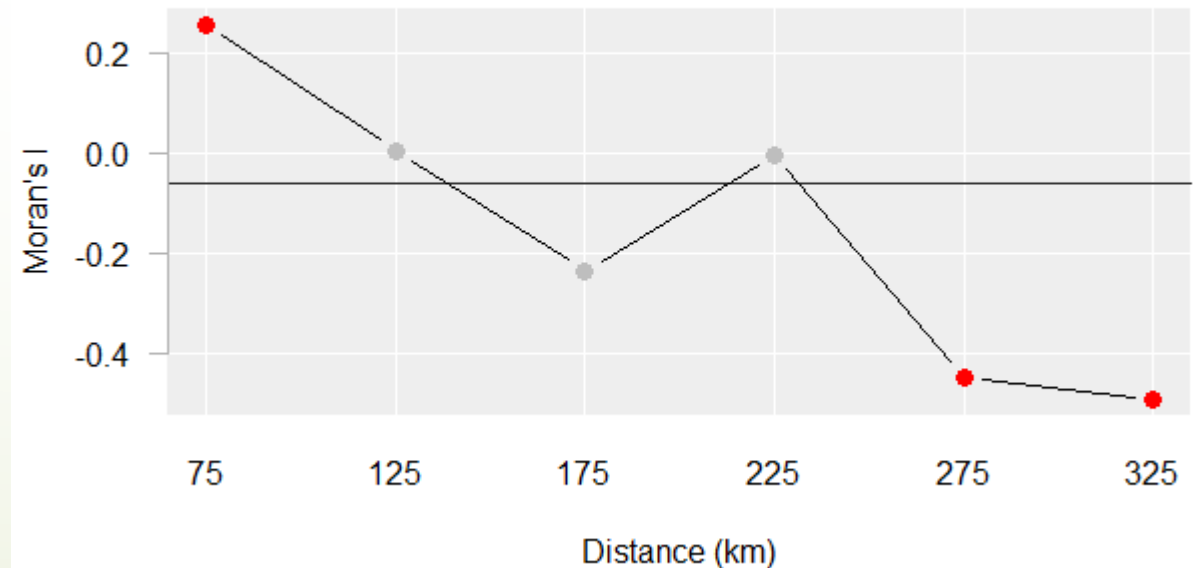
Sofletea et al., Forests 2020, 11, 1047

*Pinus sylvestris*



# Анализ пространственной автокорреляции, SAA

- ▶ Анализ пространственной автокорреляции (Spatial AutoCorrelation, SAA) - способен выявлять ассоциации между генетической структурой популяции, представленной, например, особями в системе географических или декартовых координат и матрицей генетических расстояний между ними.
- ▶ Выявление систематической пространственной структуры популяции фактически означает более близкое родство соседей, что может отражать и семейную (демовую) структуру как индикатор того или иного типа системы скрещивания, ограниченного радиуса разноса гамет и зигот, так и в перспективе потенциал для инбридинга.
- ▶ SAA это корреляция между значениями переменной и пространственными координатами точек, позитивная корреляция показывает группировку генетически более родственных особей вместе в пространстве, то есть отклонения от случайного распределения генотипов (Getis, 2008; Diniz et al., 2003).





# Цифровое генетическое картографирование

- Суть метода в отображении на карте значений признака, это могут быть как
  - частоты аллелей, генотипов, гаплотипов и т.д., так и рассчитанные показатели генетического разнообразия:
  - аллельное богатство  $A$ ,  $A_e$ ,
  - наблюдаемая  $H_o$  и ожидаемая  $H_E$  гетерозиготность
  - интегральные характеристики (главные компоненты РС, главные координаты РС<sub>0</sub>, частоты реконструированных кластеров семейства программ *STRUCTURE* и др.).
  
- «Тепловая карта» (heatmap) строится по опорным точкам с эмпирически оцененными значениями признака с помощью интерполяции.

# Цифровое генетическое картографирование


- ❑ Долгое время, начиная с А.С. Серебровского – автора учения о генофондах, генетические данные не визуализировались на картах, оставаясь в рамках таблиц, аналитической математики и статистики.
- ❑ 1970-х гг. с развитием компьютерного моделирования на основе баз данных о частотах аллелей и генотипов (Cavalli-Sforza, 1966) стали появляться первые геногеографические карты для генов человека (Bodmer et al., 1976).
- ❑ За ними последовали «синтетические» карты (Menozzi et al., 1978), представляющие данные анализа многофакторного анализа десятков генов (метод главных компонент).
- ❑ Отечественная школа цифровой картографии (Балановская, Нурбаев, 1995) активно использовала алгоритмы интерполяции, разработанные в геодезии и картографии (Сербенюк и др., 1990, 1991; Кошель, Мусин, 2000).
- ❑ [Рычков Ю.Г., Жукова О.В., Шереметьева В.А., Брук С.И., Афанасьева И.С., Балановская Е.В., Тихомирова Е.В., Сигнеев В.И., Лебедева И.А., Шнейдер Ю.В., Петрищев В.Н., Раутиан Г.С., Назарова А.Ф., Сыскова Н.Н., Бородина С.Р., Удина И.Г. Генофонд и география народонаселения. 2000. М.: Наука.](#)

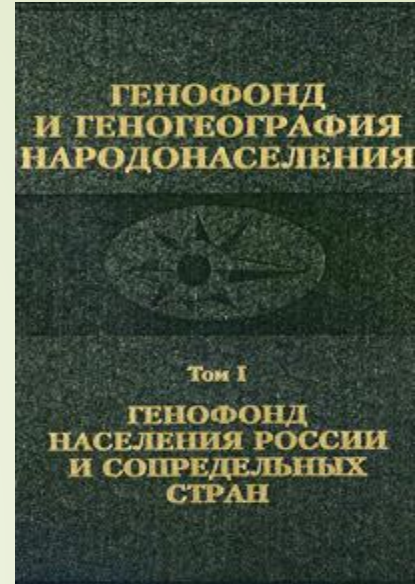
РФФИ грант 97-04-62064

Представлена исчерпывающая сводка имеющихся в литературе и новых данных по распространению частот генов и фенотипов по всем изученным в народонаселении России и сопредельных стран генетическим маркерам человека. Вся накопленная за многие десятилетия генетическая информация о народах России и сопредельных стран - от первых определений групп крови АВ0 и цветовой слепоты, осуществленных в начале XX в., до современных данных о многообразии эритроцитарных и лейкоцитарных антигенов, ферментов и других белков крови вплоть до первых характеристик полиморфизма ДНК в населении страны - обобщена в предлагаемом томе в форме, облегчающей использование этих данных в генетике человека, антропологии, медицине и в других смежных науках.



# Цифровое генетическое картографирование

- ❑ Теория и практические основы построения геногеографических карт - Кошель С.М. Геоинформационные технологии в геногеографии // Современная географическая картография /под ред. И.К.Лурье и В.И.Кравцовой. — М: Дата+, 2012. — С. 158–166.
- ❑ Программное обеспечение для построения геногеографических карт:
- ❑ Пакет GG MAG (на основе MAG  Сербенюк и др., 1991; Кошель, 2000), разработанный под ОС MS DOS и совместимый с международной разработкой ПО Surfer.
- ❑ Балановская Е. В., Балановский О. П. Русский генофонд на Русской равнине. — М.: ООО «Луч», 2007. — 416 с.

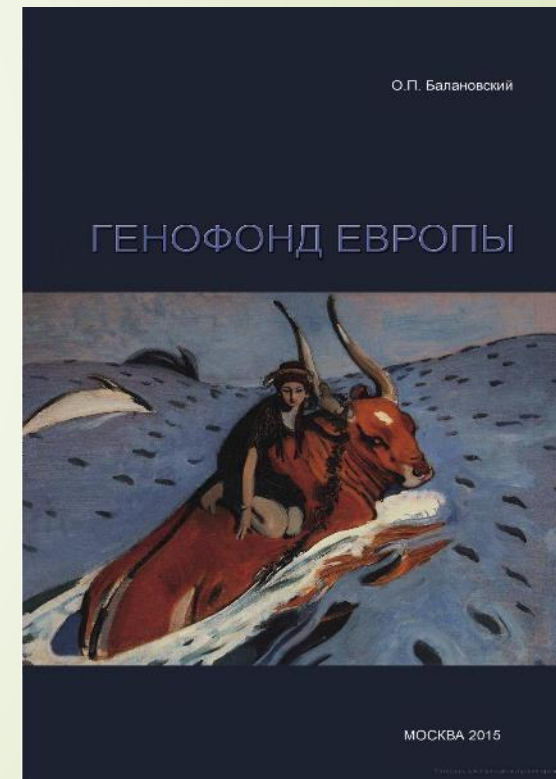
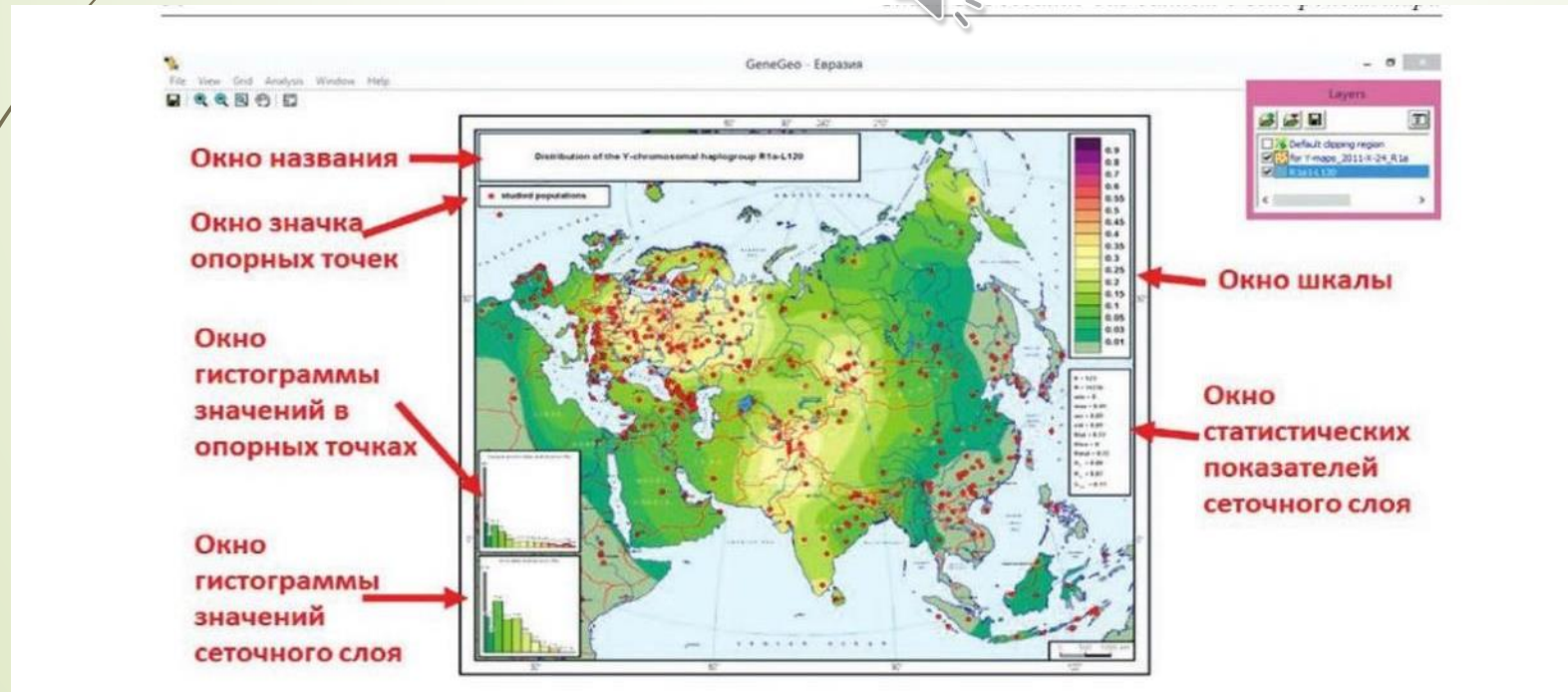




# Цифровое генетическое картографирование

□ Пакет GenGeo (Кошель, 2012; Балановский, 2015) под MS Windows - не только средство визуализации в виде «тепловых карт», в первую очередь - средство анализа пространственно распределённых, в основном частотных, данных по трёхмерной цифровой модели (ЦМ) карты:

- - блок моделирования
- - блок пространственного и статистического анализа
- - картографический блок





# Цифровое генетическое картографирование

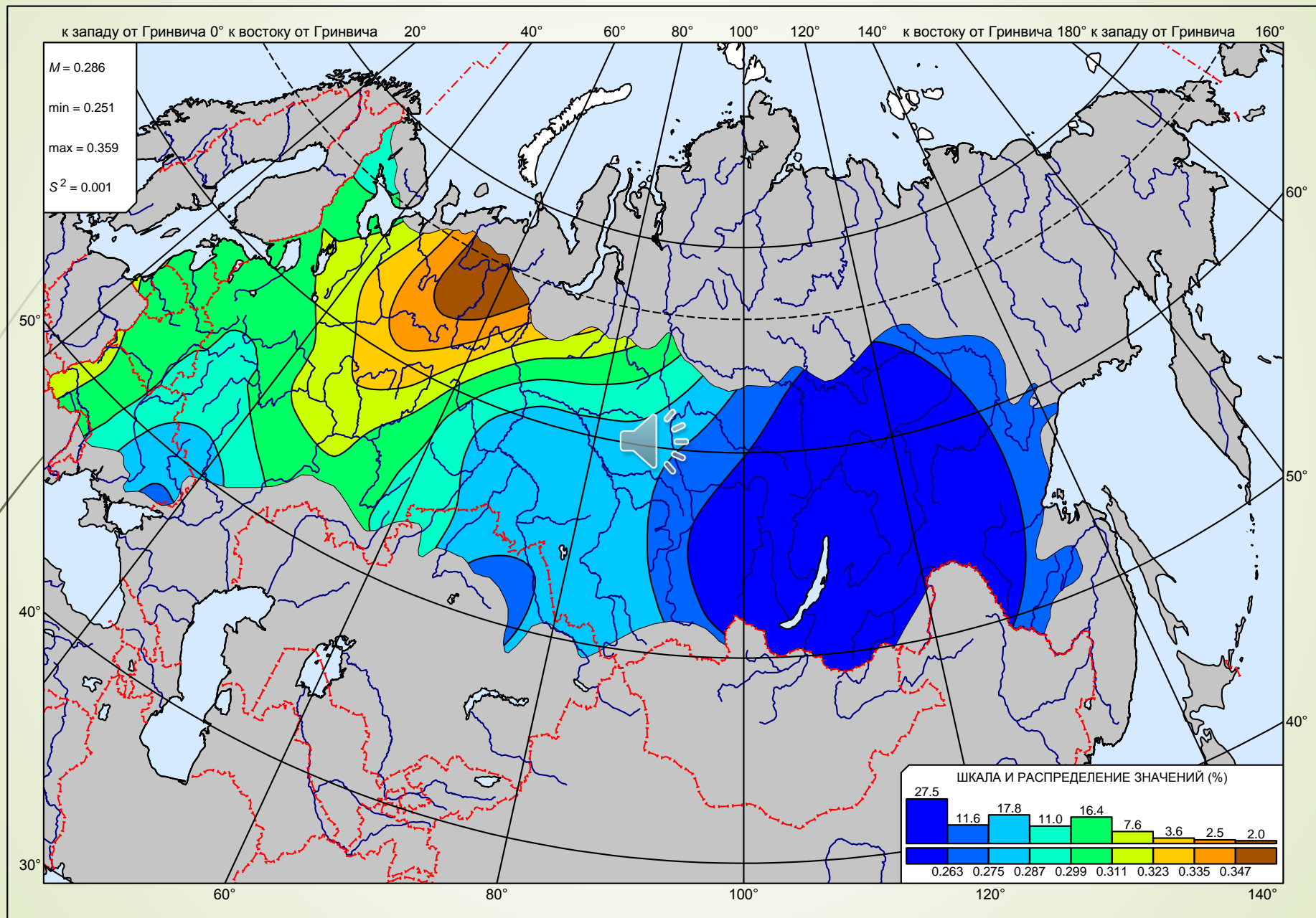
- Из современных программных продуктов следует отметить также **MD-Studio** (Евсюков и др. 1997, *Генофонд и генгеография...* 2003), для которого есть возможность прорисовки контуров видовых ареалов (функция, невостробованная в анализе человеческих популяций, но абсолютно необходимая в лесной генетике, где ареалы видов отличаются от контуров континентов и островов)

*Евсюков А.Н.* Материалы и методы генетического картографирования // *Генофонд и геногеография народонаселения* / Под ред. Ю.Г. Рычкова: Том II. Геногеографический атлас населения России и сопредельных стран. С-Пб, 2003. С. 24–38.



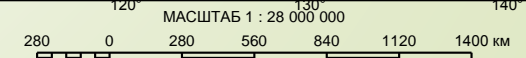
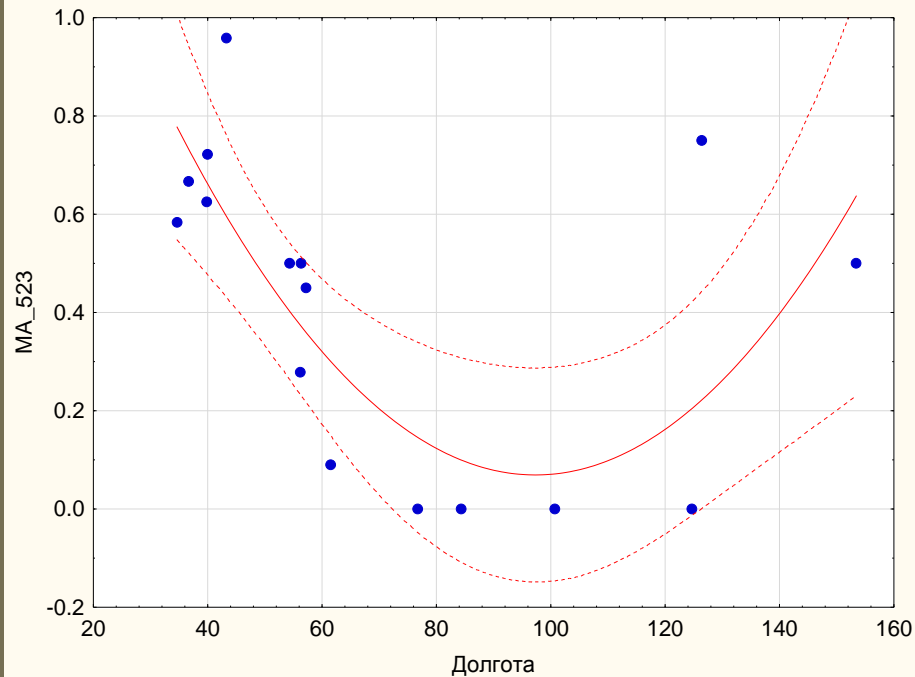
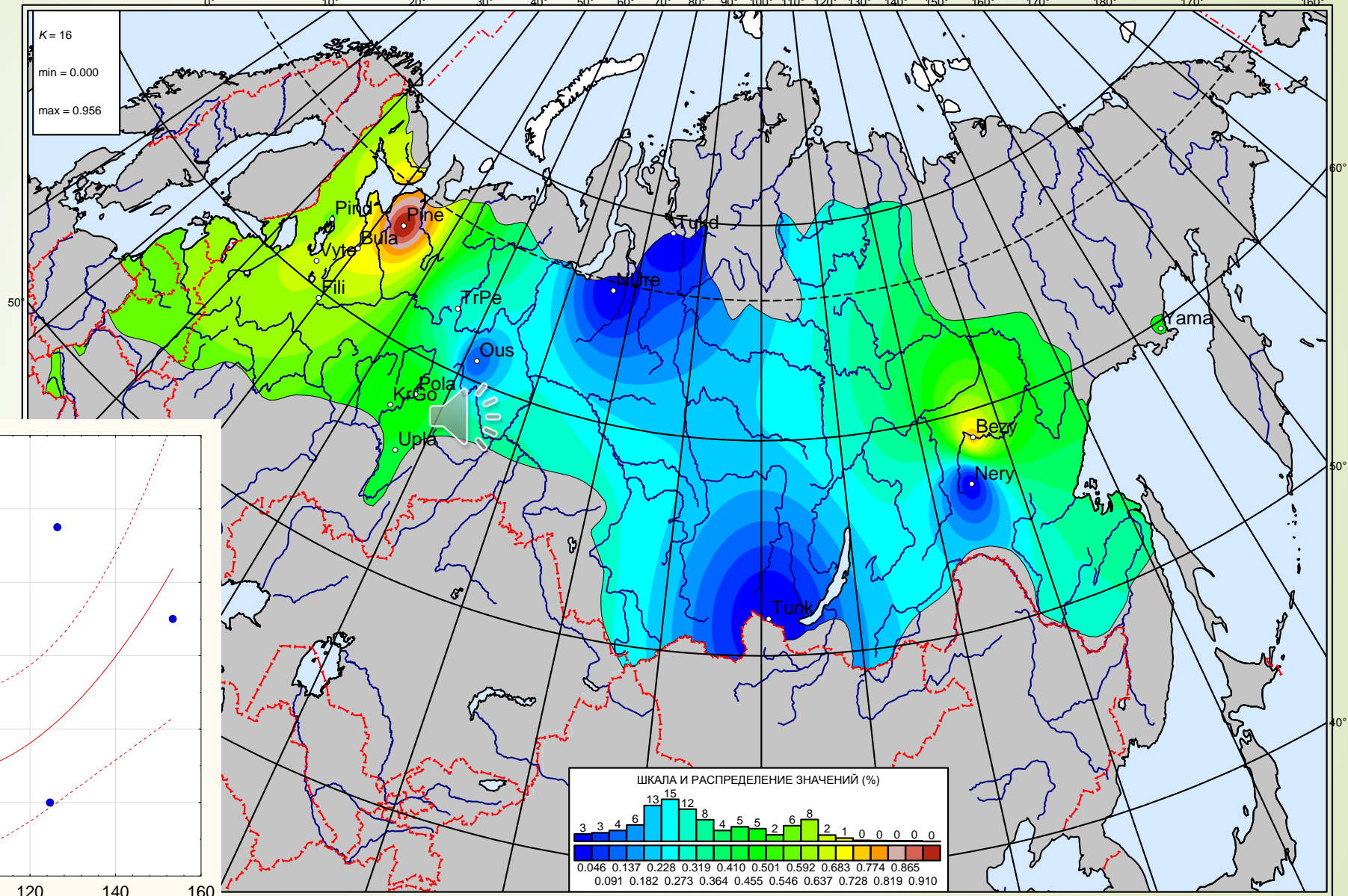
# Аллозимная гетерозиготность у *Pinus sylvestris*, результаты MD-Studio

## УРОВЕНЬ ГЕНЕТИЧЕСКОГО РАЗНООБРАЗИЯ



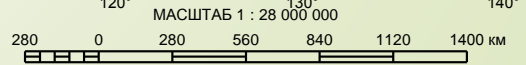
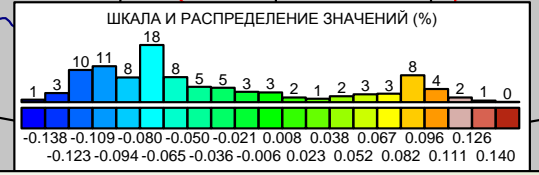
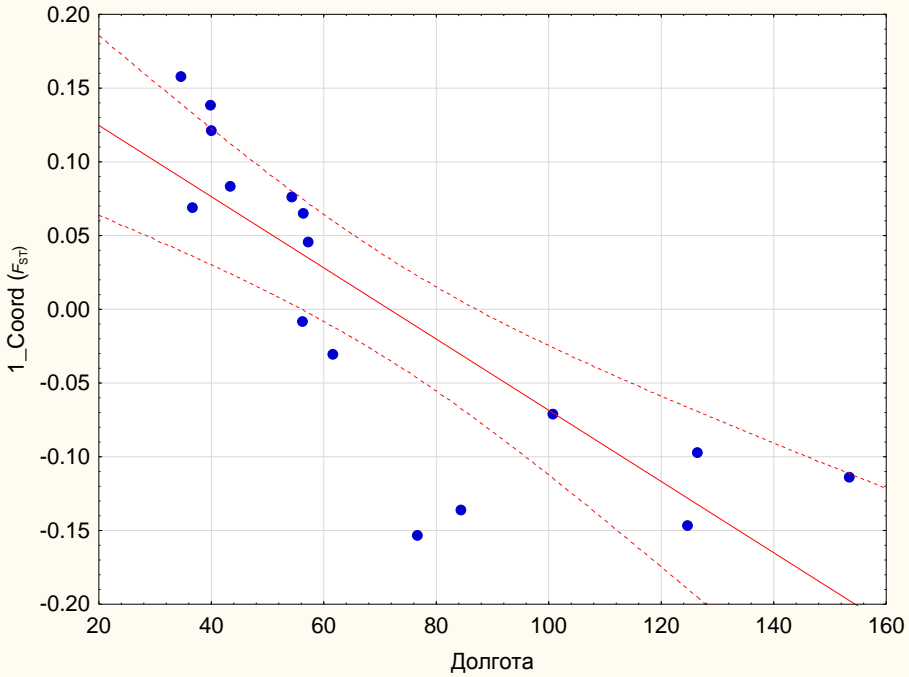
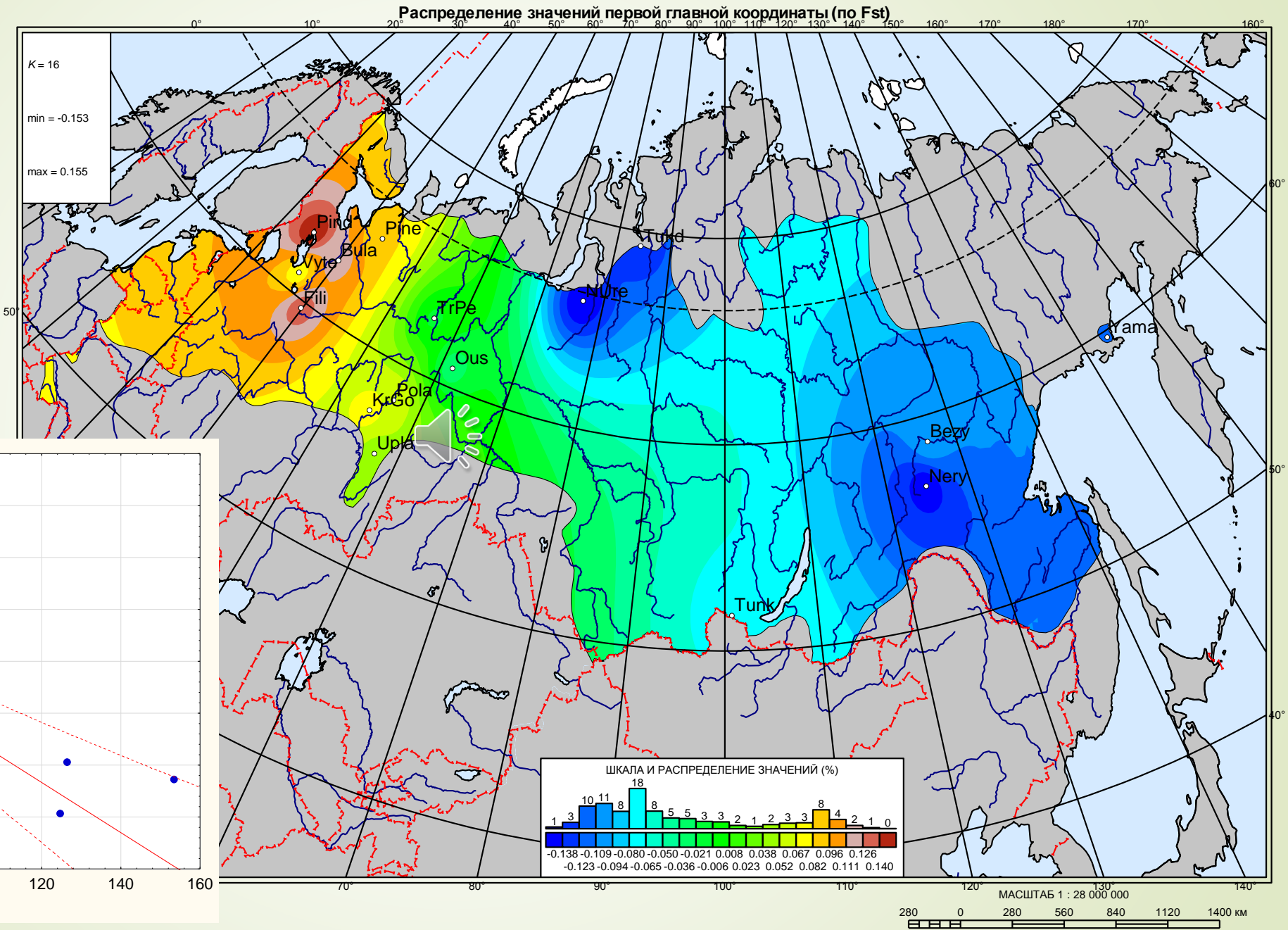
# Распределение частоты аллеля 1 снипа MA523 в популяциях ели

Распределение частоты отсутствия сайта рестрикции по локусу MA\_523



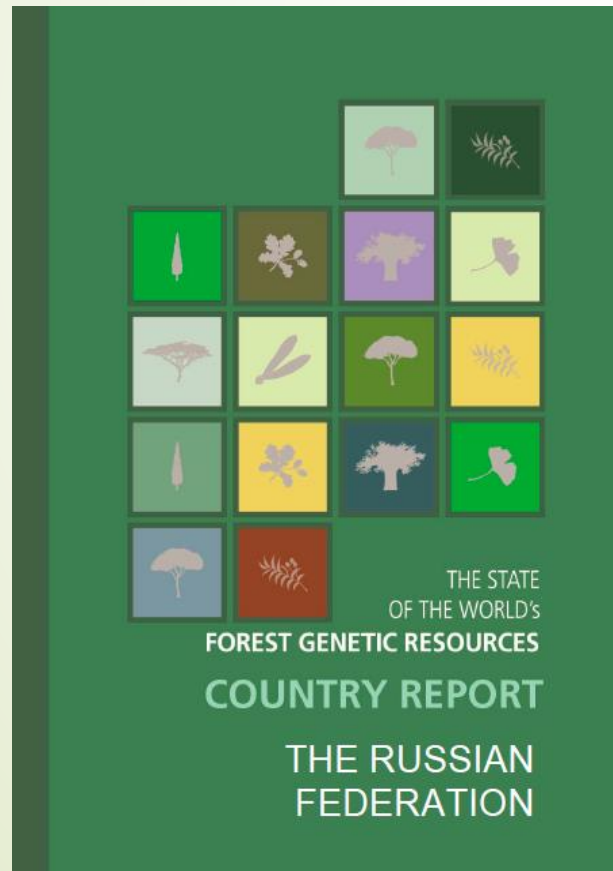


# Распределение 1й главной координаты по 9 SNP в популяциях ели





# Применение геногеографических и генокартографических технологий в лесной науке и лесном хозяйстве России



Страновой доклад РФ ФАО (The State of the World's Forest Genetic Resources Country Report. The Russian Federation, 2014).

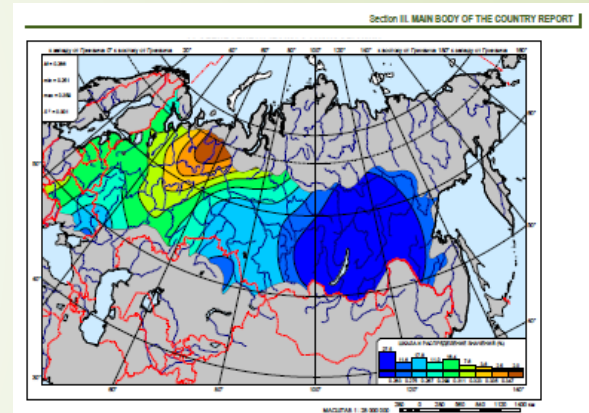


Figure III.1.1. Distribution of the average expected heterozygosity over the area of common pine

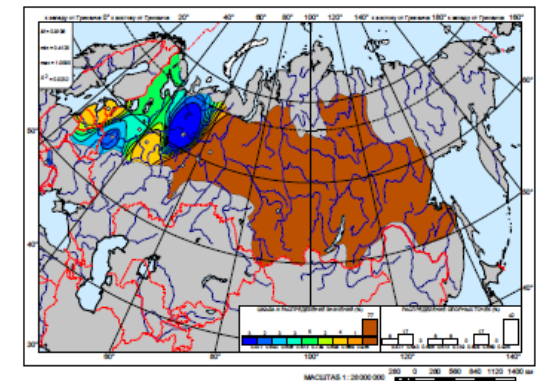


Figure III.1.2. Haplotype frequency distribution by *nad1* mtDNA locus of Norway and Siberian spruce

# Заключение

- ❑ Цифровое генетического картографирование – высокоэффективный метод анализа и визуализации генофондов
- ❑ Эти методы необходимо применять в лесной генетике и практике лесного хозяйства
- ❑ Необходимо понимание важности применения этих методов на всех уровнях
- ❑ ... и, конечно, финансирование для получения исходных молекулярно-генетических данных и создания геногеографических атласов.





Спасибо за внимание!

