

Вадим Веренич

*Актуальные проблемы изучения так
называемого динарского кластера
гаплогруппы I2a*

Исследовательские заметки о происхождении и миграции так называемого «динарского кластера» гаплогруппы I2a

Редакторский вариант: 20.01.2015

Гаплогруппа I2 (общие сведения и обзор публикаций).	3
Структура гаплогруппы I: структура SNP-дерева	4
Описание «динарской» субклады гаплогруппы I2a1b2a1-L147.2/CTS5966	6
Генеогеографический анализ	8
Распространение гаплогруппы I2 в Восточной Европе и России	9
Частоты распространения и разнообразие гаплогруппы I2 в Восточной Европе.	11
Филогенетический анализ	12
Графо-филогенетический анализ	12
Филогенетический анализ в программе Fluxus Network	15
TNT-дерево и возраст по данным rho-статистики	16
Картографический анализ	18
Оценка возраста	20
Анализ генетической изменчивости	23

Гаплогруппа I2 (общие сведения и обзор публикаций).

Гаплогруппа I2 человеческой Y-хромосомы маркируется принятыми в современной номенклатуре Y-хромосомного дерева Международной организации генетической генеалогии [ISOGG-2014] снипами L68/PF3781/S329, M438/P215/PF3853/S31. В настоящий момент есть все основания полагать с ближайшей родственной гаплогруппой I1, является одной из старейших европейских гаплогрупп, и в настоящее время встречается практически исключительно в Европе, либо в тех местах, которые были заселены выходцами из Европы. Вместе с гаплогруппой I1, она охватывает примерно 20% современного мужского населения Европы.

По расчетам популяционных генетиков, возраст этой гаплогруппы достигает примерно 20 000 лет; так к примеру, возраст популяции гаплогруппы I, вычисленный на основании расчетов дисперсии значений STR-локусов Y-хромосом гаплогруппы I оценивался в $24\ 000 \pm 7100$ лет, полученный возраст очень близок ко времени дивергенции населения Европы в палеолите (23000 ± 7700 лет до настоящего времени). Поэтому есть веские основания полагать, что эта гаплогруппа присутствовала в генофонде европейцев уже как минимум со времен финальной фазы палеолита и мезолита, возникнув из пула европейских Y-хромосом не ранее средне-верхнего палеолита, и не позднее последнего максимума оледенения, т.е. нижнего предела предполагаемого возраста мутации M170 (Semino et al. 2000). Полученный интервал оценок времени возникновения гаплогруппы I позволяет выдвинуть предварительное предположение о том, что первоначальный этап распространения гаплогруппы в Европе может быть связана с процессом культурной диффузии общеевропейского комплекса граветтских технологий примерно 28,000-23,000 лет назад (Djindjian 2000; Perles, 2000). С точки зрения генографии, географическое распространение гаплогруппы I в Европе носит бимодальный характер с двумя ярко выраженными эпицентрами в Скандинавии и на Балканах. Гаплогруппа I-M170, в самом обобщенном масштабе, т.е. вместе с дочерними субкладами, является преобладающей Y-хромосомной гаплогруппой на Балканах (усредненное значение по популяциям Югославии - 49%), в других европейских популяциях встречается с переменными частотами, более 40% у саамов, 10% в южноевропейских популяциях (греки, итальянцы) и в 20% на западе Европы (голландцы, французы). В данной работе мы не будем рассматривать гаплогруппу I1, которая была предметом отдельных исследований. Стоит отметить отличия от гаплогруппы I1, субклады (ветви) гаплогруппы I2 имеют более сложную и древнюю структуру. Поэтому в целях более детального и аккуратного анализа структуры гаплогруппы I2a, мы будем придерживаться иного сценария, который существенно отличается от сценария происхождения гаплогруппы I1, предложенного ранее в исследованиях посвященных генетической истории гаплогруппы I1. Частоты распространения гаплогруппы I2a в целом весьма информативны в плане изучения происхождения этой гаплогруппы. Согласно собранным исследователями STR-полиморфизмов статистическим данным о разнообразии Y-хромосомы в европейском генофонде, гаплогруппа I2a преимущественно представлена в восточной и южной Европы. Эта гаплогруппа достигает наивысших частот в юго-восточной Европе - Хорватии (31%) и Боснии (40%), составляя при этом примерно 80% -90% от общего вклада всей гаплогруппы I в генофонд этих балканских стран. Распространение гаплогруппы внутри самих балканских стран также имеет определенную специфику: например, в Хорватии гаплогруппа I (в целом, то есть на уровне I-M170) встречается статистически значительно чаще на трех южных островах - Брач, Хвар и Корчула, - чем в материковой части Хорватии и северном острове Крк (Barac et al. 2003¹), т.е. эта гаплогруппа наблюдается гораздо чаще в восточных и южных районах Хорватии и в то же время также наиболее часто встречаемой гаплогруппой в Боснии и Герцеговине (Marjanovic et al. 2005).

Наряду с этим, в западной Европе гаплогруппа I2a представлена преимущественно субкладой I2a1-M26 (Semino et al. 2000; Bosch et al. 2001; Capelli et al. 2003; Maca-Meyer et al. 2003), встречающейся с относительно низкой частотой ($\leq 5\%$), за исключением Сардинии (41%), Кастилии (19%) (Flores et al. 2004), Беарне (8%), и в стране басков (6%). Благодаря высоким частотам распространения в Сардинии, эту субкладу (I2a-M26) иногда называют «сардинской», однако исследование древних образцов ДНК жителей неолитической Европы, свидетельствуют скорее всего о том, что первоначальным ареалом распространения I2a1-M26 в Европе был пиренейский регион (Гаскония и земля басков).

¹ В ранних работах динарский субклад определялся снип-мутацией (M423), которая была обнаружена в 2007 году (Underhill et al. 2007), таким образом она является первой из обнаруженных мутаций, которые определяют I2a1b. Снип-мутация M423 расположена на снип-мутационном дереве ниже SNP-мутации P37.2 (определяющей всю гаплогруппу I2a в целом). SNP-мутация M423 произошла перед обоими разделением ветвей субклад I2a1b 1 (островные кластеры Isles) и I2a1b2a1 («динарский» кластер). По оценке Нордведта, это произошло не позднее, чем 8000 лет назад. Традиционно считается, что основатель «динарского» субклада жил, вероятно, на Балканах или в бассейне реки Дунай. Возраст динарского кластера I2a1b2a1 первоначально оценивался в 3600 лет, а затем был скорректирован на 2550 года (в настоящее время постоянно пересчитывается). Исходя из этого наблюдения, коллектив популяционных генетиков, которые исследовали гаплогруппу I еще в 2004 году (Rootsi et al. 2004), предположил, что, как представители гаплогруппы I1, так и гаплогруппы I2a реколонизировали Европу из области балканского ледникового убежища (рефугиума) после окончания последнего ледникового периода. Спустя 10 лет, в свете новых полученных данных, эта модель вызывает определенные критические замечания. Тем не менее, в силу тех или иных причин, описанная выше модель по-прежнему пользуется популярностью у академических популяционных генетиков (ср. например Regueiro et al. 2012).

Генография субклад I2a носит нон-инклюзивный характер по отношению других субклад I2a, иными словами ареалы распространения отдельных субклад I2a практически не пересекаются. Например, субклад I2a1-M26 и парагруппа I2a* (последний с маргинальной частотой) сосуществуют западу от итальянских Апеннин, однако к востоку от Адриатики встречается преимущественно I2a1b-M423. Распространение I2a1b-M423 характеризуется резким градиентом частот: частоты Y-хромосом этой субклады резко уменьшаются по мере удаления к западу от Балкан – так, например, этот субклад практически отсутствует у итальянцев, немцев, французов, и швейцарцев. При движении на северо-восток уменьшение частот гораздо более плавное, I2a1b-M423 (точнее ее динарская субклада) сохраняет заметные частоты среди населения, говорящего на славянских языках.

Как уже отмечалось выше, в восточной Европе гаплогруппа I2 представлена практически исключительно своей субкладой (подгруппой) I2a1b, маркируемой мутацией M423. Заслуживает особого внимание то обстоятельство, что эта подгруппа, в свою очередь, на территории восточной Европы представлена практически исключительно своим «динарским STR-кластером» (который соответствует субкладу I2a1b2a1-L147.2/CTS5966 в номенклатуре ISOGG-2014). В динарский кластер входят гаплотипы Y-хромосом с характерным аллельным мотивом STR-локусов («динарский модельный гаплотип»: 16-24-11-11-13, определяется соответственно по 6 - локусом DYS19-390-391-392-393, соответственно)². Например, большинство гаплотипов (126/221) из хорватской выборки гаплотипов гаплогруппы I разделяют подобный STR-мотив (Barac et al. 2003).

В принятой академической популяционной генетикой модели высокая частота гаплогруппы I2a (в том числе и I2a-1b2a1-L147.2/CTS5966) на Балканах принимается как указание на автохтонное происхождение этой гаплогруппы на Балканах, а ныне живущие на Балканах представители этой гаплогруппы считаются потомками древнейшего населения этого региона. В дальнейшей части работы, мы покажем, что характерный градиент частотного распространения гаплотипов I2a1b2a1-L147.2/CTS5966 в совокупности с статистическим анализом разнообразия гаплотипов I2a1b2a1-L147.2/CTS5966 (в северо-восточной Европе разнообразие гаплотипов выше, чем на Балканах и юго-восточной Европе), свидетельствуют в пользу альтернативной модели. Согласно этой модели, распространение наиболее часто встречающегося в Европе динарского кластера гаплогруппы I2a1b2a1-L147.2/CTS5966 имело экспансивный характер, и было (предположительно) связано с распространением носителей славянской пражско-корчакской культуры с территории нулевой фазы этой культуры (в Полесье и Карпатах) в направлении Карпат, центральной Европы и Балкан в период славянского расселения в Европе (5-8 веках нашей эры). Согласно расчетам времени коалесценции гаплотипов «динарского кластера» (примерно 2000 лет до настоящего времени), начальный период экспансии этого кластера в Европе приходится на первые века нашей эры

Структура гаплогруппы I: структура SNP-дерева

По состоянию на начало 2014 года в структуре надежно выделяются следующие филогенетически значимые субклады (принятые в ISOGG-2014, цифро-буквенное обозначение субклад приведено в соответствии со стандартом ISOGG-2014). Из краткой аннотации и описания гаплогруппы видно, что каждая из которых имеет характерный географический ареал распространения, практически непересекающийся с ареалами распространения других субклад I2a. Второй важной особенностью представляется хронология возрастов субклад этой гаплогрупп – в представленной ниже таблице указан возраст генетического разнообразия, и большинство из этих возрастов попадает в интервал между 8000-2000 г. до нашего времени, то есть примерно в период между началом неолита в Европе и финалом железного века. Верхний потолок этого интервала перекрывается с возрастом генетического разнообразия всей гаплогруппы I2a-P37.2 в целом (8000±4000 лет), при этом считается что ветвь I2a-P37.2 выделилась из родительской гаплогруппы I2 еще в мезолитический период времени (10700±4800 лет до настоящего времени)(Rootsi et al. 2004):

<ul style="list-style-type: none"> I2a1b2a1 (CTS176/S2621, CTS1293/S2632, CTS1802/S2638, CTS5375/S2679, CTS5985/S2687, CTS7218/S2702, CTS8239/S2715, CTS8486/S2722, CTS11030/S2768, L178/S328⁺) – эта субклада включает в себя наиболее распространенную в Восточной Европе разновидность гаплогруппы I2a - «динарский кластер» I2a1b2a1a (I2a1b2a1a-L147.2/CTS5966), встречающийся в Центральной Европе кластер I2a1b2a1*: Возраст генетического разнообразия составляет примерно 2500-3000 лет (Nordtvedt 2014). К этой субкладке примыкает экзотическая субклада I2a-1b2a1-CTS4022+ CTS10936+ CTS11768⁺, (STR-кластер «Disles»), встречающийся как на континенте (Польша) так и на Британских островах
<ul style="list-style-type: none"> I2a1b1 (L161.1/S185) - не встречающийся практически нигде за пределами британских островов и бывших английских колоний кластер гаплотипов. Субклад разделяется на три больших кластера А, В, С, встречающиеся в специфических районах Британии, Шотландии и Ирландии. Возраст генетического разнообразия примерно 6000 лет (Nordtvedt 2014)
<ul style="list-style-type: none"> I2a1a2a2 (L880) – субклада встречающаяся с низкой частотой в северной Франции. . Возраст генетического разнообразия примерно 7000-8000 лет (Nordtvedt 2014)
<ul style="list-style-type: none"> I2a1* - реликтовая субклада, обнаруженная у небольшого количества жителей Франции.

² В последующем, этот динарский модельный гаплотип (сокращенно ДМГ) был дополнен энтузиастами-исследователями генетической генеалогии для использования в коммерческих панелях STR-маркеров (главным образом для различных панелей FTDNA – 12, 37, 64, 111 маркеров и т.д.). Например, в стандартной записи 111 маркерной панели FTDNA ДМГ значения маркеров следующие 13-24-15-10-12-15-11-13-12-13-11-30-17-8-10-11-11-25-15-20-29-12-14-14-15-10-10-21-21-14-12-18-18-34-35-12-10-11-8-15-16-7-11-10-8-11-9-12-22-22-16-10-12-12-12-8-10-22-20-13-14-10-13-11-11-12-11-31-14-8-15-11-27-19-12-11-11-12-11-9-12-11-10-11-12-31-11-12-22-15-11-10-22-15-18-11-25-16-11-14-23-12-21-18-12-15-17-9-11-11. Этот модельный гаплотип был обнаружен не только на Балканах, но и по всей Восточной Европе

<ul style="list-style-type: none"> I2a1a2a1a (L233/S183, L624) – «западная» или «альпийская» ветвь. . Возраст генетического разнообразия примерно 8000-6000 лет (Nordtvedt 2014). Встречается на западе и северо-западе Франции
<ul style="list-style-type: none"> I2a1a1 (L158/PF4073/S433, L159.1/S169.1, M26/PF4056) – субклада с высокой частотой распространения в Сардинии, составляет около 40 % всех мужских линий на Сардинии; кроме того, она также обнаружена с низкой или умеренной частотой в стране басков и на Иберийском полуострове, а также составила 1,6 % (1/64) в выборке албанцев, живущих в Македонии. Предполагается, что эта гаплогруппа распространилась в южной Европе в ходе миграций неолитического населения, что частично подтверждается результатами изучения древней ДНК. . Возраст генетического разнообразия примерно 18000-16000 лет (Nordtvedt 2014)
<ul style="list-style-type: none"> I2a2b (L38/S154, L39/S155, L40/S156, L65.1/S159.1, L272.3) - субклада, встречающаяся с незначительной частотой в германогворящих регионах (преимущественно в долине реки Рейн), где она могла присутствовать со времен мезолита. Эта субклада была обнаружена в 4 скелетных останках мужчин, найденных в пещере Лихтенштейн – археологическом памятнике бронзового века в центральной Германии, где также были найдены артефакты культуры полей погребальных урн (De Beule 2008, 2009a, 2009b). . Возраст генетического разнообразия примерно 5000 лет (Nordtvedt 2014).
<ul style="list-style-type: none"> I2a2a (L34/PF3857/S151, L36/S152, L59, L368, L622, M223, P219/PF3859/S24, P220/S119, P221/PF3858/S120, P222/PF3861/U250/S118, P223/PF3860/S117, Z77) – наиболее представленная из всех субклад I2 в мужском генофонде Западной Европы. Ареал распространения I2a2 коррелирует с историческими границами распространения носителей германских языков. Гаплогруппа I2a2 обнаружена более чем у 4 % населения лишь в Германии, Нидерландах, Бельгии, Дании, Англии (кроме Уэльса и Корнуолла), Шотландии, на южной оконечности Швеции и Норвегии, а также в провинциях Нормандия, Мэн, Анжу и Перш на северо-западе Франции, в Провансе на юго-востоке Франции, в исторических областях Италии – Тоскане, Умбрии и Лациуме; в Восточной Европе она встречается с крайне низкой частотой в Молдавии, в Рязанской области и в Мордовии. Возраст генетического разнообразия примерно 8000-9000 лет (Nordtvedt 2014).
<ul style="list-style-type: none"> I2b* (L416,L417,L418) – крайне незначительная по размерам (минорная) субклада I2, встречающаяся в малых количествах в Италии, Германии, Шотландии. Возраст генетического разнообразия примерно 3000-2000 лет (Nordtvedt 2014).
<ul style="list-style-type: none"> I2c* (L596, L597) - минорная субклада, возникающая, как предполагается в рейнском регионе Германии во времена мезолита. Возраст генетического разнообразия примерно 5000-4000 лет (Nordtvedt 2014).

Таблица 1. Таблица с кратким описанием основных субклад I2a (цифрно-буквенное обозначение субклад приведен в соответствии номенклатуре ISOGG второй половины 2014 года)

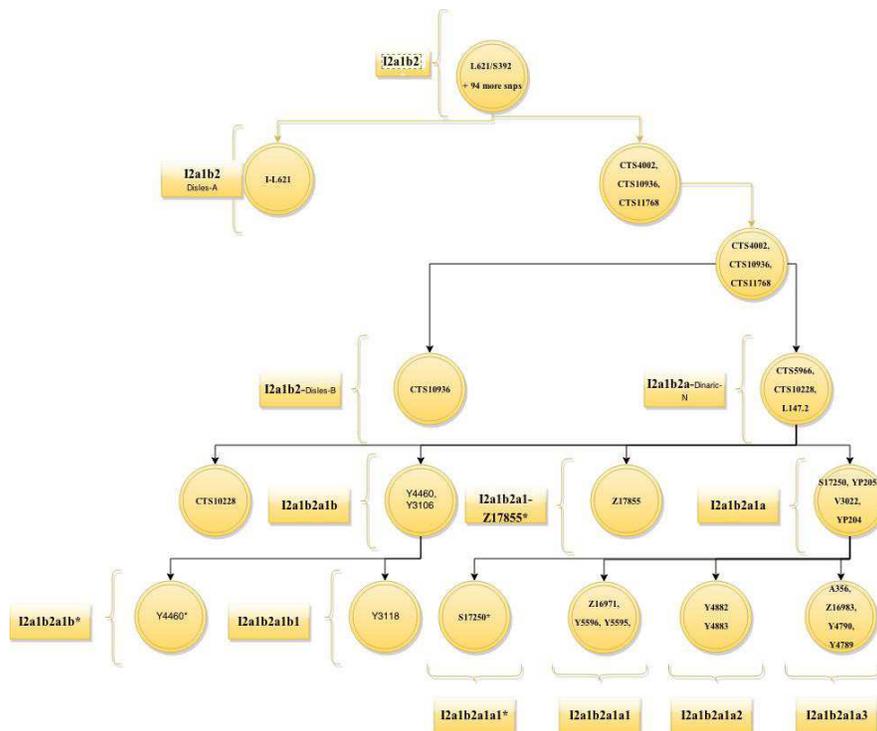


Рисунок 1. Схема SNP в структуре филогенетического древа I2a (версия Зденко Марковича из FTDNA проекта I2a, осень 2014 года) с сопровождающей аннотацией ISOGG-2014.

На схеме снипов (Рис. 1) хорошо заметна третья закономерности структуры субклад (подветвей) гаплогруппы I2a. Самые «древние» ветви (т.е субклады гаплогруппы близкие к корневому родительскому уровню дерева) географически локализируются в окрестностях Альп, западной Франции, долины Рейна и так далее. В этой связи находка древнейшего образца ДНК с Y-гаплогруппой I2a (возрастом в 8000 лет)³ в мезолитическом пещерном жилище Лошбур (Люксембург) может свидетельствовать не только о древнем присутствии этой гаплогруппы в регионе, но и о том, что процесс распространения I2a в начался где-то в конце мезолита из региона между Рейном и французскими Альпами.

Описание «динарской» субклады гаплогруппы I2a1b2a1-L147.2/CTS5966

В традиционной принятой в академической популяционной генетике «балканский» пик частоты распространения гаплотипов так называемой «динарской субклады»⁴ гаплогруппы I2a1b2a1-L147.2/CTS5966 объясняется непрерывной генетической приемственностью населения этого региона как минимум со времен мезолита, а то и неолита.

В 2009 году Кен Нордтведт предложил разделить этот субклад на 2 разновидности по характерным значениям 2 DYS-маркеров: северный динарский кластер Dinaric-N (преобладающий к северу от реки Дунай) и южный динарский кластер Dinaric-S, наиболее часто встречающийся на территории бывшей Югославии. По расчетам точкой коалесценции генетических линий, проведенными Нордтведтом в 2009 году. «северный динарский субклад» (время интеркладовой коалесценции 1900 лет до настоящего времени) старше «южного динарского субклада» (время интеркладовой коалесценции 1700 лет до настоящего времени).

Определяющие маркеры DYS (Y-STR локусы) в сравнении с другими субкладами I2a:

Таблица 2. Модальные значения DYS-локусов в двух условных субкластерах динарской субклады

DYS #	Dinaric-S	Dinaric-N
(1)	385a ...14 14
(2)	439 13 13
(3)	389ii .. 31 31
(4)	442 11 11
(5)	395b ...15 15
(6)	511 9 9
(7)	557 16 16
(8)	534 12 12
(9)	450 7 7
(10)	44410 10
(11)	48130 30
(12)	52021 21
(13)	5659 9
(14)	44116 16
(15)	45615 15
(16)	46112 12

³ Образец Лошбур (Loschbour) принадлежит к ветви, параллельной современным дочерним субкладам, которая отделилась от основной ветви не позднее чем 10 000 лет тому назад. Примечательно, что к этой исчезнувшей ветви принадлежат и несколько образцов ДНК жителей шведской мезолитической стоянки Motala (Motala 3, Motala 12), которая существовала синхронно Лошбуру.

⁴ Уже само условное обозначение субклады – «динарская» – носит условный характер. Название связано с «динарским модальным гаплотипом» описание которого приводится в старой статье (Varas et al.2003). Благодаря усилиям Нордтведта, название было популяризировано в среде любителей (генетических генеалогов). Название Dinaric представляет собой сокращение от «Динарские Альпы». Динарские Альпы расположены на севере Италии, расширяясь к югу вдоль всего восточного побережья Адриатического моря на расстояние примерно 400 км.

Наиболее важными отличающиеся значения DYS маркеры у I2a1b2a1-Dinaric-N и I2a1b2a1-Dinaric-S показаны ниже в таблице (Табл. 3), однако как показали последующие эксперименты с филогенией гаплотипов динарского субклада, это разбиение субклада на две группы не было надежным в силу высокой скорости мутации DYS449. ненадежность подобного разделения была окончательно доказана весной 2014 года, когда группа энтузиастов обнаружила первые

важные сипы, которые разделяют динарскую субкладу. Наиболее важным из этих сипов является S17250. На настоящий момент, у 5 представителей «северо-динарской» группы обнаружено предковое значение S17250 – и у 7-представителей «мутировавший» аллель (S17250 +), в то время как все 5 протестированных представителей южно-динарской группы имеют «мутировавший» аллель S17250 +⁵. Таким образом, более раннее разделение субклада I2a1b2a1 по 2 DYS-локусам не отражает истинной филогении: по состоянию на конец 2014 года субклад разделяется на 2-3 группы, однако в настоящий момент из-за недостатка статистических данных трудно сказать, какие из данных мутаций являются приватными (генеалогическими), а какие актуальны для этнопопуляционных построений (Рис. 2 и 3).

Таблица 3. Различающиеся значения в модальных гаплотипах Dinaric-N и Dinaric-S

DYS #.Dinaric-N ..Dinaric-S
448 ... 19 20
449 ... 30 32

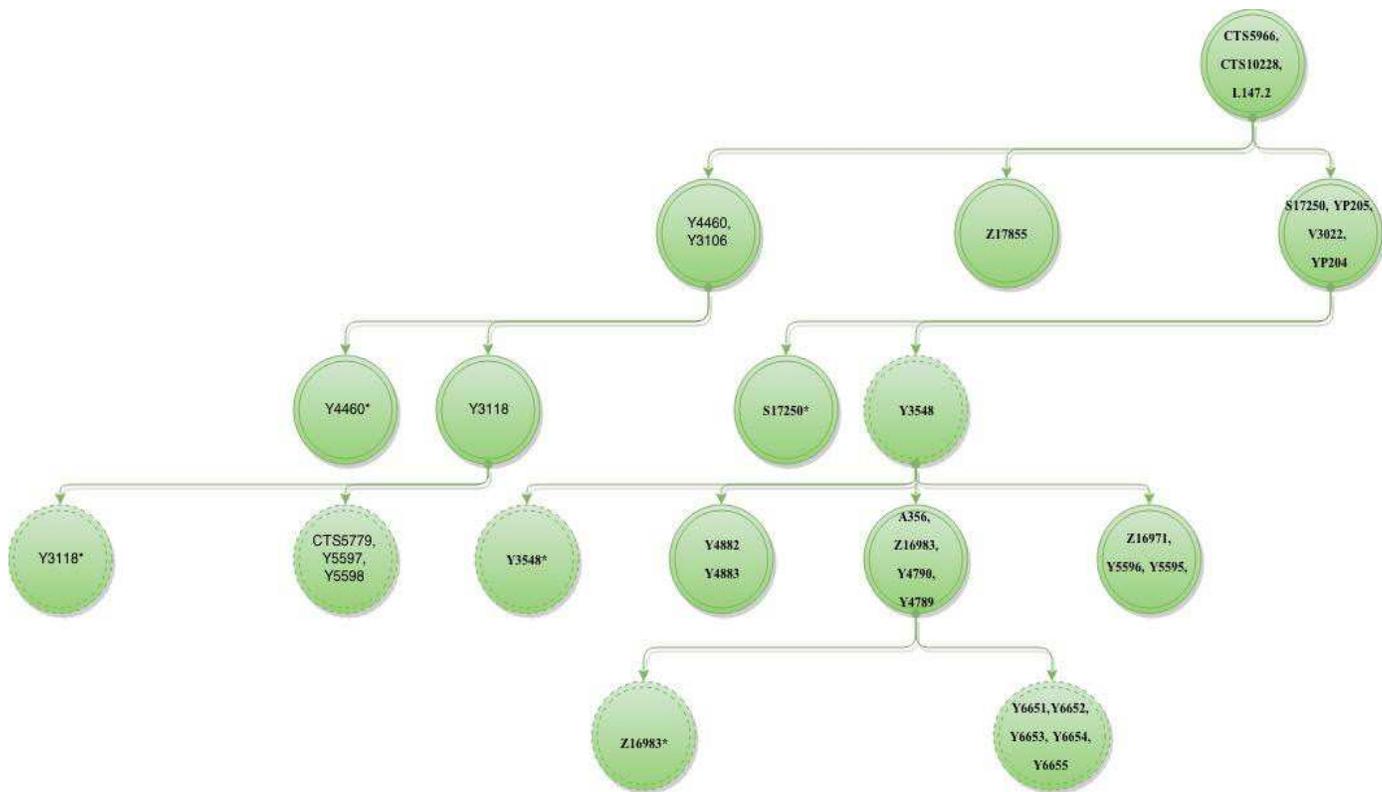


Рисунок 2. Сипы, обнаруженные внутри «динарской» субклада I2a1b2a1 (по данным Yfull, начало 2014 года)

⁵ Сообщено в личной переписке координатором проекта FTDNA I2a Лоуренсом Майкой (2011).

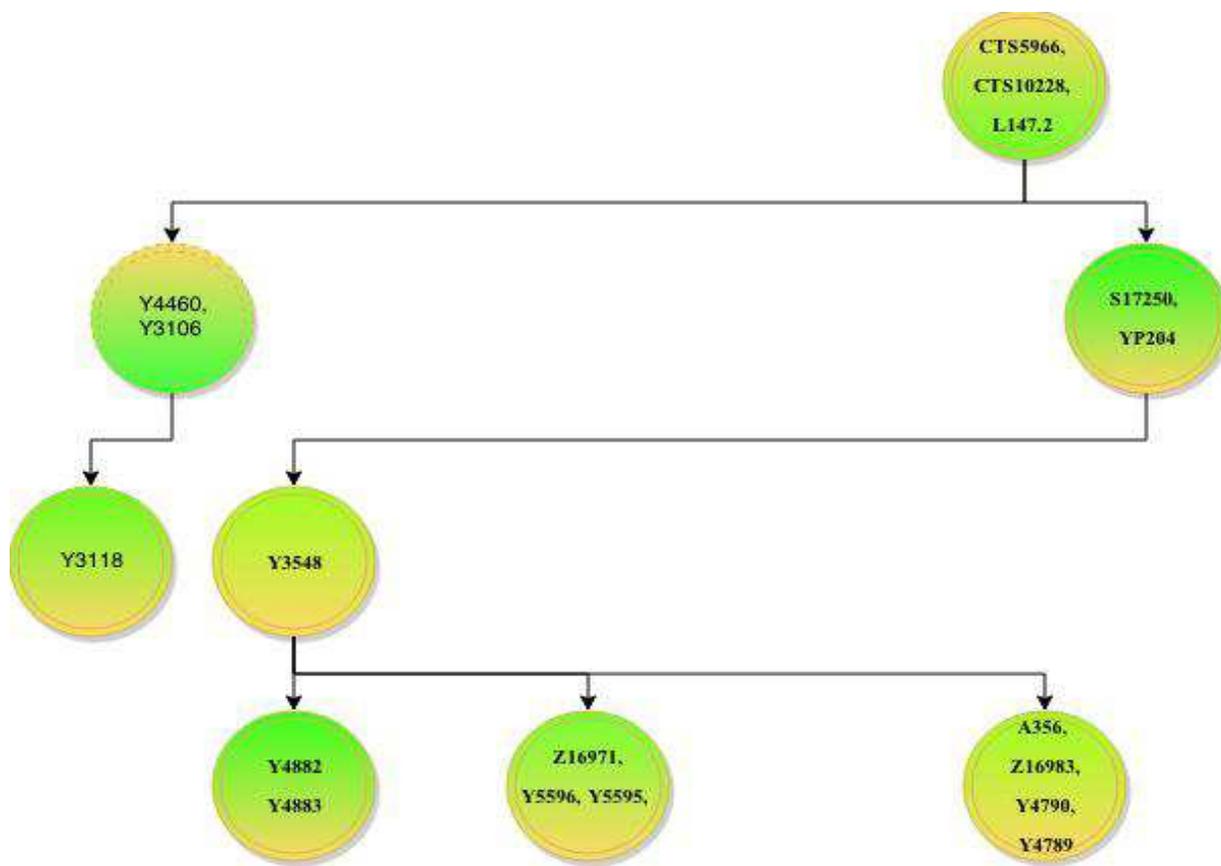


Рисунок 3. Снимки, обнаруженные внутри «динарской» субклады I2a1b2a1 (по данным Yfull, конец 2014 года). В силу того что лежащие ниже CTS5966 снимки (S17520, Y3106, Y3118, YP206, Y3548, Y4790, Y4882, Z16971) были обнаружены относительно недавно, в настоящее время нет никакой возможности сказать какие именно снимки имеют популяционную значимость, а какие снимки являются приватными. Обращает на себя внимание снимок Y3548, из-за введения которого топология экспериментальная версия древа субклад I2a1b2a1 (в версии Yfull) отличается .

Генеогеографический анализ

Для выявления устойчивых черт структуры генофонда, для объективного анализа закономерностей огромную помощь оказывает полисистемный подход. Его сущность - в параллельном анализе нескольких массивов разнородных данных с использованием одних и тех же методов. Современная генеогеография основывается на изучении и сравнении полиморфизма генома человека в разных популяциях людей. С помощью построения различных генетических карт она позволяет определять генетические расстояния между разными существующими на нашей планете группами людей и на этом основании вычислять время их разделения в ходе эволюции, то есть позволяет воспроизводить сюжеты из истории человечества. Картографический метод исследования в генеогеографии создает специальные карты, которые отражают результаты генетических процессов, возникающих в ходе формирования и развития народонаселения. Статистический и филогенетические методы дополняют и уточняют результаты применения картографического метода.

Распространение гаплогруппы I2 в Восточной Европе и России

В популяционной генетике регион Балкан традиционно рассматривается в качестве своего рода миграционного шлюза миграций с Ближнего Востока в Европу. В качестве примера используется «сводная» гаплогруппа IJ-M429, которая как считается, попала в Европу через Балканский полуостров незадолго до последнего ледникового максимума (Battaglia et al., 2009; Semino et al., 2000; Underhill et al., 2007). Одна из ветвей потомков носителей Y-хромосомной гаплогруппы IJ-M42, в последующем дала в Европе начало гаплогруппы I-M258, впоследствии разделившись на несколько «сестринских» гаплогрупп I1-M253, I2a-P37.2 и I2a2-M223 (название субклад данно по ISOGG-2014; в оригинальной статье Underhill et al. 2007 используются старые обозначения). Гаплогруппа I, как предполагается в профильной литературе, сыграла центральную роль в процессе заселения Европы из отдельных ледниковых «рефугиумов». Кроме того, предполагается что фаза экспансии гаплогрупп I1-M253 и I2a-P37.2 со столь различной филогеографии, началась не позднее периода раннего голоцена что соответствует началу европейского мезолита (Rootsi et al. 2004). Вычисленный с использованием 15 Y-STR локусов возраст гаплогрупп I-M258 ($n = 39$) была оценен в $20\ 000 \pm 4\ 000$ лет, полученная оценка возраста хорошо согласуется со временем дивергенции гаплогруппы I в Европе (Rootsi et al. 2004) – $23\ 000 \pm 7\ 700$ лет). С другой стороны данный интервал возраста хорошо укладывается во временный интервал распространения в Европе гравецких технологий (археологически начало распространение гравецких технологий датируется примерно 24 тыс. до настоящего времени, т.е верхним палеолитом, и распространение гравецких технологий по всей Европе связано с миграцией людей с Ближнего Востока, Анатолии на Балканы примерно ~ 28 000 - 22 000 лет тому назад.

В ранних работах, в которых использовались выборки гаплотипов с небольшим числом Y-STR локусов, приводились утверждения о высоком разнообразии гаплотипов I2a2-P37.2/M423 в Сербии (индекс разнообразия: $0,9977 \pm 0,0094$). Эти утверждения легли в основу гипотезы о том, что мутация P37.2 присутствовала на Балканах еще до последнего ледникового максимума (Rootsi et al. 2004; Semino et al., 2000). Кроме того, возраст гаплотипного разнообразия гаплогруппы I2a1b-M423 у балканских популяций составляет примерно 9000 лет назад, и соответствует, как уже сообщалось ранее, эпохе раннего голоцена или европейского мезолита (Battaglia et al., 2009; Pericic et al., 2005; Rootsi и др., 2004; Underhill et al., 2007). Ранние статьи о гаплогруппе I подтверждали заключение авторов статьи (Battaglia et al. 2009), о том что значительное присутствие «автохтонной» балканской гаплогруппы I2a1b-M423 в генофонде жителей Восточной Европы является свидетельством перехода потомков балканских охотники-собирателей к земледельческому образу жизни. Однако, в свете последних данных популяционной генетики, а также благодаря исследованиям энтузиастов, мы можем утверждать что эта предковая ветвь (субклада) I2a1b-M423 выделилась из родительской гаплогруппы I2a1 около 8000 лет назад в эпоху финального мезолита предположительно на территории северо-западной Европы, т.е в другом регионе Европы. В плане расчета времени расхождения дочерных субклад для I2a-P37.2 особых изменений нет – как и предполагалось ранее, оно произошло около $10,7 \pm 4,8$ тыс. лет назад (Rootsi et al. 2004). В ходе анализа древней ДНК жителей мезолитической северной и западной Европы выяснилось, что эта гаплогруппа была широко распространена в северо-западе Европы в эпоху мезолита (Lazaridis et al. 2014), от западной до северо-восточной Европы (мезолитические захоронения Лошбур, Люксембург и Мотала, Швеция). В основных своих чертах, Эта картина хорошо согласуется с более ранними выводами популяционных генетиков, согласно которым, начальная фаза распространение данной субклады произошло «не ранее голоценового перехода и не позднее раннего неолита» (Pericic et al. 2005), при этом возраст разнообразия для I2a1b M423 составляет $8,8 \pm 3,6$ тысяч лет (Underhill 2007). Из вышеприведенной выше информации следует, что благодаря находкам древней ДНК удалось уточнить место первоначального распространения гаплогруппы I2a1b M423 (северо-западная часть Европы), по-крайней мере одной из ее древнейших исчезнувших ветвей, выделившейся из родительской гаплогруппы примерно 8700-9000 лет назад (см. Рисунок 3., где желтым цветом обозначен общий кластер образца Loshbour и ближайшего к нему образца HGDP0887_Rus из панели референсной русской популяции в проекте HGDP). Видимо, в этом же регионе и произошло выделение предковой ветви I2a1b2a1-L147.2/CTS5966, так как ближайшая к динарской субкладе сестринская субклада I2a1b2a*-CTS54002 (так называемый Disles) встречается как в Польше, так и на британских островах, а следующая по удаленности клада (островная субклада – Isles- I2a1b1 (L161.1/S185)) встречается практически исключительно на британских островах.

В свежей работе (Gamba et al. 2014) опубликованы данные о самых ранних из обнаруженных в балканском регионе образцов древней ДНК с гаплогруппой I2: KO1 с неолитической стоянки Tiszaszöls-Domaháza (5780-5650 до нашей эры) и NE7 с неолитической стоянки Арс-Берекалжа (4360-4490 до нашей эры). В контексте этой статьи важно отметить, что оба образца (образец NE7-I2a2a-L1228 и образец KO1 – I2a-L68) принадлежат к параллельным ветвям, ни одна из которых не является предковой ветвью динарской субклады I2a1b2a1. Исходя из этого можно сделать осторожное предположение о том, что представители динарской субклады I2a1b2a1-L147.2/CTS5966 появились на Балканах гораздо позднее неолитического периода.

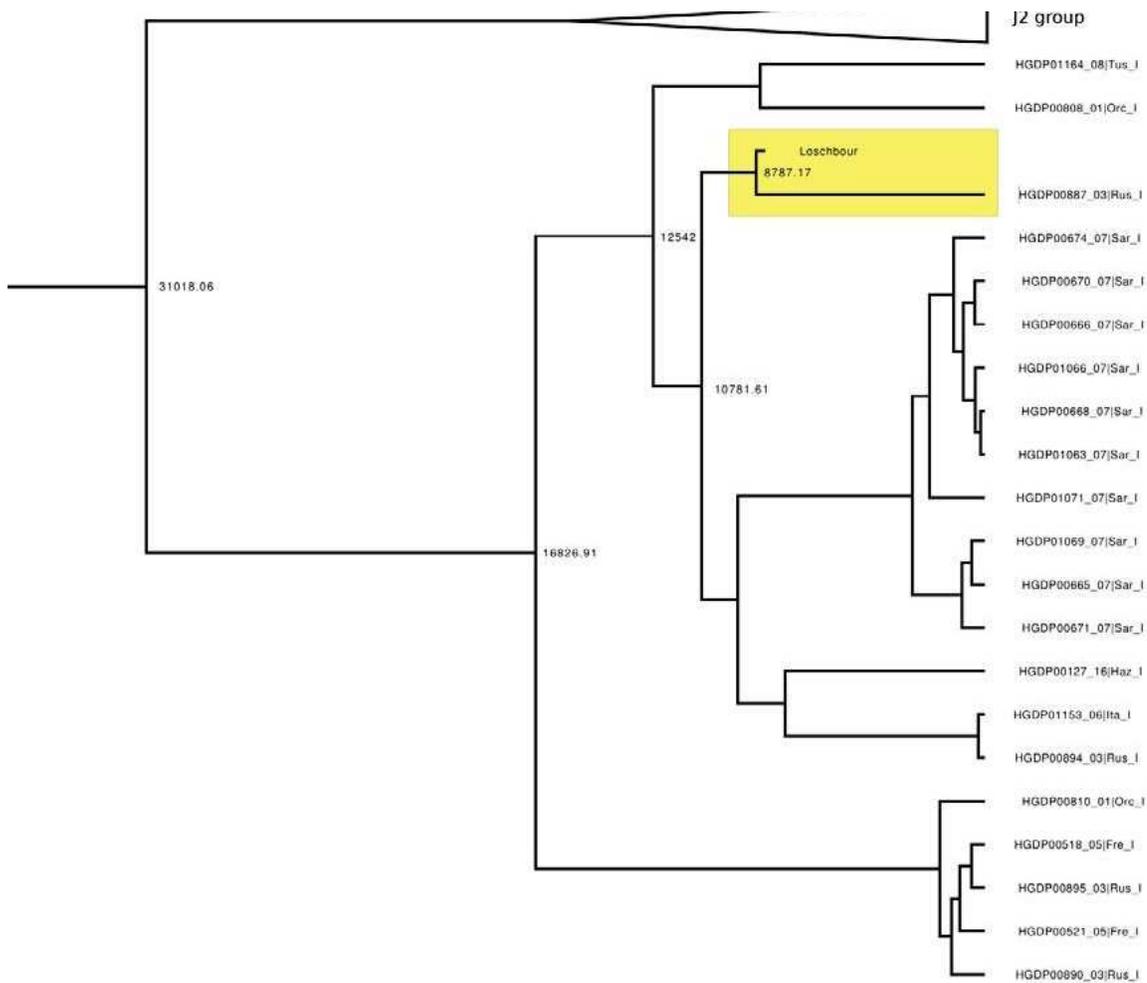


Рисунок 4. Положение древней мезолитической клыды I2a1b* на дереве гаплогруппы I2a (из работы Lazardis et al. 2014)

Частоты распространения и разнообразие гаплогруппы I2 в Восточной Европе.

Современные европейские частоты гаплогруппы I2 характеризуются резким градиентом с пиковым центром частот на Балканах, а также в районе между Карпатами и Полесьем: градиент уменьшения частот направлен в сторону северо-востока Европы. Из всех субклад родительской гаплогруппы в Восточной Европе наиболее представлена одна из разновидностей гаплогруппы I2a1b с характерным «динарским гаплотипным мотивом» («динарский кластер» -I2a-

1b2a1-L147.2/CTS5966, в короткой записи I2a-L147.2/CTS5966); эта разновидность наиболее часто встречается среди южнославянских народов, а также (в меньших количествах) среди румын, молдаван, украинцев, венгров, южных литовцев, албанцев, греков, жителей северо-востока Италии, в западной Анатолии и на Северном Кавказе.

Отметим основные характеристики распространения гаплогруппы I2a1b. Наибольшие частоты распространения I2a1b встречаются у жителей Боснии-Герцеговины и хорватов Далмации от 40 до 60%, у сербов и македонцев 20-30%, примерно столько же в Молдавии у гагаузов. В Украине частота распространения гаплогруппы I2a1b достигает 15-16%. В Беларуси регионы с наибольшей частотой – это западная ее часть (22%) и восточное Полесье (27%), в центральной части около 14-16%, и резкое снижение к северу страны, где частота падает до 7%. В России частотный градиент также направлен с юго-запада (Белгородская, Курская области 18-20%) на север с понижением почти до нуля. В центральных областях России I2a1b не имеет резких всплесков частот и держится на уровне в среднем 5-10%. В остальных европейских странах распространенность гаплогруппы I2a1b крайне низкая. В Скандинавии от 0 до 1%, в Чехии 5-8%, в Турции в среднем 1-2%. В Евразии и на Ближнем Востоке, а также в Африке гаплогруппа I2a1b (как и вся гаплогруппа I2a в целом) имеет нулевые частоты. Подобная картина распространения частот гаплогруппы характерна для случаев массовой экспансии (распространения) определенной генетической линии.

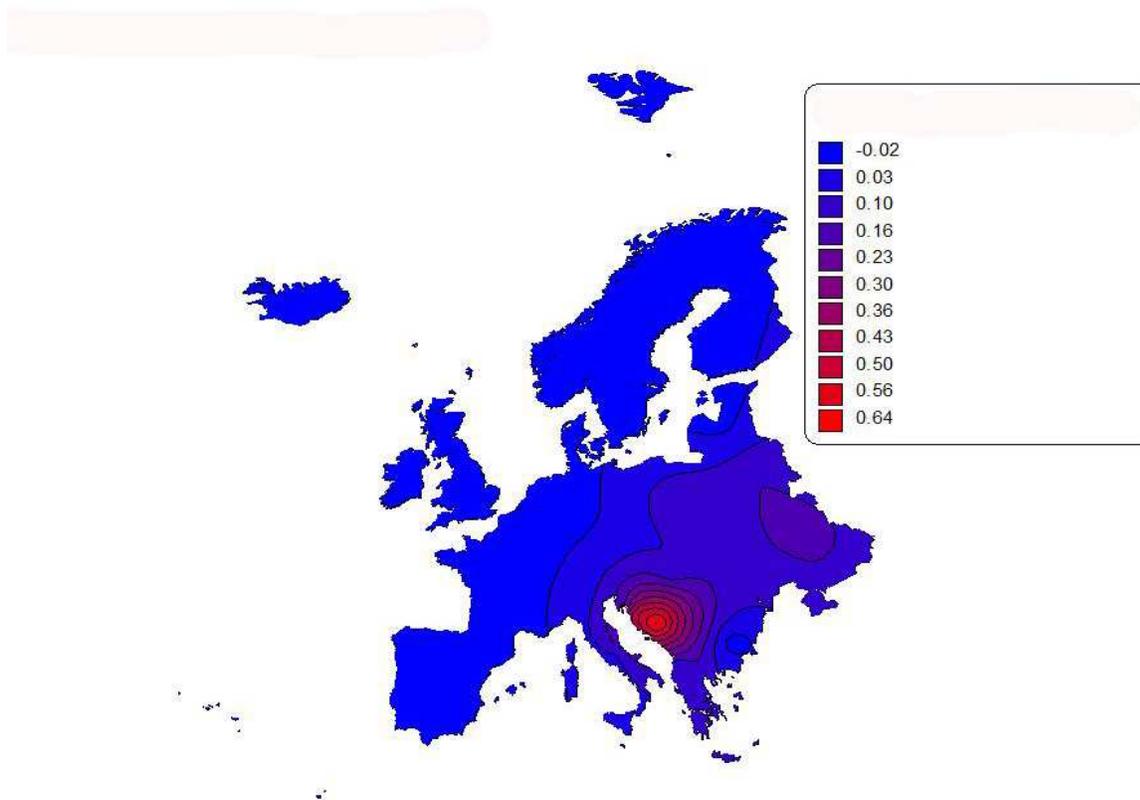


Рисунок 5. Картографическая визуализация частот распространения гаплогруппы I2a1b2a1 в Европе, с очевидным бимодальным пиком частот на Балканах и Карпатах (программа Golden MapViewer 7). При построении карты использовался kriging-алгоритм пространственной нормализации значений частот гаплогруппы I2a1b2a1

Филогенетический анализ

Наряду с стандартным картографическим методом, одним из наиболее адекватным и продуктивным подходом в изучении генетической структуры населения является филогенетический метод с использованием маркеров Y-хромосомы. Постоянно растущий уровень филогенетического разрешения, а также покрытие популяционными исследованиями все большего количества географических групп популяций

Графо-филогенетический анализ

Применение филогенетического анализа гаплотипов с большим количеством DYS-маркеров позволяет выяснить генетическую историю происхождения и по-новому взглянуть на пути миграции динарского кластера I2a1b. Существует альтернатива традиционному методу построения медианных сетей гаплотипов в Fluxus Network 4.6.1.1: эта альтернатива является единственно приемлимой в случаях, подобных нашему, то есть когда вместо филогенетической схемы гаплотипов нужно создать гаплотипные сети (haplotype networks), но из-за больших объемов данных построить их в стандартной поггенепрограмме Fluxus-Network в течении оптимального времени невозможно из-за большой вычислительной сложности задач. В этих случаях вместо медленного алгоритма редукции штейнеровских деревьев в Fluxus Network можно использовать комбинацию бесплатного программного обеспечения Arlequin (Excoffier, Lischer 2010), HapStar (Teacher, Griffiths 2011), Gephi (Bastian et al. 2009): эта комбинация **делает возможным** реконструировать возможные варианты эволюционно-филогенетической сети гаплотипов гаплогруппы. Используемый в программе Gephi алгоритм Force Atlas 2, позволяет разглядеть мелкие детали размещения отдельных гаплотипов, и в конечном итоге дает уникальную структуру графа, и эта структура существенным образом отличается от привычной структуры сети гаплотипов в стандартных вариантах программы Fluxus Network, которые публикуются в обычных статьях популяционных генетиков. Исходя из этого, мы ограничились применением другого алгоритма Force Atlas и последующей группировкой перекрывающихся узлов графа сети в одну группу. Благодаря этому приему, на выходе был получен более приемлимый, с точки зрения академического стандарта, графический вариант. Нужно отметить, что хотя полученный вариант и не является классическим штейнеровским МР-деревом (подобно МР-дереву гаплотипов в программе Fluxus Network), однако оно дает наглядное представление о характерных особенностях субструктуры и топологии кластеров динарского кластера (Рисунок 6) При построении дерева использовалась выборка из 580 16-маркерных (расширенный Yfiler формат DYS456, DYS389I, DYS390, DYS389II, DYS458, DYS19, DYS385, DYS393, DYS391, DYS439, DYS635, DYS392, YGA-TAN4, DYS437, DYS438, DYS448) гаплотипов из следующих баз данных референтной базе данных ДНК-маркеров населения Республики Беларусь, YHRD (Willuweit, Roewer 2007) [принадлежность к гаплогруппе-I2a1b2a1-L147.2/CTS5966 определялась предиктором Ydna.ru]. Кроме гаплотипов из вышеупомянутых баз данных, были использованы гаплотипы из приложений к следующим публикациям (Berger et al., 2005; Ljubković et al., 2008; Roewer et al., 2008; Karachanak et al., 2013; Mirabal et al., 2010). Территориальный охват выборки примерно соответствует географическому ареалу Восточной Европы. Предварительный тест на молекулярную дисперсию (AMOVA) не проводился, так как выборка организовывалась по принципу максимального Y-хромосомного генотипического сходства: поэтому генетическая вариативность между группами по этническим и географическим признакам не должны существенно отличаться от прогнозируемого.

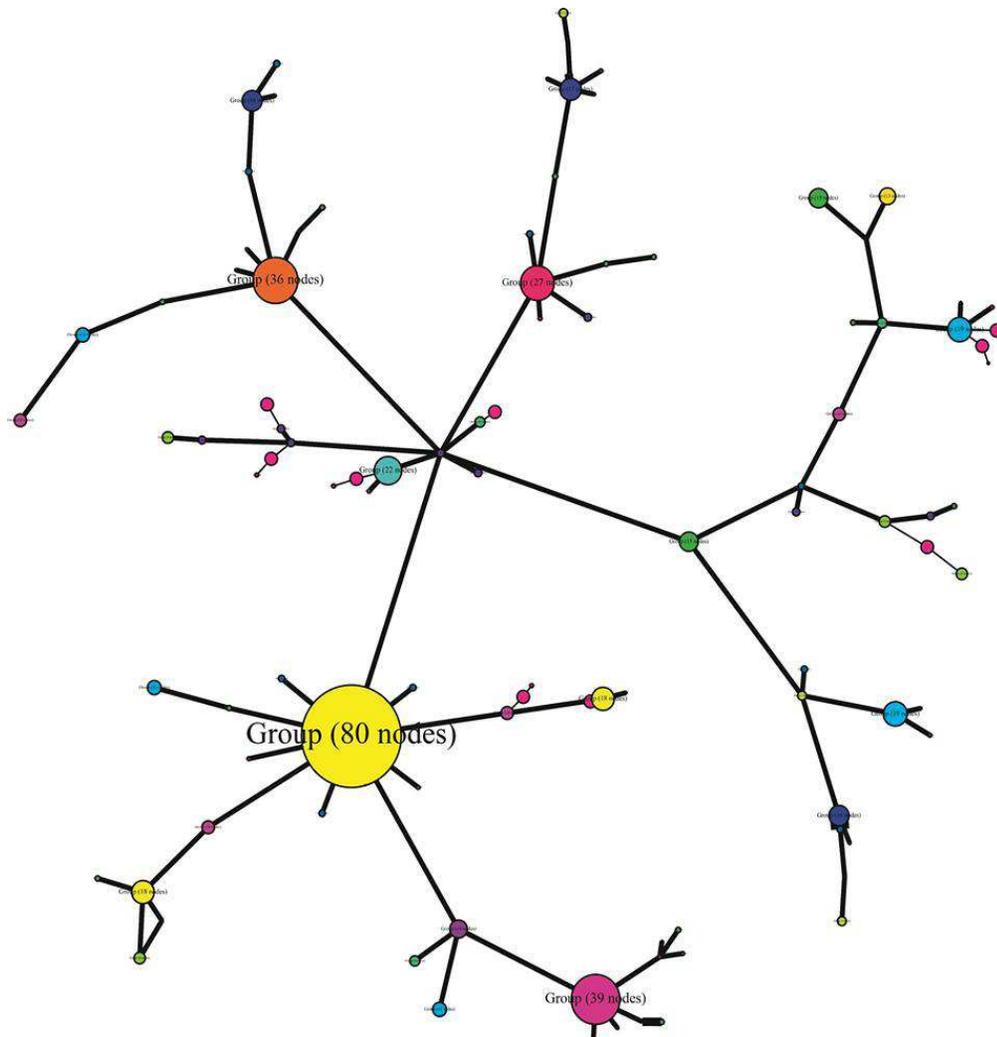


Рисунок 6. Визуализация MST (minimum spanning tree, минимальное остовное дерево) динарского кластера гаплогруппы I2a, полученного в программе Arlequin при вычислении матрицы попарной дисперсии аллелей в гаплотипах (Excoffier et al. 1992). В самом центре графа представлены гаплотипы из Беларуси и Украины. К центральной лежит на ребре графа связанного с группой Group (22 nodes, 22 узла), гаплотипы которого характеризуются характерной мутацией recLOH (DYS385a,b (14;14)). Наибольшее количество гаплотипов (80) находится в «желтом скоплении», где представлены в основной своей массе представители так называемого «южнодинарского» разнообразия (кластера Dinaric-S) (см. подробнее выше), и включает в себя преимущественно босняки, хорваты, сербы.

Примечательно что большое скопление гаплотипов вокруг большого узла ($n=80$, узел отмечен желтым цветом) имеет типичную форму филогенетического старкластера (Zerjal et al. 2003). Подобная форма филогенетических кластеров в основном наблюдается в тех случаях, когда происходит быстрый демографический рост одной конкретной филогенетической линии, и этот рост приводит к появлению серии одновременных мутационных событий. Исходя из этого можно предположить что экспансия носителей I2a-1b2a1-L147.2/CTS5966 на территории Балкан сопровождалась быстрым демографическим ростом популяции.

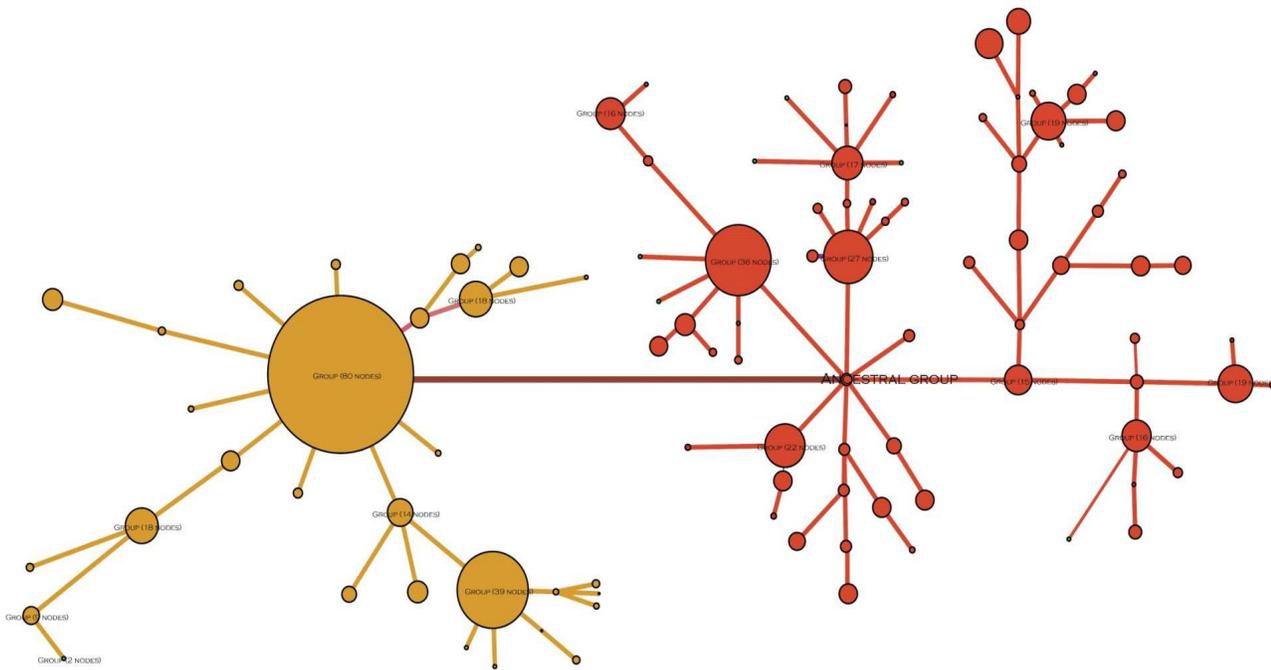


Рисунок 7. Пост-процессинг(Рисунок.6) вышеприведенного графа в Adobe Illustrator и Adobe Photoshop.

Филогенетический анализ в программе Fluxus Network

При построении дерева использовалась та же самая выборка из 580 [n=580] * 16-маркерных (расширенный Yfiler формат DYS456, DYS389I, DYS390, DYS389II, DYS458, DYS19, DYS385, DYS393, DYS391, DYS439, DYS635, DYS392, YGATAH4, DYS437, DYS438, DYS448) гаплотипов из базы данных YHRD [принадлежность к гаплогруппе -I2a1b2a1-L147.2/CTS5966 определялась предиктором Ydna.ru]. Филогенетические взаимоотношения микросателлитных гаплотипов внутри гаплогрупп и их разнообразие мы реконструировали и с помощью традиционного метода медианных сетей [Bandelt et al., 1999], реализованного в пакете программ Network 4.516 (<http://www.fluxus-engineering.com/sharenet.htm>). Поскольку полная сеть содержит в себе массу минимальных вероятных филогенетических деревьев, то наиболее вероятные филогенетические деревья конструировали путем нахождения минимального связного графа (штейнеровского дерева) (Polzin, Daneschmand 2003), используя дефолтные значения параметров Star Contraction (Forster et al. 2001) и epsilon (эпсилон). Мы выбрали графический вариант топологии штейнеровского дерева, которое наиболее точно согласуется с топологией минимального остовного дерева (см. выше Рис. 6 и Рис.7).

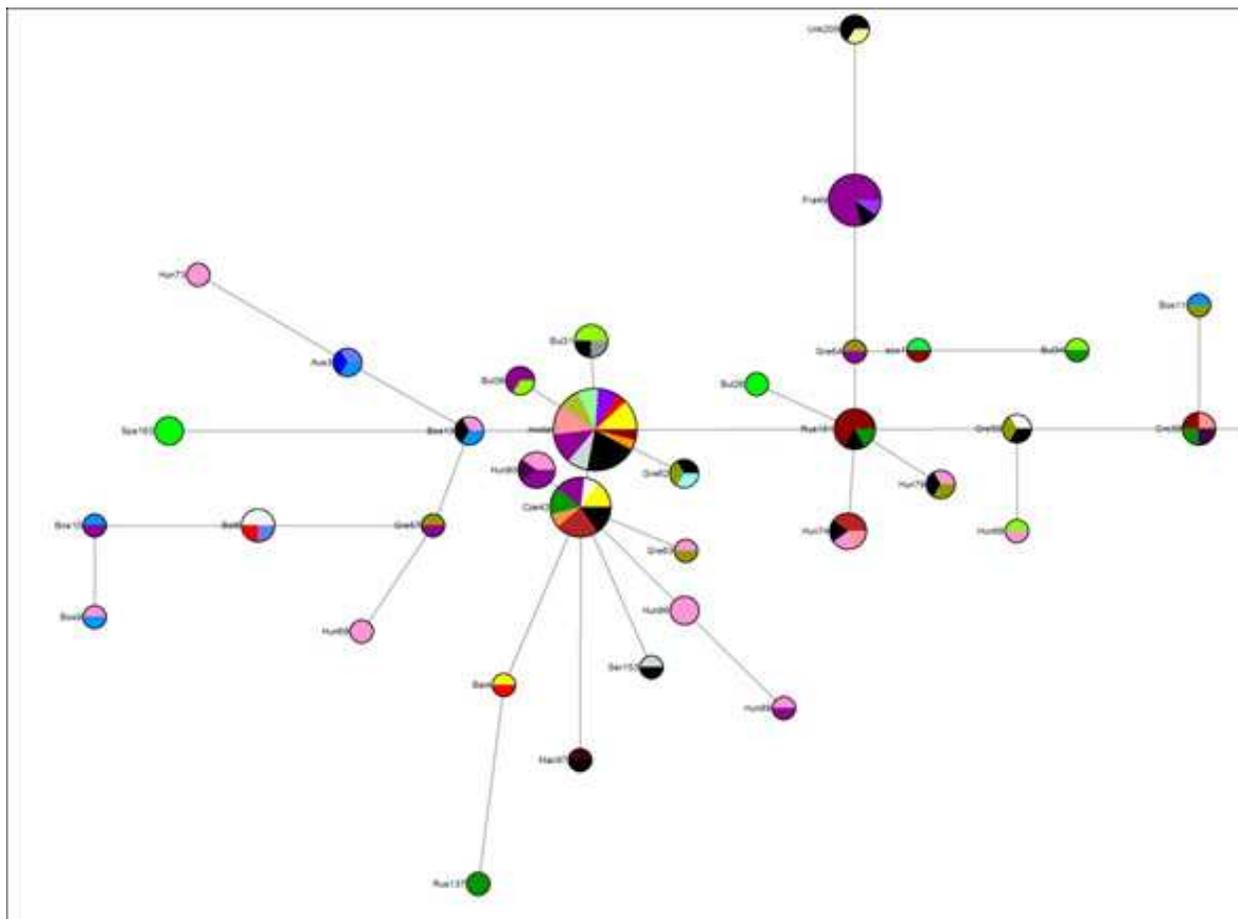


Рисунок 8. Филогенетическое MP-дерево (наиболее парсимоничное дерева) выборки из 580 [n=580] * 16-маркерных (расширенный Yfiler формат DYS456, DYS389I, DYS390, DYS389II, DYS458, DYS19, DYS385, DYS393, DYS391, DYS439, DYS635, DYS392, YGATAH4, DYS437, DYS438, DYS448) гаплотипов из базы данных YHRD. Размер узлов пропорционален количеству гаплотипов, длина ребра прямо-пропорциональна количеству мутаций. Этническая принадлежность гаплотипов обозначена особой цветовой легендой: красный (Беларусь), синий (Босния), темно-синий (Франция), салатовый (Болгария), оливковый (Греция), розовый (Венгрия), серый (Сербия), фиолетовый (Польша), черный(неизвестный этнос), оранжевый (Турция), коричневый (Украина), голубой (Италия), белый (Германия), желтый (Чехия), темно-фиолетовый (Хорватия), темно-зеленый (Россия), темно-серый (Узбекистан), темно-синий (Австрия), темно-салатовый (Македония), светло-фиолетовый (Румыния), темно-синий (Словения).

При применении методов редукции MP-деревьев филогенетических сетей гаплотипов, мы обнаружили 111 одинаково вероятных вариантов штейнеровских MP-деревьев медианной сети гаплотипов I2a2. При построении деревьев использовалась дополнительные возможности упрощения топологии сети гаплотипов в Fluxus Networks: к ним относятся в первую очередь MJ-эвристика, MP-пост обработка получившейся сети с учетом вычисленных весов (weighted heuristic, с неравновесной метрикой) STR-локусов, а также редукция штейнеровских деревьев.

TNT-дерево и возраст по данным rho-статистики

Как известно, парсимонистские филогении, основанные на коммерческих 17,25,37 локусах У-хромосомы, зачастую носят случайный и ненадежный характер, и поэтому Для решения этой проблемы мы использовали другую выборку гаплотипов ($n=378$) с большим числом Y-STR (панели 67 Y-STR маркеров и 111 Y-STR маркеров), гаплотипы собраны в открытых проектах **FTDNA**. Из собранных в первичную выборку гаплотипов в первую очередь отбирались гаплотипы с включенными в FTDNA SNP-ами гаплогруппы I2a (L68, P215, L596, L597, L415, L416, L417, L460, P37.2, L158, L159.1, M26, L160, M161, L178, M423, P41.2, L161, L621, L233, L35, L37, L181, P214, P216, P217, P218, L34, L36, L59, L622, M223, P219, P220, P221, M284, L126, L137, L369, L701, P78, L699, L703, Z161, L623, L801, P95, L38, L39, L40, L65.1, L272.3), затем внутри отфильтрованной выборки были отдельно сгруппированы гаплотипы по значениям двух маркеров, с характерными для северно-динарского и южно-динарского кластеров. В отдельную группу были размещены гаплотипы с значениями профильных маркеров, близкими к модальным значениям двух ближайших к «динарской» субкладе субклад – I2a1b2a1-CTS54002/CTS10936 – Disles-B и I2a1b2-L621/S392 – Disles-A (см. Рисунок 1). В неопределенных случаях, исключались гаплотипы I2a1b2a1 с британских островов. Наконец, в качестве outgroup-узла мы использовали модальный гаплотип гаплогруппы I1 - сестринской гаплогруппы I2. В итоге получилась выборка, включающая в себя в основном гаплотипы венгров Баната, Словакии, Трансильвании и Буковины; словаков, поляков, карпаторусинов, хорватов, словенцев, сербов, украинцев, белорусов, русских. Кроме этого, в выборке представлено несколько ашкенази, незначительное количество немцев, австрийцев, албанцы, итальянцы.

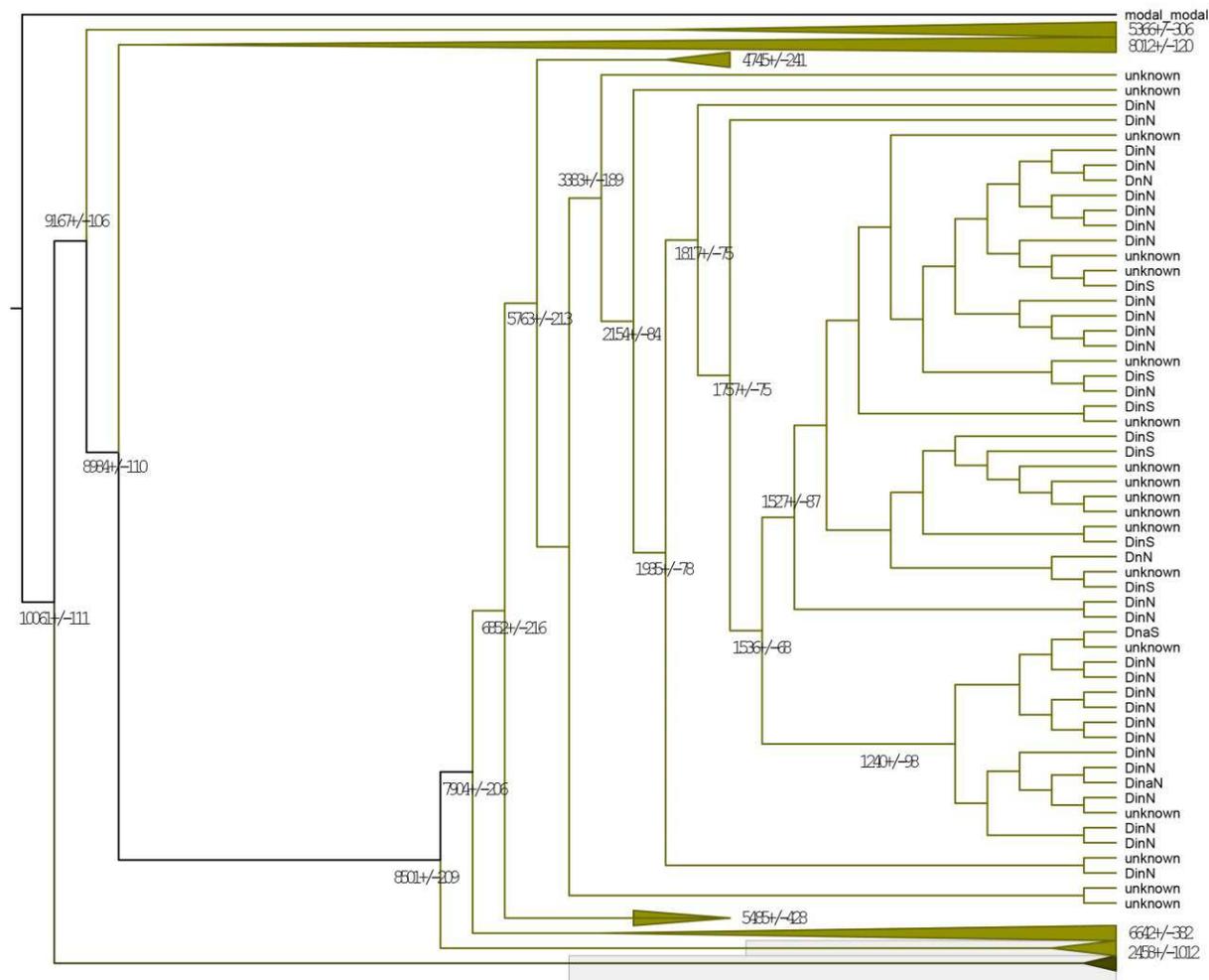
С помощью филогенетического программного обеспечения **TNT** и неравновесной метрики весов Y-SNP/Y-STR мы определили 5 (пять) штейнеровских деревьев минимальной стоимости Для выбора наиболее достоверного варианта из пяти оптимальных топологий филогенетических деревьев мы применили совокупность четырех признаков:

- 1) величина μ (суммарное число мутаций гаплотипа-потомка, накопленное от условного предкового гаплотипа), должна совпадать с величиной μ_i .
- 2) все вычисленные штейнеровские деревья (деревья с максимальной парсимонией) в случае достоверности построений, должны минимально различаться по топологии.
- 3) оценка частоты соответствующей партии (ветви) среди всех изученных MP-деревьев. Оценка этого параметра позволяет выявить «случайность» размещения конкретной «ветви» в «конкретном» древе. Согласно эмпирическим наблюдениям, средняя частота партий достоверных по другим признакам деревьев составляет ≥ 0.95 . То же самое можно сказать о достоверности ветвей: степень «типичности» партий гаплотипов в древе (то есть встречаемость их в других деревьях той же степени оптимальности) имеет большое значение при рассмотрении топологий; значительное число редких партий в древе свидетельствует о случайности его топологии, и в случае когда это характерно для всех оптимальных деревьев, можно говорить об отсутствии явного филогенетического «сигнала» в исходных данных.
- 4) оценка расстояния Робинсона-Фолдса между построенными деревьями: средняя дистанция Робинсона-Фолдса между деревьями не должна превышать выверенный порог средних расстояний $\langle 0.15-0.2 \rangle$, за пределом которого достоверная реконструкция мутационного дерева STR невозможна. В нашем случае, среднее расстояние между деревьями динарской субклады составляет округленно 0.10-, т.н. не превышает выверенный порог средних расстояний в 0.15-0.2.

Из возможных вариантов деревьев мы выбрали тот вариант, который наилучшим способом отвечают всем из перечисленных параметров (Рисунок 9, из-за большого масштаба дерева на рисунке отображен только интересующая нас субклада I2a1b2a1-L147.2/CTS5966). Обращает на себя внимание очевидная закономерность: при использовании 67-111 Y-STR локусных гаплотипов кластерное разбиение DinN («северный динарский кластер») и DinS («южный динарский кластер») не соблюдается, так как гаплотипы обоих кластеров размещаются вперемешку.

Это дополнительное подтверждение ненадежности разделения гаплогруппы на кластеры в соответствии с определенными аллельными значениями Y-STR локусов.

Оценка возраста субклады с помощью оценки генетической дистанции rh (ρ -статистики), рассчитываемой как среднее число мутационных шагов от гаплотипа-основателя (так называемого модального гаплотипа) ко всем остальным гаплотипам в составе медианной сети. Частоту мутаций микросателлитных локусов в данном случае мы определяли как равную $2 \cdot 10^{-3}$ на локус за 30 лет. Расчеты с помощью ρ -статистики показали, что возраст субклады I2a1b2a1-L147.2/CTS5966 составил 3400 ± 200 лет, время разделения I2a1b2a1-L147.2/CTS5966 и ближайшей субклады I2a1b2a* (кластер Disles)- 5750 ± 200 , время разделения I2a1b2-L621 (монофилетически включающей в себя распространенную динарскую субкладу и редкую субкладу «Disles») и «островной британской» субклады I2a1b1



4.0

Рисунок 9. Наиболее вероятная топология дерева гаплотипов гаплогруппы I2 с фокусом на структуру субклады I2a1b2a1-L147.2/CTS5966

Картографический анализ

В рамках полисистемного подхода мы провели филогенетический анализ, который позволяет взглянуть на вопрос происхождения гаплогруппы I2a2 (и всей гаплогруппы I2a, в целом) с точки зрения геногеографии. Приведенные выше данные о характере распространения и взаимного положения различных субклад гаплогруппы I2a (см. Таблицу 1), а также положение субклады I2a1b, к которой принадлежали древнейшие из найденных образцов гаплогруппы I2a (Loschbour и Motala, см. Рисунок 4) общая предковая популяция как континентальной клады I2a-1b2a1, так и клады I2ab1) с высокой степенью вероятности могла жить в Западной Европе – в регионе Альпийских гор. Развивая данный сценарий, можно предположить что предковая популяция древнеевропейских охотников-собирателей I2a1b (а также, видимо, и родственная ей предковая популяция различных клад гаплогрупп I2a2) начала свое движение с Альп в сторону северо-западной и центральной Европы сразу после окончания последнего ледникового максимума (т.е. в мезолите), следуя за отступающим на север ледником и тундровой флорой.

С целью проверки этой версии, было принято решение использовать алгоритм Cartographer (Maddison, Maddison 2014a) реализованный в пакетном приложении к программе Mesquite (Maddison, Maddison 2014b), который позволяет производить оценку наиболее вероятного (с точки зрения парсимонии) места «происхождения» субклады, исходя из нынешней географической дистрибуции (географического распространения) двух филогенетически ближайших клад. В данном случае мы определили наиболее вероятное место «происхождения» I2a1b2a1 на основании сопоставления географических координат таксонов-гаплотипов и топологии дерева I2a1b2a1, которое укоренялось outgroup-корнем – модальным гаплотипом «сардинской» гаплогруппы I2a1a). Несмотря на то, что наибольшая плотность субклады I2a1a приходится на Сардинию, при размещении модального гаплотипа I2a1a на Пиренейском полуострове, мы руководствовались хорошо аргументированным и эмпирически подтвержденным мнением Нордтведта, согласно которому самые старые по возрасту кластеры гаплогруппы I2a1a приходятся на северный регион Пиренейского полуострова, и частично южную Францию

Поэтому в качестве координат аутгруппа-корня (модального гаплотипа I2a1a) были приняты координаты географического центра Пиренейских гор. Для картографического анализа в картографическом пакете Cartographer программы Mesquite, мы использовали дополнительную выборку из 687 37-локусных гаплотипов I2a1b2a1 из проектов FTDN. Оптимальный вариант филогенетического дерева этой выборки был выявлен путем нахождения 20 независимых совпадений 20 парсимонически лучших деревьев в программе TNT (Goloboff et al. 2008). Затем это дерево было выгружено из TNT в формате .tre и загружено в картографический пакет Мескита, где оно было спроецировано на карту мезолитической Европы с отмеченными на ней археологическими стоянками времен финального палеолита и мезолита (мы использовали карту из атласа «Археология Западной Европы. Каменный век.А» (Монгайт 1973), карта приведена в проекции Меркатора). Благодаря наличию встроенного алгоритма для определения максимально вероятной (в терминах парсимонической комбинации географически близких ныне живущих таксонов) локации «предковой популяции» субклады I2a-1b2a1. На реконструированной карте место «происхождения» субклады разместилось чуть южнее швейцарских Альп (обозначено большой красной точкой) (Рисунок 10), примерно в 420 километрах к юго-востоку от места обнаружения древнейшего (8000 лет до настоящего времени) образца ДНК мужчины, принадлежавшего к гаплогруппе I2a1b (Loschbour-Heffingen, Luxembourg) и примерно в 650 километрах к югу от Лихтенштейнской пещеры, в которые были обнаружены 4 скелета мужчин с древнейшими (3000 лет до настоящего времени, культура полей погребения) из найденных образцов гаплогруппы I2a2b, а также в 920 к юго-западу от местонахождения неолитического поселения Арс-Верекалја (Венгрия), где был обнаружен древнейший из найденных образцов I2a2a (возрастом в 6700 лет).

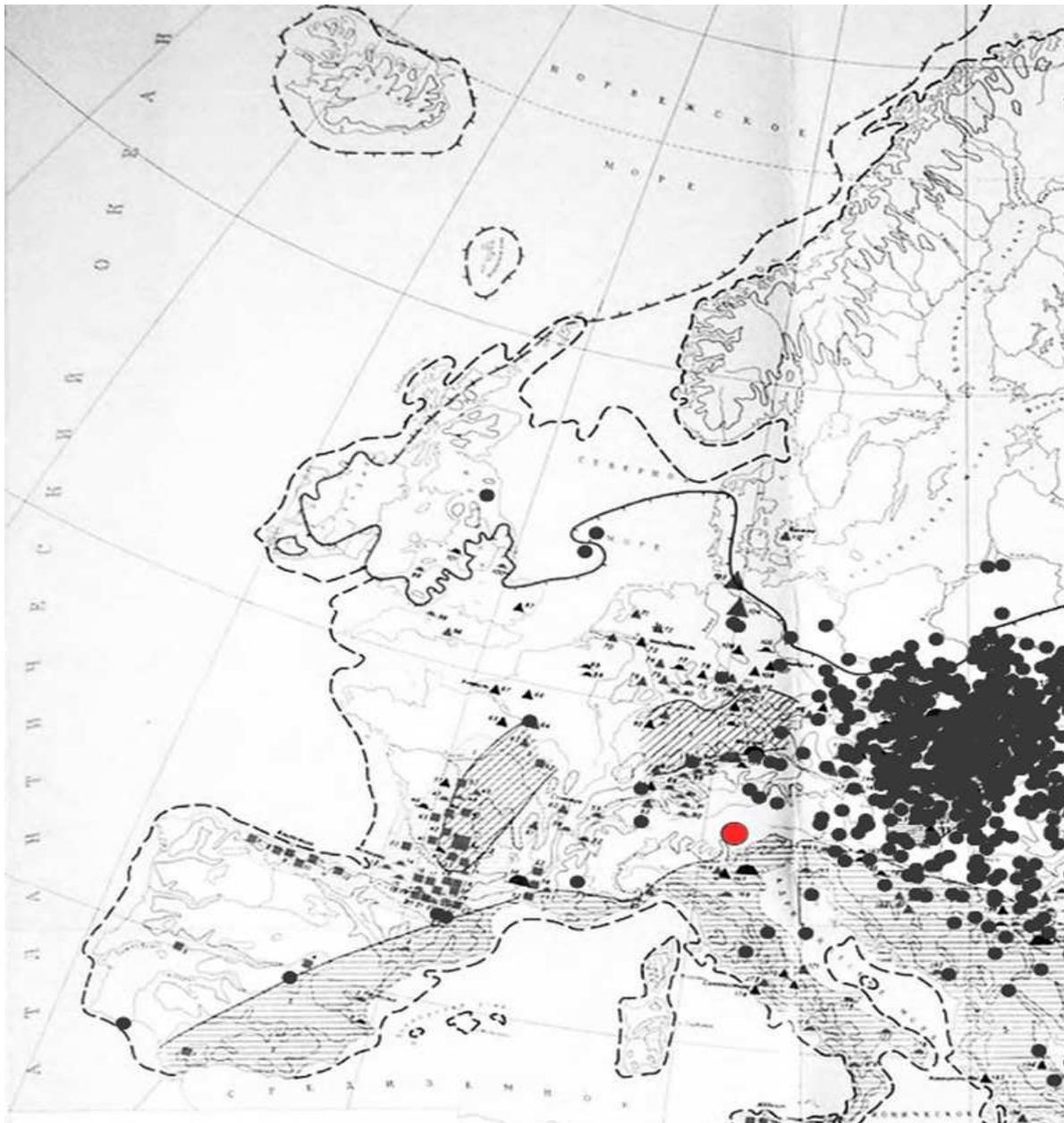


Рисунок 10. Карта археологических культур Европы каменного века с нанесенной на нее проекцией филогенетического дерева гаплогруппы I2a1b2a1. Каждый из гаплотипов-таксонов дерева обозначен черным кружком, красный кружок обозначает наиболее вероятную реконструкцию региона происхождения гаплогруппы I2a1b*

Оценка возраста

Оценку времени, за которое было накоплено наблюдаемое разнообразие гаплотипов в гаплогруппе, т.е. ее возраст мы оценивали двумя методами: 1) на основании средних квадратных различий в числе повторов по каждому локусу между всеми хромосомами и гаплотипом-основателем, 2) с помощью оценки генетической дистанции ρ , рассчитываемой как среднее число мутационных шагов от гаплотипа-основателя ко всем остальным гаплотипам в составе медианной сети. При расчетах возраста субклады обнаружилась удивительная аномалия – при солидном возрасте самой гаплогруппы I2a, динарская субклада I2a1b2a1 довольно «молода». TMRCA, или время жизни «последнего общего предка» по стандартным расчетам возраста с использованием «генеалогических» скоростей, которые используются в генетической генеалогии, составляет в среднем примерно 3000 лет до настоящего времени.

Ниже приведена таблица с оценками TMRCA I2a1b2a1 разными методами и разными скоростями мутации (для пересчета числа поколений в года на поколенный интервал было принято 30 лет, за исключением BATWING для которого использовался интервал в 25 лет) и с использованием разных панелей Y-STR локусов FTDNA. Для расчета использовались 774 гаплотипов I2a-1b2a1 с разным количеством маркеров (37, 67 и 111 Y-STR локусов), взятые из открытых публичных проектов FTDNA. Среднее медианное значение интервала полученных возрастов составляет 2757 лет, стандартное отклонение 404 лет (2757 ± 404), верхний предел интервала TMRCA перекрывается с интервалом TMRCA, вычисленного с помощью ρ -статистики 3400 ± 200 лет:

TMRCA -12 локусов с использованием метода Кена Нордтведта (Nordtvedt 2008) и скоростей мутаций Чандлера (Chandler 2006) :								
Ч и с л о поколений	82							
В годах	2461							
TMRCA -37 локусов с использованием метода Кена Нордтведта и скоростей мутаций Чандлера :								
Ч и с л о поколений	73							
В годах	2190							
TMRCA -67 локусов с использованием метода Кена Нордтведта и скоростей мутаций Чандлера/Литтла (Little 2007) :								
Ч и с л о поколений	77							
В годах	2360							
TMRCA -12 локусов с использованием метода Кена Нордтведта и скоростей мутаций SMGF (2004) :								
Ч и с л о поколений	57							
В годах	1700							
TMRCA -37 локусов с использованием метода Кена Нордтведта и скоростей мутаций Чандлера без локусов CDYa, CDYb и DYS464 :								
Ч и с л о поколений	103							
В годах	3080							
TMRCA -67 локусов с использованием метода Кена Нордтведта и скоростей мутаций Чандлера без локусов CDYa, CDYb и DYS464 :								
Ч и с л о поколений	103							
В годах	3085							
TMRCA -37 локусов с использованием метода Кена Нордтведта и скоростей мутаций Чандлера без локуса DYS464 :								
Ч и с л о поколений	79							
В годах	2375							
TMRCA -37 локусов с использованием метода Кена Нордтведта и скоростей мутации Чандлера без локусов CDYa, CDYb :								
Ч и с л о поколений	88							
В годах	2637							
TMRCA -12 локусов с использованием метода Джеймса Хильда и скоростей мутации Чандлера :								
Ч и с л о поколений	89							

В годах	2657							
TMRCА -37 локусов с использованием метода Джеймса Хильда и скоростей мутации Чандлера:								
Ч и с л о поколений	98							
В годах	2940							
TMRCА -67 локусов с использованием метода Джеймса Хильда и скоростей мутации Чандлера:								
Ч и с л о поколений	96							
В годах	2895							
TMRCА -37 локусов с использованием метода Джеймса Хильда и скоростей мутации Чандлера без локусов CDYa, CDYb и DYS464:								
Ч и с л о поколений	110							
В годах	3297							
TMRCА -67 локусов с использованием метода Джеймса Хильда и скоростей мутации Чандлера без локусов CDYa, CDYb и DYS464:								
Ч и с л о поколений	102							
В годах	3072							
TMRCА -12 локусов с использованием метода Джеймса Хильда и скоростей мутации Чандлера без локусов DYS385ab:								
Ч и с л о поколений	86							
В годах	2577							
TMRCА -37 локусов с использованием метода Джеймса Хильда и скоростей мутации Чандлера без локусов DYS 385, DYS 459, DYS 464, YCA II, и CDY:								
Ч и с л о поколений	103							
В годах	3082							
TMRCА -67 локусов с использованием метода Джеймса Хильда и скоростей мутации Чандлера без локусов DYS 385, DYS 395, DYS 413, DYS 425, DYS 459, DYS 464, YCA II, и CDY:								
Ч и с л о поколений	100							
В годах	3004							
TMRCА -12 локусов с использованием метода Кена Нордтведта и скоростей мутации Чандлера без локусов DYS385ab:								
Ч и с л о поколений	77							
В годах	2309							
TMRCА -37 локусов с использованием метода Кена Нордтведта и скоростей мутации Чандлера без локусов DYS 385, DYS 459, DYS 464, YCA II, и CDY:								
Ч и с л о поколений	99							
В годах	2981							
TMRCА -67 локусов с использованием метода Кена Нордтведта и скоростей мутации Чандлера без локусов DYS 385, DYS 395, DYS 413, DYS 425, DYS 459, DYS 464, YCA II, и CDY:								
Ч и с л о поколений	102							
В годах	3074							
TMRCА -8 локусов с использованием метода BATWING (Wilson et al .2003), гамма распределением скоростей мутаций, N, alpha, beta:								
Ч и с л о поколений до коалесценции	139							
В годах	3475							

Таблица 4. Таблица с рассчитанными значениями TMRCА – времени жизни последнего общего предка выборки гаплотипов I2a1b2a1.

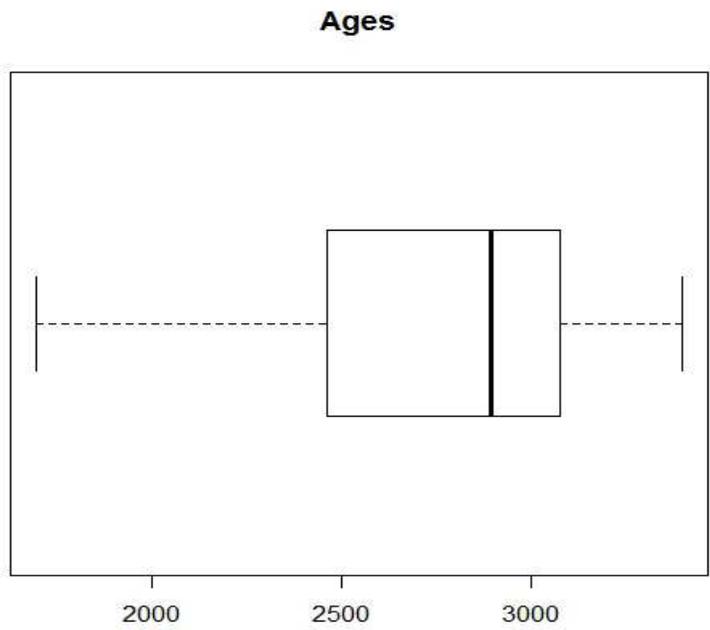


Рисунок 11. Диаграмма разброса полученных значений TMRCA (Таблица 4) , на которой изображены медиана, нижний и верхний квантили, минимальное и максимальное значение интервала возрастов и выбросы.

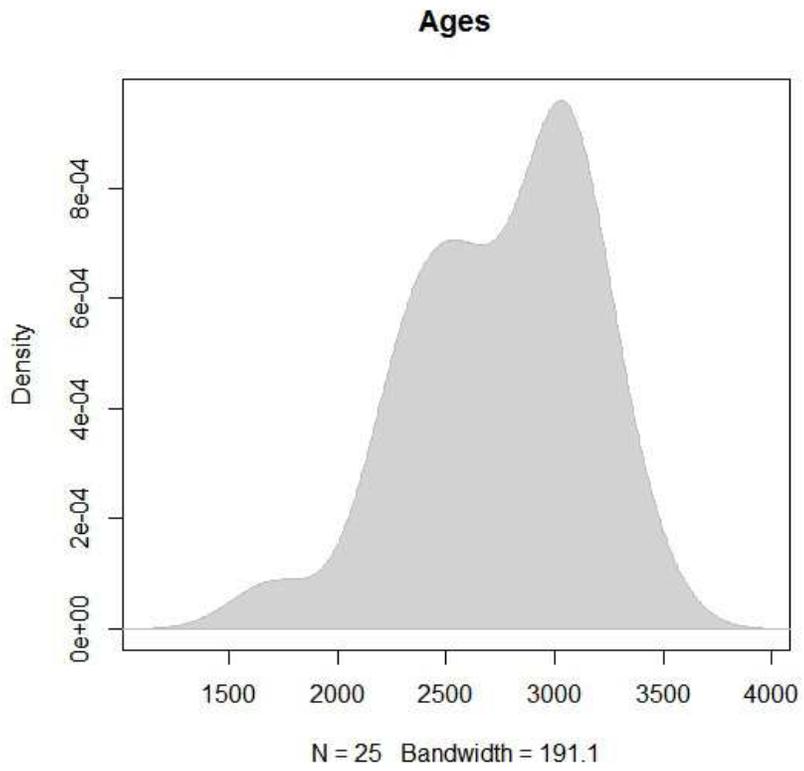


Рисунок 12. Гистограмма распределения плотности полученных значений TMRCA (Таблица 4).

Статистический анализ

Анализ генетической изменчивости

В рамках предложенного выше полисистемного подхода, для расчета генетической вариативности микросателлитных гаплотипов и дифференциации популяций были использованы стандартные статистические методы. Анализ генетического разнообразия имеет существенное значение для определения места происхождения отдельных генетических линий (Животовский 2006): традиционно считается, что в том месте или популяции, где возникла данная гаплогруппа, ее частота и STR-дисперсия (или возраст STR-изменчивости) максимальны по сравнению с другими популяциями (Sengupta et al., 2006). При их несовпадении, когда максимум частоты гаплогруппы приходится на один географический регион, а максимум дисперсии – на другой, место возникновения гаплогруппы становится неопределенным, но при необходимости сделать предварительное заключение, предпочтение следует отдавать дисперсии как статистически и эволюционно более устойчивому показателю. Как станет ясным из дальнейшего изложения материала, в случае обсуждаемой гаплогруппы I2a1b2a1 несовпадение популяций (мест) с наивысшей частотой распространения гаплогруппы с популяциями (местами) наивысшего разнообразия гаплотипов. Так, безусловный пик-максимум распространения I2a1b2a1 приходится на Балканы (особенно на регион Боснии-Герцеговины), а один из пиков значений дисперсии (разнообразия) гаплотипов - на полесско-карпатский регион Восточной Европы. Исходя из общего правила, процитированного выше, в этой работе предпочтение было отдано дисперсии как более устойчивому индикатору, в то время, как частоты гаплогрупп больше подвергнуты флуктуациям.

Так, именно на высоком STR-разнообразии субклад всей гаплогруппы I2a на Балканах (индекс гаплотипного разнообразия-среди боснийцев (0,93) и хорваты (0,85)), наряду с максимальной величиной частоты распространения I2a на Балканах, авторы ранних работ основывали свое мнение о том, что гаплогруппа I2a присутствовала на Балканах уже во времена последнего ледникового максимума (Underhill et al. 2007; Semino et al. 2000). Однако, как было показано в последующих работах, не менее высокие значения STR-разнообразия субклад всей гаплогруппы I2a наблюдались также в популяциях чехов и словаков (0,90), гораздо более низкие значения разнообразия были обнаружены в популяциях турок (0,76) и молдаван (0,41) (Cinnioglu et al. , 2004).

В целях проверки данных выводов, в программе Arlequin был проведен анализ молекулярного и стандартного разнообразия (Таблица 5), а также вычислена генетическая дистанция гаплотипов представителей субклада I2a1b2a1. Комплексный анализ трех параметров позволяет, наряду с филогенетическим анализом, определить место вероятное происхождения субклада, а также предположить характер и степень влияния популяционных эффектов. В упомянутых анализах использовалась выборка из 624 гаплотипов взятых из открытых проектов FTDNA, Ysearch и академических выборок (каждый гаплотип включает от 17-67 стандартных Y-STR локусов соответствующих панелей FTDNA) этого субклада (плюс некоторое количество гаплотипов из научных выборок). Совокупная выборка гаплотипов была разделена на этнопулы, согласно задекларированной национальности носителя гаплотипа (Рисунок 13) (в целях избежания статистических флуктуаций связанных с разной длиной гаплотипов, в пулах было зафиксировано примерно одинаковое количество гаплотипов разной длины).

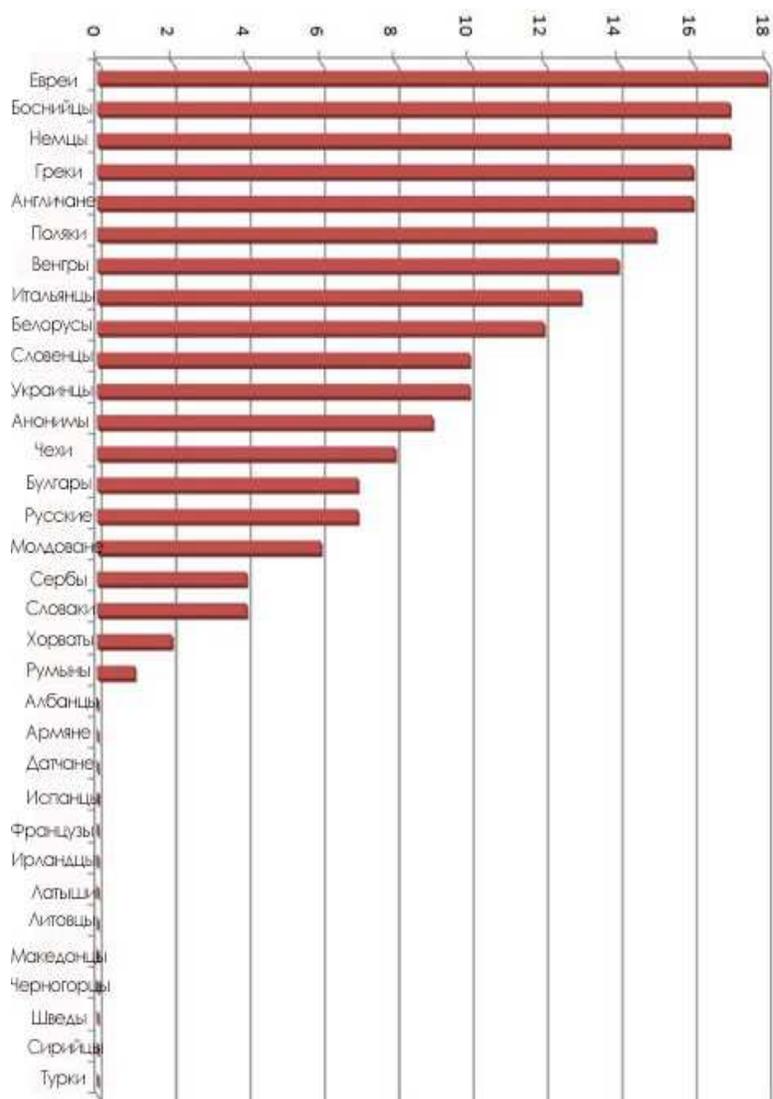


Рисунок 13. Визуализация генетической дистанции между этнопулами по вычисленным в программе Arlequin значениям

Полученные этнопулы или этнические кластеры имеют специфический: например в выборке выделяется характерный еврейский кластер (украинские и белорусские евреи-представители этого кластера имеют редкие характерные значения Y-STR локусов: DYS537 = 11, DYS464a = 13, DYS456 = 14, DYS458 = 18, DYS576 = 19, DYS570 = 16) с самой высокой интерпопуляционной дистанцией (Рисунок 13) от остальных кластеров и самым низким разнообразием (индекс стандартного разнообразия 0.7238 +/- 0.1206; индекс молекулярного разнообразия 0.100952 +/- 0.080414) (Таблица 5) указывает на недавнее происхождение кластера как следствие чистого эффекта основателя, имевшего место примерно 500 лет назад.

Боснийско-герцеговинский кластер (второй после еврейского кластера по величине интерпопуляционной дистанции), но с более высоким уровнем разнообразия указывает на более удаленный по времени популяционный эффект основателя. В то время, как географически близкий к боснийско-герцеговинскому кластеру хорватский кластер I2a-1b2a1 характеризуется относительно низким уровнем молекулярного разнообразия (многие хорватские гаплотипы I2a1b2a1 имеют характерный гаплотипный мотив DYS19=14), и незначительной дистанцией с другими популяциями, что отражается в более низком значении индекса стандартного разнообразия (что подразумевает изменение размера эффективной популяции). Данный эффект можно объяснить кумулятивным действием двух параллельных популяционных эффектов - эффекта основателя и эффекта «бутылочного горлышка» (Рисунок 13 и 14; Таблица 5).

Популяция	Индекс стандартного разнообразия (Garza and Williamson 2001.)	Индекс молекулярного разнообразия (Nei 1987; Tajima 1993)
Чехи	1.0000 +/- 0.0171	0.366959 +/- 0.259787
Молдоване	1.0000 +/- 0.0340	0.350649 +/- 0.229447
Русские	1.0000 +/- 0.0071	0.337968 +/- 0.197325
Сербы	0.9643 +/- 0.0772	0.305195 +/- 0.199028
Украинцы	0.9977 +/- 0.0058	0.300929 +/- 0.177936
Словенцы	1.0000 +/- 0.003	0.286325 +/- 0.173064
Венгры	0.9913 +/- 0.0077	0.275443 +/- 0.160077
Румыны	0.9524 +/- 0.0955	0.273810 +/- 0.183153
Англичане	0.9867 +/- 0.0092	0.267089 +/- 0.144476
Белорусы	0.9723 +/- 0.0257	0.265613 +/- 0.16387
Немцы	0.9921 +/- 0.0035	0.260986 +/- 0.151255
Поляки	0.9844 +/- 0.0089	0.257135 +/- 0.149835
Литовцы	1.0000 +/- 0.5000	0.250000 +/- 0.288675
Словаки	0.9919 +/- 0.0099	0.229420 +/- 0.134482
Болгары	0.9286 +/- 0.0844	0.220238 +/- 0.149337
Боснийцы	0.9632 +/- 0.0282	0.213596 +/- 0.133580
Греки	0.9800 +/- 0.0175	0.201667 +/- 0.126114
Черногорцы	1.0000 +/- 0.5000	0.166667 +/- 0.204124
Итальянцы	0.9265 +/- 0.0579	0.146324 +/- 0.104366
Хорваты	0.9643 +/- 0.0772	0.130952 +/- 0.106487
Евреи	0.7238 +/- 0.1206	0.100952 +/- 0.080414
Албанцы	-	-
Без указания этноса	-	-
Армяне	-	-
Датчане	-	-
Испанцы	-	-
Французы	-	-
Ирландцы	-	-
Латыши	-	-
Мексиканцы	-	-
Шведы	-	-
Сирийцы	-	-
Турки	-	-

Таблица 5. Значения индексов молекулярного и стандартного разнообразия гаплотипов I2a1b2a1, сгруппированных по этническому признаку (значения рассчитаны в программе Arlequin) и отсортированных по убыванию значения индекса молекулярного разнообразия

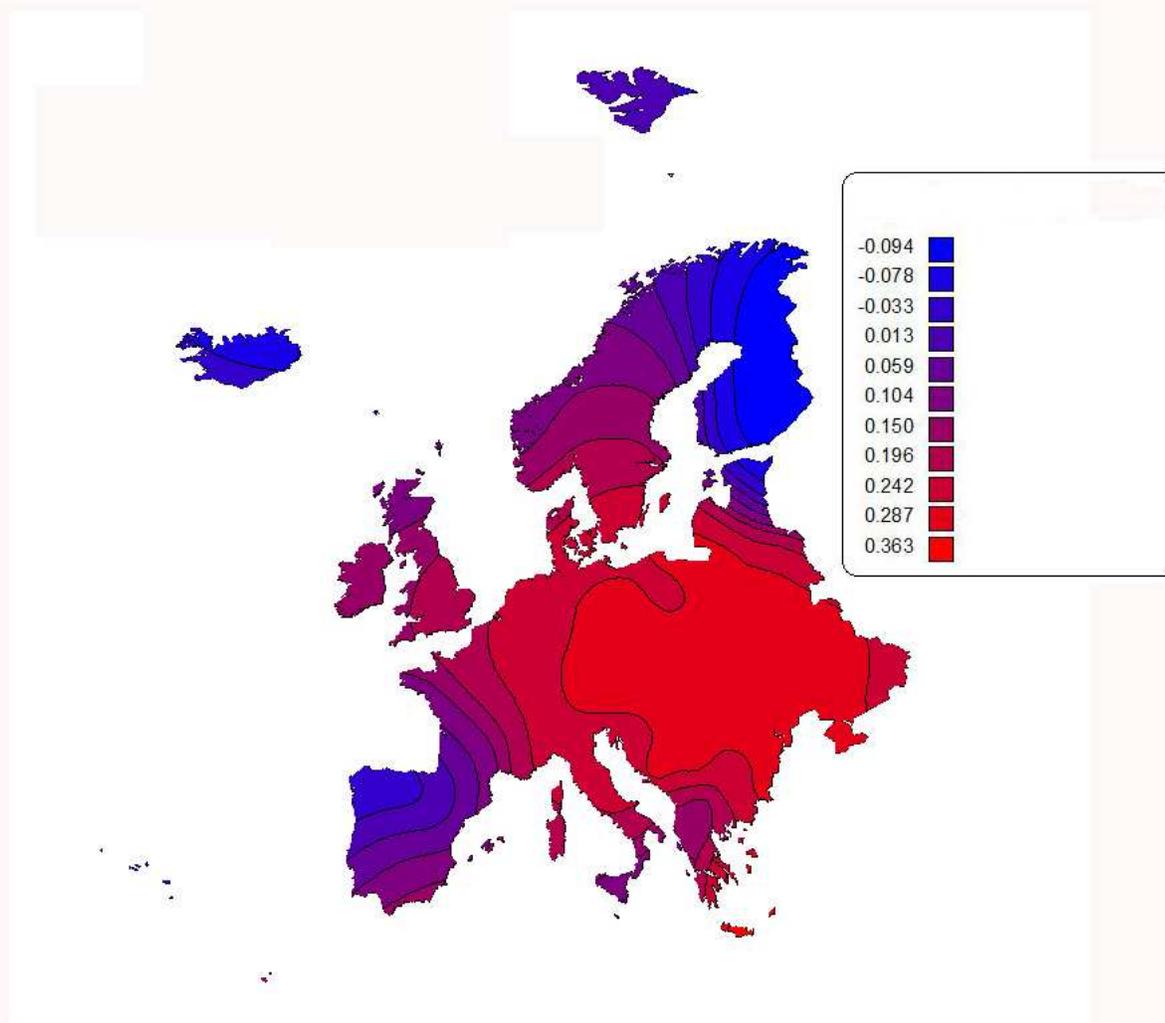


Рисунок 14. Картографическая визуализация значений гаплотипного разнообразия гаплогруппы I2a1b2a1 в Европе (программа Golden MapViewer 7). При построении карты использовался kriging-алгоритм пространственной нормализации значений индекса гаплотипного (молекулярного) разнообразия гаплогруппы I2a1b2a1

Источник разнообразия	Степени свободы	Сумма квадратов	Компоненты разнообразия	Процент разнообразия
Между группами	14	8.808	0.00100 Va	0.2
Между популяциями внутри групп	17	9.365	0.00471 Vb	0.96
Внутри популяции	575	278.462	0.48428 Vc	98.83
Сумма	606	296.636	0.48999	100
Индексы фиксации				
FSC:		0.00964		
FST:		0.01166		
FCT:		0.00204		
Статистические тесты значимости (1023 пермутации)				
Vc и FST:		P(случайная величина < наблюдаемой величины)= 0		
		P(случайная величина = наблюдаемой величине)= 0		
		P-value = 0.00000+-0.00000		
Vb и FSC:		P(случайная величина < наблюдаемой величины)= 0.00587		
		P(случайная величина = наблюдаемой величине)= 0		
		P-value = 0.00587+-0.00295		
Va и FCT:		P(случайная величина < наблюдаемой величины)= 0.28543		
		P(случайная величина = наблюдаемой величине)= 0		
		P-value = 0.28543+-0.01317		

Таблица 6. Тест AMOVA (Weir and Cockerham 1984; Excoffier, Smouse and Quattro 1992; Weir 1996) в выборке гаплотипов I2a1b2a1, разбитой на группы по лингвистическому признаку

Источник разнообразия	Степени свободы	Сумма квадратов	Компоненты разнообразия	Процент разнообразия
Между группами	12	7.86	0.00207 Va	0.42
Между популяциями внутри групп	19	10.314	0.00378 Vb	0.77
Внутри популяции	575	278.462	0.48428 Vc	98.81
Сумма	606	296.636	0.49013	100.00
Индексы фиксации				
	FSC:	0.00774		
	FST:	0.01192		
	FCT:	0.00422		
Статистические тесты значимости (1023 пермутации)				
Vc и FST :		P(случайная величина < наблюдаемой величины) = 0		
		P(случайная величина = наблюдаемой величине) = 0		
		P-value = 0.00000+-0.00000		
Vb и FSC:		P(случайная величина < наблюдаемой величины) = 0.00489		
		P(случайная величина = наблюдаемой величине) = 0		
		P-value = 0.00489+-0.00203		
Va и FCT:		P(случайная величина < наблюдаемой величины) = 0.1437		
		P(случайная величина = наблюдаемой величине) = 00		
		P-value = 0.14370+-0.01079		

Таблица 7. Тест AMOVA (Weir and Cockerham 1984; Excoffier, Smouse and Quattro 1992; Weir 1996) в выборке гаплотипов I2a1b2a1, разбитой на группы по антропологическому признаку

Попарное сравнение результатов теста AMOVA (Таблицы 6 и 7) по 2 группировкам популяций -лингвистической (популяции объединены в группы по принадлежности к той или иной языковой группе) и антропологической (популяции объединены по принципу генетической близости антропологических признаков их представителей) - показывает, что динамика разнообразия отдельных этнопулов гаплотипов I2a1b2a1 лучше объясняется лингвистическими, чем антропологическими признаками. Как и следовало ожидать в случае гомогенной выборки гаплотипов, в обеих группировках 98% разнообразия приходится на вариации внутри популяций, т.е. между отдельными гаплотипами (P-value = 0.00000+-0.00000). Самые заметные различия результатов наблюдаются при сравнении процентов генетического разнообразия гаплотипов, приходящегося на межпопуляционное разнообразие внутри групп. При лингвистической группировке популяций на межпопуляционное разнообразие внутри групп приходится 0.96% разнообразия, однако как показывают тесты статистической значимости, среднее значение p, т.е. вероятности верности нулевой гипотезы (о наличии вариации между популяциями I2a1b2a1 внутри лингвистических групп) больше 0.005. Поэтому, с точки зрения статистики случайных чисел, следует признать этот результат статистически незначимым, т.е. случайным. Процент межпопуляционного разнообразия между популяциями внутри «антропологических» групп ниже - 0.77% вместо 0.96% в лингвистической группировке. При этом, в антропологической группировке вероятность верности нулевой гипотезы, т.е. того, что этнопопуляции при сравнении внутри групп являются разнообразными в плане полиморфизмов Y-STR гаплотипов I2a1b2a1, составляет 0.00489, т.е. является статистически значимым.

Полученные данные свидетельствуют в пользу нашего предположения о том, что гаплогруппа I2a1b2a1 не является автохтонной гаплогруппой Балкан и Динарских Альп, в противном случае наблюдалась бы более существенная корреляция между популяцией носителей этой гаплогруппы и антропотипом⁶. Поскольку популяции I2a1b2a1 лучше коррелируют с языком, а не с антропотипом, то можно сделать два вывода: 1) экспансия носителей субклада I2a1b2a1 произошла недавно, т.к. не утерялась связь представителей генетической линии с языком; 2) поскольку славянские языки были явно привнесены на Балканы, то нужно признать, что I2a1b2a1 были в числе генетических линий, представители которых привнесли славянские языки на Балканы.

⁶ Дуализм антропологических параметров популяций Динарских Альп и языковой принадлежности - хорошо известный исследователям факт.

Гипотетическое направление миграций можно восстановить с помощью проекции градиентов значений индекса молекулярного разнообразия на векторную карту со стрелками, отображающими направление уменьшения значений индекса молекулярного разнообразия. Примечательно, что изображенные на Рисунке 16 стрелки направления уменьшения разнообразия гаплотипов в значительной мере перекрываются с маршрутами славянской экспансии (Рисунок 15)⁷.

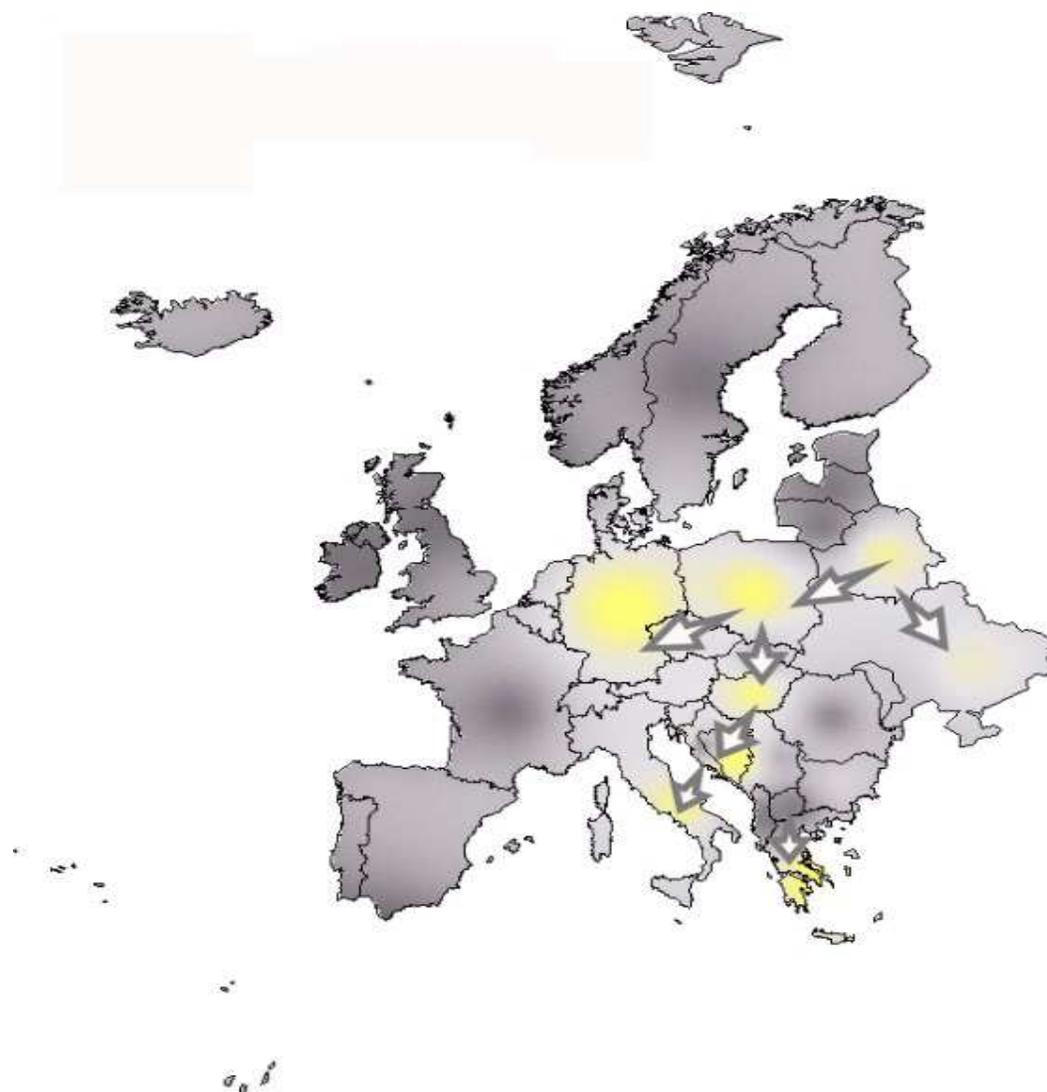


Рисунок 15. Направление исторических миграций славян в (желтым цветом в программе Golden Surfer 8 были отмечены критические зоны со значимыми значениями молекулярного разнообразия гаплотипов I2a1b2a1)

⁷ Исключение составляет стрелка, направленная из Боснии в Венгрию, однако здесь, скорее всего отражается хорошо известный историкам процесс омадьяривания славянского населения Боснии, которая довольно долгое время находилась в составе венгерского баната Боснии, а также процесса массовой миграции бошняков-католиков в Венгерское королевство после захвата турками Боснии в середине 15 века.

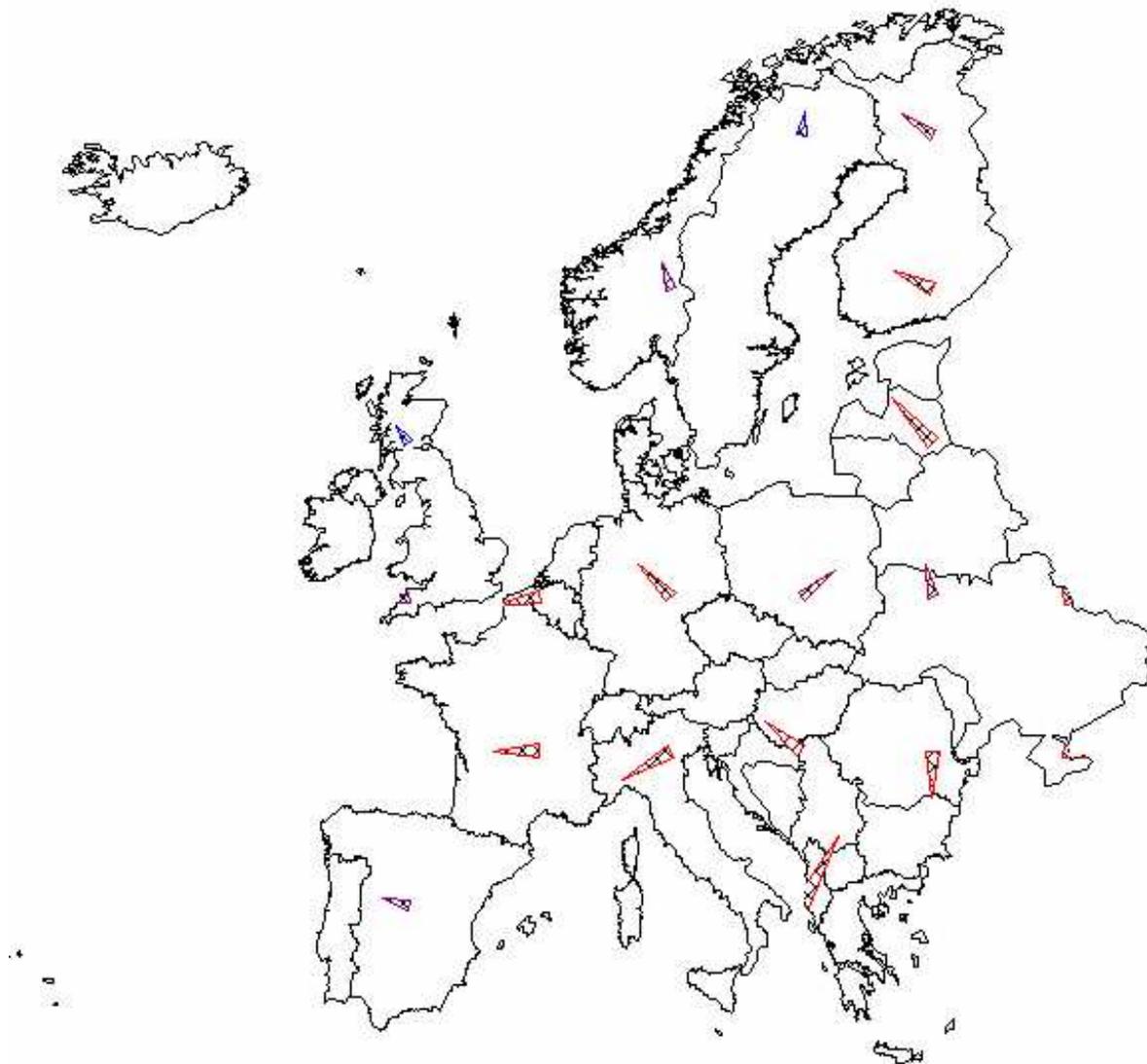


Рисунок 16. Векторная карта, сгенерированная в программе Golden Surfer 8 на основании интерполяции значений индексов молекулярного разнообразия гаплотипов, стрелками показано направление уменьшения гаплотипного разнообразия I2a1b2a1

Многомерное шкалирование

Важной проблемой в изучении гаплогруппы I2a1b2a1 является вопрос выявления этнической структуры в общем пуле гаплотипов. Одним из самых надежных инструментов для ответа на этот вопрос является многомерное шкалирование, целью которого является обнаружение структуры в наборе значений некоторой меры расстояния между объектами. В качестве объекта мы использовали 42 популяционные группы, в качестве меры расстояния – значения рассчитанной в **Arlequin** матрице Fst-расстояний между 42 популяционными группировками 774 гаплотипов динарской субклады I2a1b2a1 (Рисунок 17).

При статистических расчетах обнаружилась следующая аномалия: скорректированный коэффициент детерминации R^2 в данном случае негативный; впрочем, этого следовало бы ожидать, так как сам коэффициент детерминации R достаточно близок к нулю ($R=0.015$). Другими словами, предложенная модель разбивки носителей по этногеографическим группам объясняет 1.5% всей статистической вариативности всей выборки гаплотипов I2a1b2a1. Кроме того, p -value=1, что подразумевает принятие нулевой гипотезы отсутствия корреляции между гаплотипами по этническому признаку. Эти данные, несмотря на использование принципиально иного метода, подтверждают аккуратность полученных значений AMOVA, согласно которым на генетическое разнообразие между этническими группами приходится только 1% всего генетического разнообразия выборки, 98% приходится на разнообразие между отдельными гаплотипами.

Говоря простыми словами, в выборке гаплотипов I2a1b2a1 отсутствует значимое разделение по этническому признаку (а популяции плохо разделяются на географические кластеры – см. кластерограмму Рисунок 18). Более надежное доказательство вышеозвученного вывода было получено в ходе проведения классического теста **Мантеля (Smouse et al. 1986)**, с помощью которого определяется наличие и надежность корреляции между матрицей географических расстояний этнопулов гаплотипов I2a1b2a1 и матрицей попарных значений Fst между группами популяций. Как вытекает из результатов теста Мантеля (Рисунок 19), значения p -value с двухсторонним критерием (two-tailed p value=0.34) значительно выше 0.05, что в свою очередь означает что в выборке I2a1b2a1 не наблюдается статистически значимой корреляции между географической и генетической дистанцией даже в столь приличной по размерам выборке (774 гаплотипа).

MDS

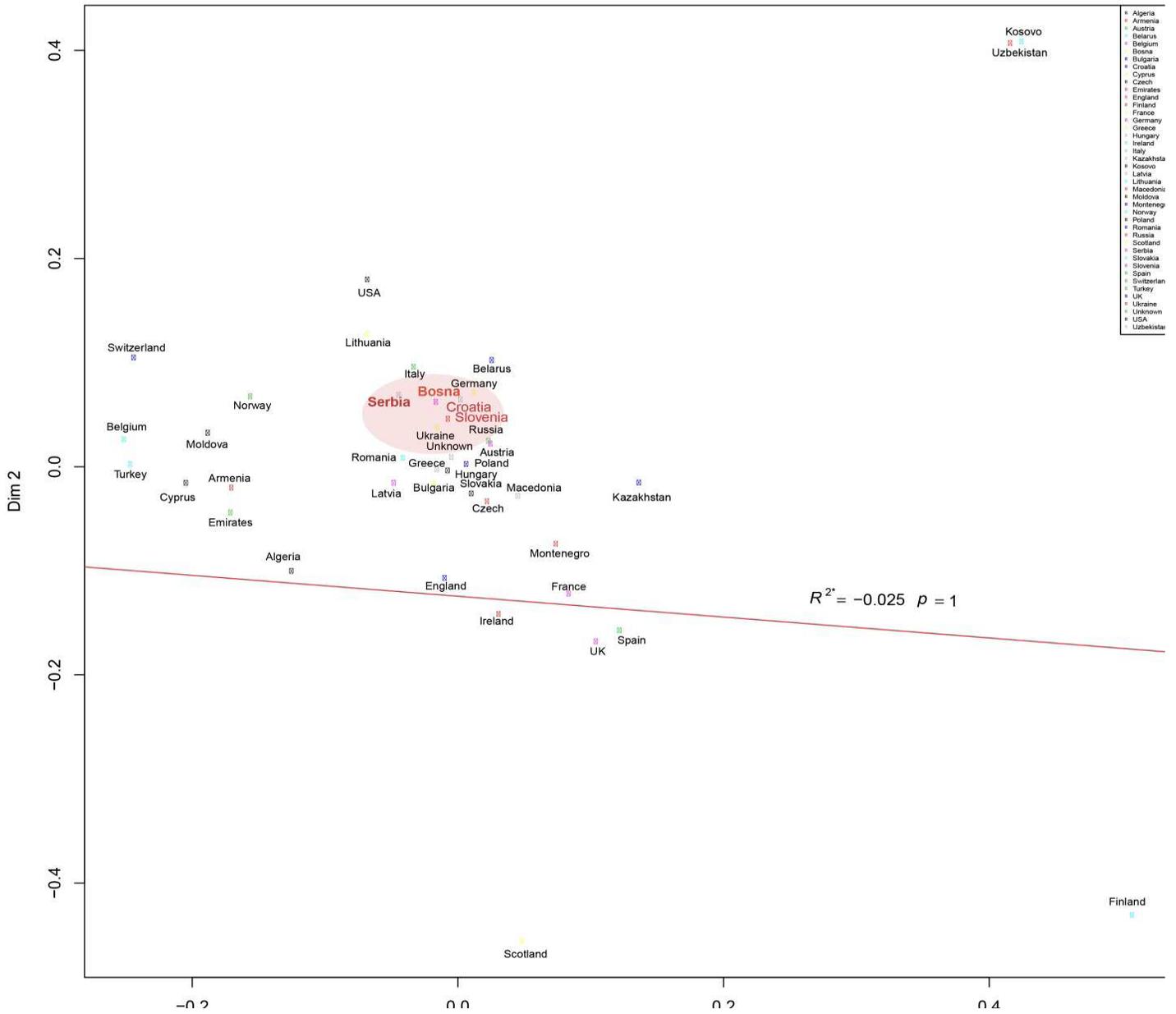


Рисунок 17. График многомерного шкалирования выборки l2a1b2a1, на котором представлено взаимное расположение этнических популяций в пространстве первых 2 измерений генетических дистанций Fst.

Cluster Dendrogram

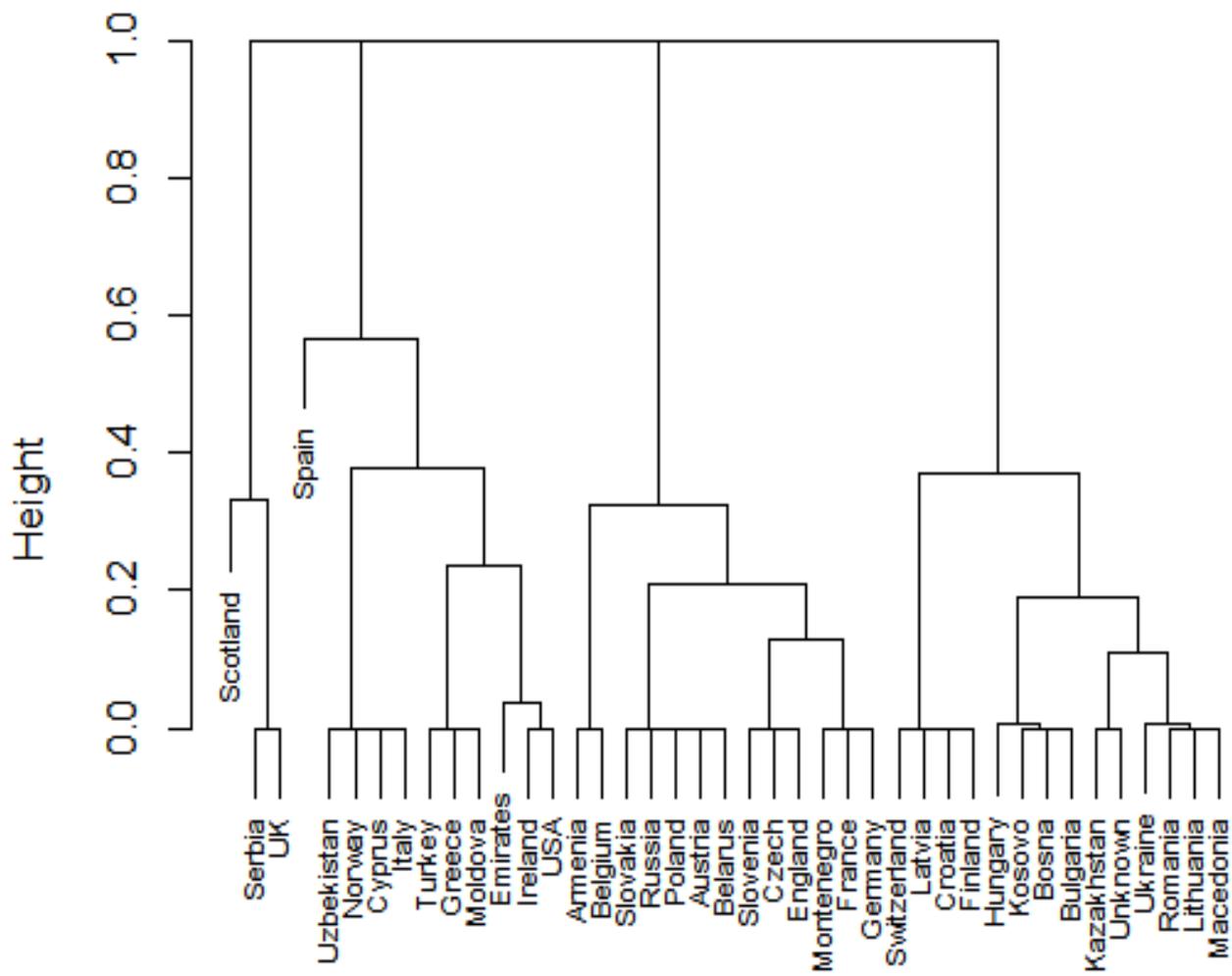


Рисунок 18. Кластерограмма этнических популяций, воссозданная на основе величин попарной дистанции между популяциями в матрице F_{st}

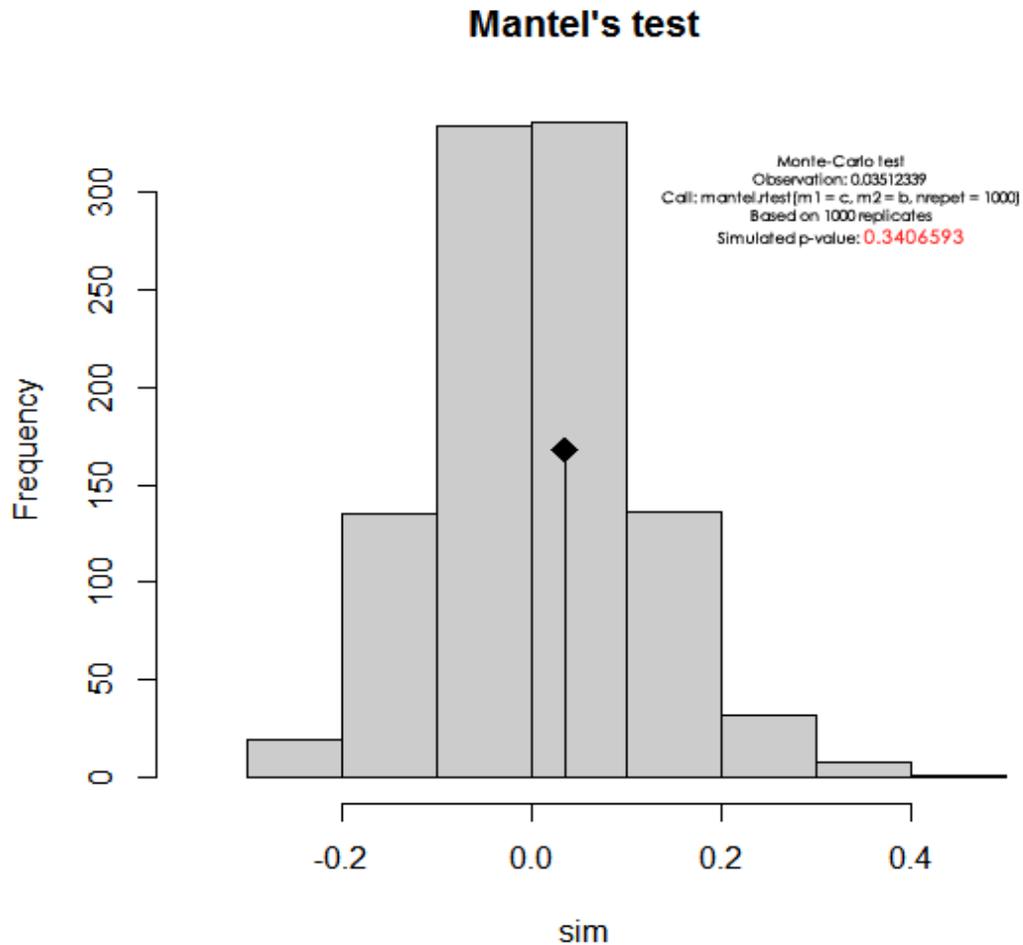


Рисунок 19. График результатов теста Мантеля на соответствие матриц генетической и географической дистанции при 1000 пермутаций по методу Монте-Карло. Низкое значение наблюдаемой корреляции (0.035) и значение $p > 0.05$ свидетельствуют о том, что между этими двумя матрицами отсутствует корреляция.

Выводы исследования

Перед тем как приступить к обсуждению результатов исследования, необходимо вкратце сформулировать в виде тезисов основные результаты анализов, проведенных в рамках системного метода:

1. Гаплогруппа I2a1b2a1 определяется снипами CTS176/S2621, CTS1293/S2632, CTS1802/S2638, CTS5375/S2679, CTS5985/S2687, CTS7218/S2702, CTS8239/S2715, CTS8486/S2722, CTS11030/S2768, L178/S328 (и пр.). Согласно данным экспериментальной филогении (экспериментальному дереву гаплогрупп) Yfull (2014) и ISOGG, эта клада разбивается на субклады новыми снипами I-S17250 (к этой субкладе принадлежат все из протестированных представителей т.н «южного динарского Y-STR кластера» и часть представителей т.н «северного динарского кластера», в то время как у остальной части обнаружено «предковое значение» аллели снипа I-S17250 –), I2a1b2a1b - Y4460 (Y3106, в эту кладу входят выходцы из Беларуси, Латвии, России, Польши), и Z17855 (в эту «безымянную» субкладу, еще не принятую в номенклатуре ISOGG, входит часть болгар, македонцев и украинцев).
2. Самый большую подгруппу образуют дочерние кластеры самой большой субклады I-S17250 (обозначенной в дереве ISOGG как I2a1b2a1a): I2a1b2a1a1-Z16971(Y5596,Y5595, обнаружена у представителей небольшого кавказско-

украинского кластера, чей возраст по расчетам дисперсии Y-STR составляет 1000 лет), I2a1b2a1a2-Y4882 (Y4883, представлена в выборке беларусов, украинцев и поляков), I2a1b2a1a3-A356/Z16983 (Y4790, Y4789, данные снипы найдены у части представителей описанного выше «южного-динарского» кластера). Данное распределение субклад гаплогруппы I2a1b2a1 показывает, что разнообразие субклад (ветвей) этой гаплогруппы гораздо выше в восточной Европе, где ареалы распространения дочерних субклад динарского субклада в значительной степени накладываются друг на друга. На Балканах и в южной Европе разнообразие дочерних субклад I2a1b2a1 ниже.

3. Исходя из имеющихся ныне фактов, можно сделать вывод о том, что так называемый северно-динарский Y-STR кластер I2a1b2a1 («Dinaric-N») соответствует той части родительской ветви I2a1b2a1-CTS5966, которая не входит в субкладу I2a1b2a1a3- Z16982/ Z16983/A356 (последняя включает значительную часть южно-динарского Y-STR кластера)⁸. У большинства представителей всей динарской гаплогруппы I2a1b2a1 обнаружены мутировавшие аллели в Y-снипах Y3548, S17250, и YP205. Следовательно, варианты Y-хромосомы гаплогруппы I2a1b2a1, не имеющие эти три снип-мутации, представляют собой самое раннее кладистическое разделение гаплогруппы I2a1b2a1-CTS5966, видимо незадолго до последующей популяционной экспансией. Прямая патрилинейная родословная носителей этих «предковых вариантов» прослеживается в регионы юго-восточной Польши и смежных регионов западной Украине. Кроме того, в ходе коммерческих исследований проекта **FTDNA I2a**, у одного из представителей субклада «Disles» (I2a1b2a-CTS10936) - ближайшего (и возможно «родительского») по отношению к корневому уровню «динарской субклады» были определены 2 предковых (-) значения снипов, играющих важное значение в кладистическом разделении «динарского субклада» - CTS10936 + CTS10228- CTS5966-. Примечательно, что этот мужчина по прямой линии тоже происходит из южной Польши. Подводя итог сказанному, можно заметить, что распространение гаплогруппы I2a1b2a1-CTS5966 началось в регионе западной Украины или юго-восточной Польше, где до сих пор сохранились «реликтовые» клады (ветви) I2a1b2a1-CTS5966 с «предковыми значениями» трех снипов.⁹
4. Таким образом, более раннее разделение субклады I2a1b2a1 по 2 DYS-локусам на два кластера – южной и северный - не отражает истинной филогении: по состоянию на конец 2014 года субклад разделяется на 2-3 группы с дальнейшим разбиением на подгруппы; однако в настоящий момент из-за недостатка статистических данных трудно сказать, какие из данных мутаций являются приватными (генеалогическими), а какие актуальны для этнопопуляционных построений.
5. Генография субклад I2a носит нон-инклюзивный характер по отношению других субклад I2a, иными словами ареалы распространения отдельных субклад I2a практически не пересекаются. Распространение I2a1b-M423 в восточной Европе (где представлена главным образом ее «динарская субклада» I2a1b2a1) характеризуется резким градиентом частот: частоты Y хромосом этой субклады резко уменьшаются по мере удаления к западу от Балкан – так, например, этот субклад практически отсутствует у итальянцев, немцев, французов, и швейцарцев. При движении на северо-восток уменьшение частот гораздо более плавное, I2a1b-M423 сохраняет заметные частоты среди населения, говорящего на славянских языках. Распространение частот гаплогруппы I2a1b-M423 в восточной Европе носит бимодальный характер – с максимальным пиком (30-50 %) на Балканах, и с менее выраженным пиком (20-30%) в карпатско-полесском регионе. Наибольшие частоты распространения I2a1b встречаются у жителей Боснии-Герцеговины и хорватов Далмации от 40 до 60%, у сербов и македонцев 20-30%, примерно столько же в Молдавии у гагаузов.
6. Благодаря находкам древней ДНК удалось уточнить место первоначального распространения гаплогруппы I2a1b M423 (северо-западная часть Европы), по крайней мере одной из ее древнейших исчезнувших ветвей, выделившейся из родительской гаплогруппы примерно 8700-9000 лет назад. Образец Лошбур (Loshbour) принадлежит к ветви, параллельной современным дочерним субкладам, которая отделилась от основной ветви не позднее чем 10 000 лет тому назад: к этой исчезнувшей ветви принадлежат и несколько образцов ДНК жителей шведской мезолитической стоянки Motala (Motala 3, Motala 12), которая существовала синхронно Лошбуру. Видимо, в этом же регионе и произошло выделение предковой ветви I2a1b2a1-L147.2/CTS5966, так как ближайшая к динарской субклада сестринская субклада I2a1b2a* (так называемый кластер Disles) встречается как в Польше, так и на британских островах, а следующая по удаленности клады (островная субклада – Isles- I2a1b1 (L161.1/S185)) встречается практически исключительно на британских островах (подобный изолированный характер может быть связан с обособлением этой ветви в мезолите как следствие гипотетического затопления Доггерланда около 8500 лет тому назад (Weninger 2008)). Кроме того, оба найденных (в захоронениях древних жителей Паноннской равнины) образцов ДНК гаплогруппы I, (образец NE7 - I2a2a-L1228 и образец KO1 – I2a-L68) принадлежат к параллельным ветвям, ни одна из которых не является предковой ветвью динарской субклады I2a1b2a1. Исходя из этого можно сделать осторожное предположение о том, что представители динарской субклады I2a1b2a1-L147.2/CTS5966 появились на Балканах гораздо позднее неолитического периода.

⁸ Кроме этих снипов, формирующих четко выраженные клады в структуре дерева I2a1b2a1, в одном из тестов Geno 2.0 у представителя этой гаплогруппы были обнаружены снип CTS8429, чье положение в структуре дерева I2a1b2a1 неизвестно. Стоит также упомянуть снип YP206 (находящийся под снипом S17250); зафиксированный в полных геномных данных двух сардинцев из работ Francalacci et al. (2013); а также одноуровневые снипу YP206 (M1345/CTS934).

⁹ В частной переписке К.Нордтведт высказал мнение о том, что первичный ареал распространения этой гаплогруппы находился в Польше, к югу от Вистулы.

7. Визуальное изучение структур минимального остовного и штейнровского деревьев филогенетической сети гаплотипов I2a1b2a1 показало, что большое скопление гаплотипов вокруг больших узлов обеих деревьев имеет типичную форму филогенетического старкстера. Подобная форма филогенетических кластеров в основном наблюдается в тех случаях, когда происходит быстрый демографический рост одной конкретной филогенетической линии, и этот рост приводит к появлению серии одновременных мутационных событий. Большинство из гаплотипов, входящий в большой узел графа, принадлежат к этнопулам жителей южной Европы (главным образом, Балкан). Исходя из этого можно предположить, что экспансия носителей I2a1b2a1-L147.2/CTS5966 на территории Балкан сопровождалась быстрым демографическим ростом популяции.
8. На реконструированной карте место «выделения» I2a1b - родительской популяции современных представителей субклада I2a1b2a1 - разместилось чуть южнее швейцарских Альп (обозначено большой красной точкой) (Рисунок 10), примерно в 420 километрах к юго-востоку от места обнаружения древнейшего (8000 лет до настоящего времени) образца ДНК мужчины, принадлежавшего к гаплогруппе I2a1b (Loschbour-Heffingen, Luxembourg) и примерно в 650 километрах к югу от Лихтенштейнской пещеры, в которой были обнаружены 4 скелета мужчин с древнейшими (3000 лет до настоящего времени, культура полей погребения) из найденных образцов гаплогруппы I2a2b, а также в 920 к юго-западу от местонахождения неолитического поселения Арс-Верекалја (Венгрия), где был обнаружен древнейший из найденных образцов I2a2a (возрастом в 6700 лет). Заметная близость реконструированного места выделения предковой линии I2a1b2a1 к местам нахождения древнейших образцов гаплогруппы I2a в Европе подтверждает правдоподобность подобной реконструкции.
9. При сопоставлении интервалов возрастов гаплогруппы (т.е. расчетных времен жизни ближайшего общего предка гаплогруппы) I2a1b2a1, мы получили среднее медианное значение интервала полученных возрастов. Оно составляет 2757 лет, стандартное отклонение 404 лет (2757 ± 404), верхний предел интервала ТМРСА перекрывается с интервалом ТМРСА, вычисленного с помощью χ^2 -статистики 3400 ± 200 лет. Примечательно также, что время расхождения Y-STR кластеров внутри филогенетического дерева лежит в интервале 1700-1300 год до настоящего времени. Это означает, что в этот период времени (т.е. между 4 и 7 веками нашей эры) гаплогруппа I2a1b2a1 находилась в активной стадии экспансии.
10. Безусловный пик-максимум распространения I2a1b2a1 приходится на Балканы (особенно на регион Боснии-Герцеговины), а один из пиков значений дисперсии (разнообразия) гаплотипов - на полесско-карпатский регион Восточной Европы. Исходя из общего правила о приоритете разнообразия (дисперсии) над частотой распространения генетических линий при определении «начальной точки экспансии», предпочтение было отдано дисперсии как более устойчивому индикатору, в то время, как частоты гаплогрупп больше подвергнуты флуктуациям.
11. Полученные этнопулы или этнические кластеры имеют специфический: например в выборке выделяется характерный еврейский кластер (украинские и белорусские евреи-представители этого кластера имеют редкие характерные значения Y-STR локусов: DYS537 = 11, DYS464a = 13, DYS456 = 14, DYS458 = 18, DYS576 = 19, DYS570 = 16) с самой высокой интерпопуляционной дистанцией от остальных кластеров и самым низким разнообразием азывает на недавнее происхождение кластера как следствие чистого эффекта основателя, имевшего место примерно 500 лет назад. Боснийско-герцеговинский кластер (второй после еврейского кластера по величине интерпопуляционной дистанции), но с более высоким уровнем разнообразия указывает на более удаленный по времени популяционный эффект основателя. В то время, как географически близкий к боснийско-герцеговинскому кластеру хорватский кластер I2a1b2a1 характеризуется относительно низким уровнем молекулярного разнообразия (многие хорватские гаплотипы I2a1b2a1 имеют характерный гаплотипный мотив DYS19=14), и незначительной дистанцией с другими популяциями, что отражается в более низком значении индекса стандартного разнообразия (что подразумевает изменение размера эффективной популяции). Данный эффект можно объяснить кумулятивным действием двух параллельных популяционных эффектов - эффекта основателя и эффекта «бутылочного горлышка».
12. Полученные в ходе анализа молекулярного разнообразия I2a1b2a1 данные свидетельствуют в пользу нашего предположения о том, что гаплогруппа I2a1b2a1 не является автохтонной гаплогруппой Балкан и Динарских Альп, в противном случае наблюдалась более существенная корреляция между популяцией носителей этой гаплогруппы и анротипом. Поскольку популяции I2a1b2a1 лучше коррелируют с языком, а не с анротипом, то можно сделать два вывода: 1) экспансия носителей субклада I2a1b2a1 произошла недавно, т.к. не утерялась связь представителей генетической линии с языком; 2) поскольку славянские языки были явно привнесены на Балканы, то нужно признать, что I2a1b2a1 были в числе генетических линий, представители которых привнесли славянские языки на Балканы.
13. По мнению авторов статьи (Boattini et al. 2011) генофонд (включая генофонд Y хромосом) изолированной популяции арберешей не только должен отражать структуру генетического пула алабанцев 500 летней давности, но и служить своего рода «эталоном» генофонда древних балканских популяций. Принимая во внимание низкую частоту встречаемости динарского субклада I2a1b2a1 у современных арбарешей, можно предположить, что субклада I2a1b2a1 встречалась в генофонде популяций южной части Балканского

полуострова 500 лет назад гораздо реже, чем сейчас. Очевидно, что в ходе дальнейших демографических процессов %-ная доля I2a1b2a1 в генофонде южнобалканских популяций значительно увеличилась.

14. Гипотетическое направление миграций можно восстановить с помощью проекции градиентов значений индекса молекулярного разнообразия на векторную карту со стрелками, отображающими направление уменьшения значений индекса молекулярного разнообразия. Примечательно, что стрелки направления уменьшения разнообразия гаплотипов в значительной мере перекрываются с маршрутами славянской экспансии.
15. Результаты теста Мантеля говорят о том, что в выборке I2a1b2a1 не наблюдается статистически значимой корреляции между географической и генетической дистанцией.

Обсуждение выводов исследования

Как представляется автору этой статьи, вопрос происхождения и миграции носителей «динарской субклады» I2a1b2a1 невозможно изучить вне интердисциплинарной подхода, т.е. подхода в котором тезисы результатов популяционно-генетического исследования будут рассмотрены в более широком ключе, с привлечением данных из исторических наук. В этой связи необходимо посмотреть на эту проблему глазами историка, тем более что время экспансии «динарцев» отлично накладывается на временной интервал экспансии славян на Балканы в «историческое время».

В силу фрагментарности обзора мы не будем касаться не менее важных вопросов о том, как и когда гаплогруппа I появилась в Европе, а также вопросов о времени и месте разделения базальной ветви I на гаплогруппы I1 и I2. В настоящее время мы не располагаем достаточным количеством фактов в пользу одной из многочисленных и равновероятных версий сценария, и поэтому просто обойдем эти вопросы стороны в надежде на предстоящее изучение древней ДНК жителей Европы среднего палеолита и мезолита. По имеющимся в настоящее время данным палеогенетики, гаплогруппа I (и одна из ее двух основных ветвей - I2a) была широко распространена в западной и северной Европе уже вскоре после окончания последнего ледникового максимума (последняя ледниковая эпоха закончилась между 15 000 и 10 000 годами до н. э., а древнейшие образцы палео-ДНК с гаплогруппой I2a1b датируются 6 тыс. до н.э.). Около 14 000 лет до настоящего времени из гаплогруппы I2a выделились I2a1a-M26/PF4056 (эта группа мигрировала к югу от швейцарских Альп в сторону Пиренейского полуострова, и I2a1b-M423, которая мигрировала на север вслед за отступающим ледником и тундровой фауной. Примерно 10 000 лет назад, где-то на территории современной Германии из I2a1b выделились линия I2a1b1 (кластер «Isles», мигрировавший на Британские острова) и линия I2a1b2 (динарский кластер и кластер Disles). Можно предполагать, что в последующее время I2a1b2 была связана со свидерской культурой - археологической культуры финального палеолита (9 – 8 тыс. до н.э.) на территории Центральной и Восточной Европы, которая была представлена стоянками тундровых охотников на северного оленя, которые использовали стрелы с кремневыми наконечниками. Гипотетическое развитие дальнейшей истории общности носителей I2a1b2 - непосредственных предков I2a1b2a1 – можно выстроить в виде цепочки приемственности культур вплоть до культур участвовавших в этногенезе славян, однако из-за недостаточных сведений о мужском палеогенофонде представителей этих культур данная реконструкция не может быть использована как основание для дальнейших заключений.

К сожалению, специальных популяционно-генетических исследований гаплогруппе I крайне мало. Достаточно сказать, что последние специальные исследования по этой гаплогруппе были опубликованы почти 10 лет назад в 2004, 2006, 2007 годах – (Rootsi et al. 2004; Roewer et al. 2005; Underhill et al. 2007). Остальные работы, особенно касающиеся Y-хромосомного разнообразия популяций на территории Балкан и бывшей Югославии, преимущественно повторяют выводы озвученные в вышеупомянутых трех работах, ничего не добавляя от себя (Marjanović et al. 2005; Perićić et al. 2005; Rebała et al. 2007). Более новые работы со специфическим фокусом на гаплогруппе I1 (Shtrunov 2010) и (De Beule 2010) были незаслуженно обойдены вниманием научного сообщества, так как были написаны «любителями» и опубликованы в неакадемических изданиях. Поэтому именно работы эстонской исследовательницы-популяционного генетика Рутси задали целый ряд основных положений-гипотез по вопросу происхождения гаплогруппы I, закрепившихся позднее в научных и научно-популярных (Википедия) источниках. Со временем данное положение стало общим местом статей популяционных генетиков при описании особенностей генофонда различных популяций. Если на заре популяционной генетики, выводы делались лишь на основании распределения частот гаплогрупп (большой группы схожих гаплотипов) Y хромосом, то с накоплением фактического материала и совершенствованием методов исследования появилась возможность кроме частоты учитывать молекулярное разнообразие и генетические дистанции. Уже без этих исследований выводы на основании только частот являются неполными, а часто и совершенно неверными.

В настоящее время, исходя из описаного выше опыта полисистемного анализа и синтеза молекулярно-генетических (филогенетических и статистических), исторических, археологических данных, а также данных археогенетики (древняя ДНК), можно предположить динарская субклада I2a1b2a1 в массовом количестве появились на Балканах не ранее 2-3 в.н.э и не позднее 6-7 в.н.э., что точно соответствует времени великого переселения народов. В числе значительных миграций населения на Балканы в этом период времени можно назвать миграцию готов, гепидов с севера, а также более позднюю по времени экспансии славян (вернее, генофонда общности, которая говорила на славянском или протославянском языках), которые и принесли с собой этот субклад на Балканы. Кроме этих двух вариантов, можно рассмотреть и третий вариант, согласно которому субклад I2a1b2a1 уже присутствовал на генофонде балканских популяций на момент начала переселения народов (эта гипотеза предполагает что первоначально популяционная общность носителей I2a1b2a1 была так или иначе с дако-фракийским кругом археологических культур). В статье 2010 года (Носевич 2010), Вячеслав Носевич напрямую связывает носителей I2a1b2a1 с фракийским субстратом в формировании славян, отмечая что более высокая концентрация I2a1b2a1 на Полесье по сравнению с Западной Украиной и Словакией говорит в пользу это предположения. Традиционно считается, что вся гаплогруппа I2a связана с кругом культур балканского неолита и в том числе, с культурами близкими к трипольцам. Такой вариант возможен, однако окончательный вариант должны дать палеогенетического исследования останков представителей этих культур. Но, если гаплогруппа I2a и присутствовали у трипольцев, их вряд ли можно считать коренным элементом (ядро которого составляли выходцы с Ближнего Востока), скорее всего они представляли остатки местного мезолитического элемента, инкорпорированного в состав этой культуры.

При взвешенном сопоставлении эмпирических статистических данных о характере распространения и разнообразия I2a1b2a1 с различными моделями миграций тезис о связи доисторических носителей I2a1b2a1 с фрако-дакийцами представляется несколько натянутым. Этот тезис завязан целиком на интерполяции современных частот гаплогрупп на гаплогруппные частоты в прошлом — однако манипуляции с современными частотами в целях реконструкции гипотетических частот распространения в генофонде древних народов всегда вызывают закономерные вопросы. По крайней мере, подобные эксперименты возможны только исходя из генетических данных полученных в результате анализа останков из захоронений соответствующих культур или народностей. В противном случае — расчеты могут оказаться совершенно произвольны, и им нельзя будет доверять. Однако даже если мы будем из реконструкции предковых частот, то увидим, что I2a1b2a1 вряд ли могла присутствовать в столь значительных количествах у фракийцев, так как и у современных «потомков фракийцев и даков» (болгаров и румын) как частоты распространения, так и уровень дисперсии гаплотипов I2a1b значительно ниже чем у тех же словаков и западных украинцев. Далее, «балкано-иллирийско-фракийская гипотеза» совершенно не объясняет высокие частоты т.н. «динарской субклады» I2a1b2a1 в Полесье. Не объясняет эта теории и того, что филогенетически I2a1b2a1 не образует удалённых кластеров, что невозможно объяснить допуская автохтонность носителей этой гаплогруппы на Балканах. Ведь за тысячи лет эта линия должна была разветвиться. Но сегодняшние её носители исключительно близки друг к другу и подавляющее их число - славяне.

Тезис о связи носителей I2a1b2a1 с германских субстратом в этногенезе славян также вызывает определенные вопросы. Можно предположить, что присутствие носителей I2a1b2a1 вполне может быть связано как с особенностью ранних (прото-)славянских популяций, так и с непосредственной инвазией северных германцев (например, готов). По крайней мере сейчас уже доказано, что более 8 тысяч лет назад параллельная ветвь носителей I2a1b-M423 проживали в том числе на юге Скандинавии, а зафиксированные в историческое время (эру великого переселения народов) места проживания готов в Восточной Европы (королевство остготов на Балканах и «черняховская» культура в западной Украине) По крайней мере, среди современных этнических скандинавских германцев гаплогруппа I2a1b2a1 не встречается, также как и в Испании, Италии (где находились королевства готов в раннем средневековье), зато с заметной частотой встречается по всему славянскому ареалу :очень высокие концентрации действительно отмечаются на Балканах, но и у западных славян присутствие его значительно, 8-10% у чехов и поляков, до 20% у словаков).

Довольно высокие частоты распространения и высокий уровень внутригаплогруппного разнообразия, специфическая топология филогенетических деревьев гаплогруппы I2a1b2a1, а также факт нахождения «реликтовых вариантов» гаплогруппы I2a1b2a1 в Польше и Западной Украины свидетельствует о формировании предкового пула этой субклады в северо-восточной Европе в 1 тыс. до нашей эры и экспансии в 1 тыс. нашей эры. В свете представленных в виде тезисов основных выводов исследования, наиболее предпочтителен вариант распространения гаплогруппы I2a1b2a1 на Балканах и всей восточной Европе во время славянской экспансии. Новизна предложенного нами варианта заключается в том, что коренным образом противоречит принятому в официальной популяционной генетике представлению о Балканах, как «прародине» популяции I2a2b (см. раздел (Гаплогруппа I2a общие сведения и обзор публикаций)). Обычно в качестве одного из главных аргументов против версии распространения гаплогруппы I2a1b2a1 вместе с миграциями славянской выдвигается тезис о однородности состава гаплогрупп у ранних славян (предполагается, что они принадлежали к различными субкладам R1a1-Z283, (Rebala et al. 2007) особенно R1a1-M458 (Underhill et al. 2009), современный эпицентр распространения которой приходится на ареал пшеворской культуры). Данный аргумент не может быть воспринят так как последние исследования генофонда неолитических культур Европы показали, что уже генофонд этих ранних культур был далеко неоднородны, а в бронзовом веке и позднее степень смешения должна была только увеличиться хотя бы в силу большей мобильности населения. Более вероятным представляется

сценарий, в котором уже ранняя славянская общность (которая определяется лингвистическими и археологическими, а не генетическими признаками) была достаточно гетерогенна и включала в себя, наряду с типичными «славянскими» субкладами R1a1-Z283 (прежде всего дочерних субклад R1a1-M458 и R1a1-Z280), и субклады других гаплогрупп, в частности и I2a1b2a1. В этом случае ранние славяне образовались в результате «наслоения» R1a-Z280 и/или R1a-M458 на I2a1b2a1. Эта (одна из возможных) модель этногенеза соответствует одной из существующих моделей происхождения славянских языков: неиндоевропейский субстрат оторвал балтов от славян и дал отчёт их самостоятельному существованию в истории как языковой группе, так и целому археологическому ряду культур связанных с ними.

Ссылки

А. Публикации

Bandelt H.J., Forster P., Röhl A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biological Evolution* 16:37-48.

Barac L., Pericic M., Klaric I.M., Rootsi S., Janicijevic B., Kivisild T., Parik J., Rudan I., Villems R., Rudan P. (2003). Y chromosomal heritage of Croatian population and its island isolates. *European Journal of Human Genetics* 11:535–542-

Bastian M., Heymann S., Jacomy M. (2009). Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*.

Battaglia V., Fornarino S., Al-Zahery N., Olivieri A., Pala M., Myres N.M., King R.J., Rootsi S., Marjanovic D., Primorac D., Hadziselimovic R., Vidovic S., Drobic K., Durmishi N., Torroni A., Santachiara-Benerecetti A.S., Underhill P.A., Semino O. (2009). Y-chromosomal evidence of the cultural diffusion of agriculture in Southeast Europe. *European Journal Human Genetics* 17(6): 820 - 830

Berger B., Lindinger A., Niederstatter H., Grubwieser P., Parson W. (2005) Y-STR typing of an Austrian population sample using a 17-loci multiplex PCR assay. *International Journal of Legal Medicine* 119:241–246.

Bosch E., Calafell F., Comas D., Oefner P.J., Underhill P.A., Bertranpetit J. (2001). High-resolution analysis of human Y chromosome variation shows a sharp discontinuity and limited gene flow between Northwestern Africa and Iberian Peninsula. *American Journal of Human Genetics* 68:1019–1029.

Capelli C., Redhead N., Abernethy J.K., Gratrix F., Wilson J.F., Moen T., Hervig T., Richards M., Stumpf M.P., Underhill P.A., Bradshaw P., Shaha A., Thomas M.G., Bradman N., Goldstein D.B. (2003). A Y chromosome census of the British Isles. *Current Biology* 13:979–984.

Chandler, John F. (2006). Estimating Per-Locus Mutation Rates. *Journal of Genetic Genealogy, A Free Open-Access Journal* 2(2): 27-34.

Cinnioglu C., King R., Kivisild T., Kalfoglu E., Atasoy S., Cavalleri G.L., Lillie A.S., Roseman C.C., Lin A.A., Prince K., Oefner P.J., Shen P., Semino O., Cavalli-Sforza L.L., Underhill P.A. (2004). Excavating Y-chromosome haplotype strata in Anatolia. *Human Genetics* 114: 127–148.

De Beule, H. (2010) Origin, Distribution and Migrations of I2b*-Subclades. *The Russian Journal of Genetic Genealogy* 1(2):12:13

Djindjian F. (2000). The Mid Upper Palaeolithic (30,000 to 20,000 bp) in France. In: Roebroeks W., Mussi M., Svoboda J., Fennema K. (eds.). *Hunters of the golden age: the Mid Upper Palaeolithic of Eurasia 30,000–20,000 BP: Analecta Praehistorica Leidensia*. Leiden: University of Leiden, 313–324.

Excoffier L., Smouse P.E., Quattro J.M. (1992) Analysis of molecular variance inferred from metric distance among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 131: 479–491.

Excoffier L., Lischer H.E.L. (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564-567.

Flores C., Maca-Meyer N., Gonzalez A.M., Oefner P.J., Shen P., Perez J.A., Rojas A., Larruga J.M., Underhill P.A. (2004). Genetic structure of Iberian Peninsula revealed by Y chromosome analysis. *European Journal of Human Genetics* 12(10):855-63.

Forster P., Torroni A., Renfrew C., Röhl A. (2001). Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. *Molecular Biology and Evolution* 18:1864-1881.

Gamba C., Jones E.R., Teasdale M.D. et al. (16 co-authors) Genome flux and stasis in a five millennium transect of European prehistory. *Nature Communications* 5, Article number: 5257 doi:10.1038/ncomms6257.

Garza J.C., Williamson E.G. (2001). Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*. 10: 305–318.

Goloboff P.A., Farris J.S., Nixon K.C. (2008). TNT, a free program for phylogenetic analysis. *Cladistics* 24: 774–786.

Hadley K., Myres N.M., Ekins J.B., Ekins J.E., Hutchinson L.A.D., Lunt, M.L., Masek S.S., Nelson A.A., Nelson M.E., Pennington K.L., Perego U.A., Peterson J.L., Sims A., Tolley T., Welch A., Woodward S. (2004) Non-paternity and Locus Specific Mutation Rates of 36 Y Chromosome STRs. Presentation on The American Society for Human Genetics 54th Annual Meeting.

- Karachanak S., Grugni V., Fornarino S., Nesheva D., Al-Zahery N., et al. (12 co-authors) (2013) Y-Chromosome Diversity in Modern Bulgarians: New Clues about Their Ancestry. *PLoS ONE* 8(3): e56779. doi:10.1371/journal.pone.0056779
- Lazaridis, I., Patterson, N., Mittnik, A. et al. (122 co-authors) (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* 513: 409–413.
- Little, L. (2007). Y-STR mutation rates. <http://freepages.genealogy.rootsweb.ancestry.com/~geneticgenealogy/> Accessed online 9.12.2014.
- Ljubković J., Stipisić A., Sutlović D. et al. (2008). Y-chromosomal Short Tandem Repeat Haplotypes in Southern Croatian Male Population Defined by 17 Loci. *Croatian Medical Journal* 49: 201–206.
- Maca-Meyer N., Sanchez-Velasco P., Flores C., Larruga J.M., Gonzales A.M., Oterino A., Leyva-Cobian F. (2003). Y chromosome and mitochondrial DNA characterization of Pasiegos, a human isolate from Cantabria (Spain). *Annals of Human Genetics* 67:329–33.
- Maddison D.R., Maddison W.P. (2014a). Cartographer, a Mesquite package for plotting geographic data. Version 1.41. <http://mesquitemproject.org/packages/cartographer>
- Maddison, D.R., Maddison, W.P. (2014b). Mesquite: a modular system for evolutionary analysis. Version 3.01 <http://mesquitemproject.org>
- Mirabal S., Varljen T., Gayden T. et al. (2010). Human Y-chromosome short tandem repeats: A tale of acculturation and migrations as mechanisms for the diffusion of agriculture in the Balkan Peninsula. *American Journal of Physical Anthropology*. 142(3):380–390.
- Nei M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Nordtvedt K. (2008). More Realistic TMRCA Calculations. *Journal of Genetic Genealogy, A Free Open-Access Journal*. 4(2): 96-104.
- Peričić M., Barać L., Martinović L., Klarić I., Rootsi S., Janičijević B., Rudan I., Terzić R., Čolak I., Kvesić A., Popović D., Šijački A., Behluli I., Đorđević D., Efremonska L., Bajec Đ.D., Stefanović B.D., Villems R., Rudan P. (2005). High-resolution phylogenetic analysis of southeastern Europe traces major episodes of paternal gene flow among Slavic populations. *Molecular Biology and Evolution* 22:1964-1975.
- Perles C. (2000). Greece, 30,000–20,000 bp. In: Roebroeks W., Mussi M., Svoboda J., Fennema K. (eds.). *Hunters of the golden age. The Mid Upper Palaeolithic of Eurasia 30,000–20,000 BP: Analecta Praehistorica Leidensia*. Leiden: University of Leiden, 375–398.
- Polzin T., Daneschmand S. V. (2003). On Steiner trees and minimum spanning trees in hypergraphs. *Operations Research Letters* 31:12-20.
- Rebala K., Mikulich A.I., Tsybovsky I.S., Siváková D., Džupinková Z., Szczerkowska-Dobosz A., Szczerkowska Z. (2007). Y-STR variation among Slavs: evidence for the Slavic homeland in the middle Dnieper basin. *Journal of Human Genetics* 52 (5): 406–414.
- Regueiro M., Rivera L., Damjanovic T., Lukovic L., Milasin J., Herrera R.J. (2012). High Levels of Paleolithic Y-Chromosome Lineages Characterize Serbia. *Gene* 498(1): 59–67.
- Roewer L., Croucher P.J.P., Willuweit S., Lu T.T., Kayser M., Lessig R., de Knijff P., Jobling M.A., Tyler-Smith C., Krawczak M. (2005). Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. *Human Genetics* 116:279-291.
- Roewer L., Willuweit S., Krüger C., et al. (2008) Analysis of Y chromosome STR haplotypes in the European part of Russia reveals high diversities but non-significant genetic distances between populations. *International Journal of Legal Medicine*. 122(3):219-223.
- Rootsi, S., Magri, C., Kivisild, T. et al. (45 co-authors) (2004). Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *American Journal of Human Genetics* 75:128–137.
- Rootsi S.; Zhivotovsky L.A., Baldovič M., Kayser M., Kutuev I. A., Khusainova R., Bermisheva M. A, Gubina M. et al. (2006). A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *European Journal of Human Genetics* 15 (2): 204–11.
- Semino O., Passarino G., Oefner P.J., Lin A.A., Arbuzova S., Beckman L.E., De Benedictis G., Francalacci P., Kouvatsi A., Limborska S., Marcikiae M., Mika A., Mika B., Primorac D., Santachiara-Benerecetti A.S., Cavalli-Sforza L.L., Underhill P.A. (2000). The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* 290:1155–1159.
- Sengupta S., Zhivotovsky L.A., King R. et al. (2006). Polarity and temporality of high resolution Y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of central Asian pastoralists. *American Journal Human Genetics* 78(2): 202-221.
- Shtrunov A. (2010) The origin of haplogroup I1-M253 in Eastern Europe. *The Russian Journal of Genetic Genealogy*. 1(2):1:11.
- Smouse P.E., Long J.C, Sokal R.R. (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *System Zoology* 35: 627-632.
- Tajima F. (1993). Simple methods for testing molecular clock hypothesis. *Genetics* 135:599-607.
- Teacher A.G.F, Griffiths D.J. (2011) HapStar: Automated haplotype network layout and visualisation. *Molecular Ecology Resources* 11(1): 151-153.
- Underhill P.A., Myres N.M., Rootsi S., Chow C.T., Lin A.A., Ollilar R.P., King R., Zhivotovsky L.A., Balanovsky O., Pshenichnov A., Ritchie K.H., Cavalli-Sforza L.L., Kivisild T., Villems R., Woodward S.R. (2007). New Phylogenetic Relationships for Y-chromosome Haplogroup I: Reappraising its Phylogeography and Prehistory. In: Mellars P., Boyle, K., Bar-Yosef O., Stringer C. (eds.). *Rethinking the Human Evolution*,

Underhill P.A., Myres, N.M., Rootsi S., Metspalu M., Zhivotovsky L.A., King R.J., Lin A.A., Chow C.E., Semino, O., Battaglia, V., Kutuev I., Järve M., Chaubey G., Ayub, Q., Mohyuddin A., Mehdi S.Q., Sengupta S., Rogaev E.I., Khusnutdinova E.K., Pshenichnov A., Balanovsky O., Balanovska E., Jeran N., Augustin D.H., Baldovic M., Herrera R.J., Thangaraj K., Singh V., Singh L., Majumder P., Rudan P., Primorac D., Villemes R., Kivisild T. (2010). Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. *European Journal of Human Genetics*, 18(4), 479 - 484.

Weir B. S., Cockerham C. C. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution* 38 (6): 1358.

Weir, B. S. (1996). *Genetic Data Analysis II: Methods for Discrete Population Genetic Data*. Sinauer.

Willuweit S., Roewer R. (2007) Y chromosome haplotype reference database (YHRD): Update. *Forensic Science International: Genetics*.

Wilson I.J., Weale M.E., Balding D.J. (2003). Inferences from DNA data: population histories, evolutionary processes and forensic match probabilities. *Journal of the Royal Statistical Society: Series A* 166: 155-188.

Zerjal T., Xue Y., Bertorelle G., Wells R.S., Bao W., Zhu S., Qamar R., Ayub Q., Mohyuddin A., Fu S., et al. (2003). The genetic legacy of the Mongols. *American Journal of Human Genetics* 72: 717-721

Животовский Л.А. (2006) Микросателлитная изменчивость в популяциях человека и методы ее изучения. *Вестник ВОГиС* 10(1): 73-96.

Монгайт А.Л. (1973). *Археология Западной Европы. Т. 1. Каменный век*. Москва: Наука.

В Базы данных и ДНК проекты

ISOGG-2014. http://www.isogg.org/tree/ISOGG_YDNA_Tree_Trunk.html Accessed 09.12.2014

FTDNA Project I2a <https://www.familytreedna.com/public/I2aHapGroup> Accessed 09.12.2014

HGDP Stanford Human Genome Diversity Project. <http://www.hgsc.org/hgdp/files.html> Accessed 09.12.2014

YHRD Y chromosome haplotype reference database. <http://yhrd.org/> Accessed 09.12.2014

YDNA predictor. <http://predictor.ydna.ru/> Accessed 09.12.2014

Ysearch. A Free Public Service from Family Tree DNA. www.ysearch.org Accessed 09.12.2014

YFull Experimental YTree v2.29. <http://www.yfull.com/tree/I-L621/> Accessed 09.12.2014

Референтная база данных частот встречаемости гаплотипов Y-STR локусов в Беларуси: <http://dnkexpertiza.org/index1.html> Accessed 09.12.2014

С Программы

Golden MapViewer 7. <http://www.goldensoftware.com/products/mapviewer> Accessed 09.12.2014

Fluxus Network: Free Phylogenetic Network Software. <http://www.fluxus-engineering.com/sharenet.htm> Accessed 09.12.2014