

Variabilidade do cromossoma Y na população do Distrito de Coimbra: tentativa de pesquisa da herança genética de eventos históricos

RUI MARTINIANO¹ / LICÍNIO MANCO¹ / AUGUSTO ABADE^{1,2}

1- Centro de Investigação em Antropologia e Saúde (CIAS), Departamento de Ciências da Vida, Universidade de Coimbra

2 - Centro de Investigação Transdisciplinar Cultura, Espaço e Memória

Introdução

A diversidade genética da população Portuguesa, assim como da Península Ibérica, é o resultado de muitos processos migratórios ocorridos num passado mais ou menos longínquo por parte de diferentes grupos populacionais. Os grandes contributos parecem ser os mais antigos, ocorridos na pré-história, nomeadamente o processo de povoamento na era Paleolítica, aquando da chegada Homem Moderno que colonizou a Europa há cerca de 40000 anos a partir do Médio Oriente (progressivamente substituindo as populações de Neandertais, que viriam a extinguir-se precisamente no oeste e sul da Ibéria), a que se seguiram migrações no sentido Oeste - Leste com o repovoamento após o Último Máximo Glacial (LGM) há cerca de 15000 anos a partir dos refúgios glaciares localizados na Península Ibérica, Itália e Balcãs. Na era Neolítica, o aparecimento da agricultura originou a chamada difusão démica a partir do Médio Oriente, uma nova migração de Este para Oeste associada à agricultura que na Península Ibérica está documentado há cerca de 7000-6000 anos (Pereira e Ribeiro, 2009).

Em tempos mais recentes, outros eventos demográficos importantes na Península Ibérica incluíram a ocupação Romana e as subsequentes invasões Germânicas após o colapso do Império Romano do Ocidente. Em 711 a Península Ibérica foi invadida por exércitos Árabes a partir do Norte de África, compostos na sua maioria por Berberes, que permaneceram em Portugal por 5 séculos até à sua expulsão final em 1249 e em Espanha até 1492 (Saraiva, 1992). Antes da invasão Árabe, a população da Península Ibérica rondaria os 7-8 milhões de habitantes e a população encontrada pelos invasores islâmicos, ao contrário de ser uniforme em termos religiosos e culturais, era constituída por uma maioria cristã mas também por uma minoria Judaica. Estes e os seus descendentes eram chamados de Judeus Sefarditas, que terão chegado durante a era Romana, voluntariamente ou enquanto escravos provenientes do Médio Oriente após a derrota de Judeia no ano 70 (Saraiva, 1992). Para a diversidade do património genético

em Portugal, não será também de desprezar o contributo das migrações Fenícias, provenientes da costa Este do Mediterrâneo, da região correspondente aos actuais países Líbano e Sul da Síria, que dominaram este mar durante o primeiro milénio AC, muito antes das invasões anteriormente referidas. Este povo estabeleceu um alargado sistema de trocas comerciais através da criação de colónias e feitorias em pontos estratégicos, nomeadamente no Norte de África (Cartago), Sicília e sul da Península Ibérica (Zalloua *et al*, 2008). Ainda mais recentemente, entre os séculos XV e XVIII, o estabelecimento de uma rede de tráfego comercial de escravos, trouxe um aumento acentuado na entrada de indivíduos subsarianos em Portugal, que chegaram a atingir 10% da população no sul do país em meados do século XVI (Saraiva, 1992).

Na reconstrução da origem e história evolutiva das populações, os dados linguísticos, antropológicos e arqueológicos eram até há bem pouco tempo as únicas fontes disponíveis. Com o desenvolvimento da genética molecular, surge uma nova ferramenta de pesquisa: os polimorfismos de DNA. Marcadores genéticos autossómicos, do DNA mitocondrial e do cromossoma Y têm sido largamente utilizados nos últimos anos para dedução de afinidades genéticas entre as populações humanas, com o objectivo de compreender e determinar as origens e a história demográfica dos diferentes grupos populacionais.

Recorrendo a diversos marcadores bialélicos (substituições de um nucleótido ou inserções/delecções de sequências) localizados na região não recombinante do cromossoma Y (Y-SNPs), é possível agrupar os indivíduos que partilhem as mesmas mutações derivadas num mesmo haplogrupo, isto é, num grupo de haplótipos que define uma mesma linhagem do cromossoma Y. Os haplogrupos principais do cromossoma Y são identificados por letras do abecedário desde o A até ao T, subdivididos num ou mais ramos possíveis de definir por uma sequência de marcadores bialélicos. Adicionalmente, a utilização de marcadores multialélicos, também chamados de microssatélites ou STRs (*Short Tandem Repeats*), que consistem basicamente em pequenas sequências repetitivas múltiplas com um número de repetições que varia de indivíduo para indivíduo, permitem definir haplótipos e medir a variação genética acumulada no interior de um haplogrupo, possibilitando também a capacidade de informar quer acerca do local e data provável da sua origem.

Em Portugal, têm sido conduzidos diversos trabalhos com polimorfismos do cromossoma Y, que têm contribuído para ajudar a compreender a história demográfica da população Portuguesa na sua componente masculina, num passado mais ou menos

recente (Beleza *et al.* (2006; Gonçalves *et al.*, 2005; Adams *et al.* 2008). Num estudo recente, Beleza *et al.* (2006) estudou uma amostra populacional de indivíduos naturais do distrito de Coimbra. Há que considerar, no entanto, as limitações deste estudo no distrito devido ao reduzido número da amostragem (apenas 20 indivíduos).

Neste sentido, e com o objectivo principal de apurar a prevalência das principais linhagens masculinas que possam estar associados a eventos históricos e demográficos de que a região foi alvo no passado, procedemos ao estudo de uma maior amostra populacional de indivíduos naturais do distrito de Coimbra recorrendo a marcadores bialélicos e a microssatélites localizados no cromossoma Y.

Metodologia

Foi estudada uma amostra populacional de 92 indivíduos do sexo masculino, sem relações de parentesco entre si, naturais do distrito de Coimbra e com ascendência paterna proveniente do distrito no mínimo até à 3^a geração. Após consentimento informado, foram colhidas amostras de células do epitélio bucal dos indivíduos voluntários e a extracção do DNA genómico foi feita recorrendo metodologias estandardizadas.

Para identificação dos haplogrupos mais frequentes na população Portuguesa foram estudados 12 marcadores bialélicos diferentes (M9, M269, M304, M269, M35, M78, M81, M123, M213, M22, M201 e M170) por polymerase chain reaction-restriction fragment length polymorphism (PCR_RFLP) ou por sequenciação automática num sequenciador automático ABI Prism 310 (Applied Biosystems), utilizando *primers* previamente descritos (Underhil *et al.*, 1997; Cruciani *et al.*, 2002; Cinioglu *et al.*, 2004; Neto *et al.*, 2007; Brion *et al.*, 2004). Foram estudados 7 microssatélites (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393) nos haplogrupos mais frequentes da população utilizando primers marcados com Cy5 e o sequenciador ALF Express II (Amersham Biosciences). A sequência dos primers utilizados foi obtida na página http://www.cstl.nist.gov/div831/strbase/ystr_fact.htm.

Para a análise estatística dos dados obtidos utilizou-se o software Arlequin ver. 3.1 (Excoffier *et al.* 2005) no que diz respeito aos cálculos de frequências alélicas e diversidade genética. As *networks* dos haplótipos utilizando os STRs, foram construídas usando o programa NETWORK 4.0 com o algoritmo *median-joining* (Bandelt *et al.*, 1999).

Resultados e Discussão

A Figura 1 apresenta a árvore filogenética dos polimorfismos bialélicos analisados neste trabalho, com os diferentes haplogrupos identificados na amostra populacional do distrito de Coimbra e respectivas frequências. O nó F-M213, que inclui os haplogrupos G, I, J, K, L, R, atinge uma frequência de 84,8%. O haplogrupo R1b1b2-M269, com uma frequência de 55,4%, surge como o mais comum, seguido dos haplogrupos J-M304 e E1b1b1-M35, com frequências de 16,3% e 15,2%, respectivamente.

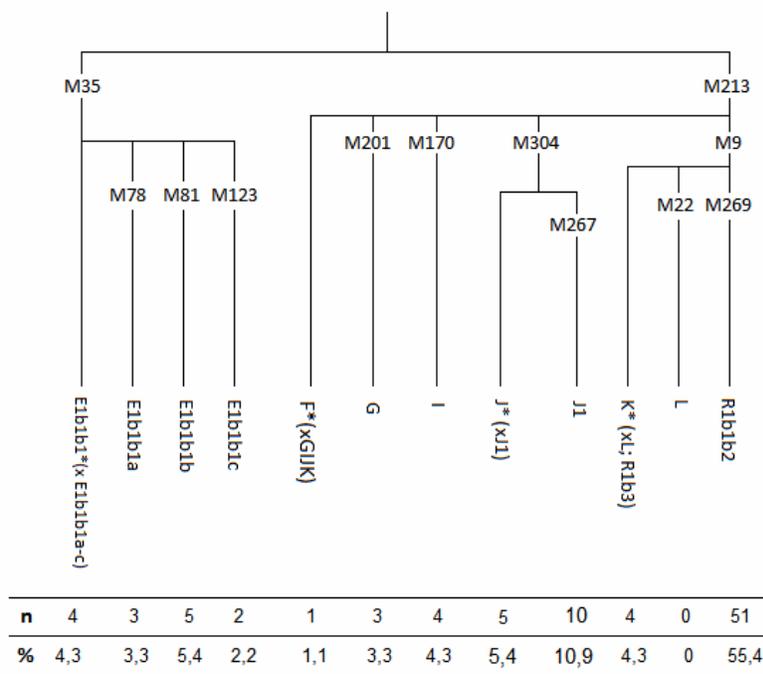


Figura 1 – Frequência de haplogrupos identificados na amostra populacional do Distrito de Coimbra.

As frequências de haplogrupos obtidas são, em termos gerais, coerentes com as verificadas por outros estudos realizados anteriormente, quer para o conjunto da população Portuguesa, quer para o distrito de Coimbra em particular (Tabela 1).

Tabela 1 - Frequências de haplogrupos mais comuns identificados em diversas regiões portuguesas e na Península Ibérica.

Pop	n	E1b1b1				F (xGIJK)	G	I	J(xJ1)	J1	R1b1b2	Ref
		(xa,b,c)	a	b	c							
D Cbra	92	4,3	3,3	5,4	2,2	1,1	3,3	4,3	5,4	10,9	55,4	[1]
D Cbra	20	5,0	0	0	5,0	0	5,0	5,0	10,0	15,0	50,0	[2]
RC*	116	0,9	2,6	5,2	2,6	0	2,6	3,4	6,0	6,0	60,3	[2]
PC	102	0	5,8	4,9	3,9	0	7,8	4,9	2,9	4,9	60,7	[3]
PN	101	4,9	7,9	5,9	-	1	4,9	6,9	3,9	1	52,5	[3]
PS	100	1	1	6	-	1	7	7	6	7	52	[3]
PN	60	2	5	3	2	-	12	2	7	2	59	[4]
PS	78	3	4	8	1	-	9	4	15	3	47	[4]
PIB	1140	1	4	4	1	-	5	6	7,7	1,2	65,2	[4]

D Cbra: Distrito de Coimbra; RC: região centro (Distritos de Coimbra, Viseu e Aveiro); PC: Portugal Centro; PN: Portugal Norte; PS: Portugal Sul; PIB: Península Ibérica.

Referências: [1] Este estudo; [2] Beleza et al., 2006; [3] Gonçalves et al., 2005; [4] Adams et al., 2008.

Para o haplogrupo mais comum R1b1b2-M269 foi estimada uma frequência de 55,4%, semelhante aos valores previamente descritos na população Portuguesa e na Península Ibérica, que variam entre 50% e 60,7% (Tabela 1). Este haplogrupo é o mais comum da Europa Ocidental, atingindo frequências que se situam entre os 70-90% nos Bascos (Alonso *et al.*, 2005) e cerca de 65% na Península Ibérica (Adams *et al.*, 2008) e a sua distribuição apresenta um gradiente decrescente no sentido Oeste para Este, que é fundamentalmente justificada pela sua origem Paleolítica e subsequente expansão pelo continente Europeu a partir do refúgio glacial Ibérico após o LGM no Paleolítico Superior. A hipótese da expansão do R1b1b2-M269 a partir dos refúgios glaciares no Paleolítico foi recentemente refutada por Balaesque *et al.* (2010) com base no estudo da variabilidade acumulada no HgR1b1b1 recorrendo aos Y-STRs. Este estudo mostrou uma maior diversidade haplotípica na população do Leste (com os valores mais elevados na Turquia) o que, em conjunto com estimativas temporais (TMRCA – *time to the most recent common ancestor*), sugere que a expansão do R1b1b2-M269 para o Oeste europeu terá ocorrido de forma rápida durante o Neolítico, associada às difusões dêmicas da agricultura.

No que diz respeito haplogrupo J-M304, foi estimada uma frequência de 16,3%, no distrito de Coimbra, inferior aos 25% previamente registados por Beleza *et al.* (2006) mas semelhante às encontradas noutras regiões do país (Tabela 2). A frequência deste haplogrupo em Portugal (média = 13,3%; Tabela 2) é superior à média encontrada na Península Ibérica (8,9%, Tabela 2) ou noutras populações Oeste-Europeias com

exceção da Itália e da Grécia (Semino et al., 2004). O haplogrupo J-M304, tem a sua origem no Médio Oriente sendo nesta região que os seus dois principais sub-grupos J1-M172 e J2-M267 se encontram representados com frequências mais elevadas e com mais alto grau de variação interna ao nível de marcadores STRs (Semino *et al.*, 2004). As suas frequências apresentam um gradiente decrescente dos diferentes países do Médio Oriente para a Europa o que tem sido associada à difusão démica dos agricultores do Neolítico (Underhill *et al.*, 2001).

O sub-haplogrupo J2-M172, registado com uma frequência de cerca de 5% no distrito de Coimbra, encontra-se também em frequências elevadas nas populações Europeias de Judeus Sefarditas e Asquenazim (23-29%) (Semino *et al.*, 2004), e também nos Judeus Portugueses de Trás-os-Montes, onde atinge uma frequência de 24,5% (Nogueiro et al., 2010). Deste modo, a presença dos haplogrupos derivados do J em Portugal, nomeadamente o hgJ2-M172, poderá também ser associada à presença Judaica no país, documentada arqueologicamente desde o 3º século.

O haplogrupo J1-M267, foi registado com uma frequência de 10,7% no distrito de Coimbra, um valor bem superior ao encontrado noutras regiões do país (Tabela 2). Este haplogrupo foi associado à expansão das tribos árabes no Levante e Norte de África onde atinge a frequência (média) de 17% (Adams et al., 2008; Semino et al., 2004) tendo-se verificado que o haplótipo STR mais comum nas populações do Noroeste de África é igual ao haplótipo modal de Árabes Muçulmanos naturais da Galileia (Norte de Israel) (Nebel *et al.*, 2002). Todavia, a ausência deste haplótipo nas amostras do distrito de Coimbra identificadas como J1-M267 (ou J(xJ1) não corrobora a possibilidade destes haplogrupos terem sido introduzidos na região pela permanência Árabe/Berbera entre os séculos VIII e XIII.

Para além dos anteriores, outros eventos históricos justificativos da presença dos haplogrupos derivados do J-M304, têm sido considerados, nomeadamente as dispersões Grega e Fenícia no Mediterrâneo (Zalloua *et al.*, 2008).

O haplogrupo E1b1b1-M35 registou uma frequência de 15,3% na população do distrito, semelhante aos resultados obtidos noutras regiões do país (Tabela 2). Este haplogrupo encontra-se amplamente distribuído pelo continente Africano, em especial no Norte e Leste de África e Médio Oriente (Semino et al., 2004). A sua origem é atribuída à região Leste Africana onde atinge a sua maior diversidade haplotípica. No continente europeu, em termos gerais, observa-se uma tendência decrescente de

frequência no sentido Este-Oeste, associada à difusão da agricultura a partir do Médio Oriente (Semino et al., 2004). Verifica-se a existência de uma profunda descontinuidade e um fluxo genético limitado do Nordeste Africano, onde atinge frequências próximas dos 75%, para a Península Ibérica, onde apresenta uma frequência média de 10% (Adams et al., 2008).

O sub-haplogrupo E1b1b1a-M78, um dos subgrupos derivados do E1b1b1-M35 com maior dispersão, ter-se-á expandido pela Europa a partir dos Balcãs por eventos migratórios múltiplos no período Neolítico ou pós-Neolítico (Semino et al., 2004). A sua presença na costa Sul da Europa foi também sugerida como um possível indicador das expansões Gregas (Semino et al., 2004). A frequência deste haplogrupo verificada no Distrito de Coimbra (3,3%) é semelhante ao conjunto do território de Portugal (Tabela 2).

O sub-haplogrupo E1b1b1c-M123 apresenta uma menor dispersão pela Europa Ocidental, que também se reflecte na amostra populacional por nós estudada, com o registo de apenas 2 indivíduos (2,2%). Cinnioğlu et al. (2004) sugere um percurso migratório para o hgE1b1b1c-M123 idêntico ao hgE1b1b1a-M78, colocando a Turquia como ponto inicial da expansão provável.

A presença do hgE1b1b1b-M81 na Península Ibérica, tem sido relacionada com as linhagens Berberes do Norte de África, uma vez que este haplogrupo é encontrado em frequências elevadas nestas populações (frequência média do hgE1b1b1b-M81 no Norte de África =54%; Adams et al., 2008) e por se saber que estes integraram maioritariamente os exércitos muçulmanos que invadiram a Península Ibérica no início do século VIII. Este subgrupo, de todos os que derivam do E1b1b1-M35, é o que atinge uma maior frequência (5,4%) na população analisada no presente trabalho, e tem sido observado com frequências variáveis (1-12%) nas populações Ibéricas (frequência média 4%) (Adams et al., 2008), ocorrendo com frequências muito baixas <6% no Oeste Europeu (Semino et al., 2004; Cruciani et al., 2004). É pois bem provável que a presença deste haplogrupo signifique um contributo genético proveniente do Norte de África durante a ocupação islâmica do distrito de Coimbra.

Para os haplogrupos G-M201 e I-M170 foram encontradas frequências de 3,3% e 4,4%, respectivamente. O haplogrupo G, apresenta frequências elevadas (>30%) na região do Cáucaso, Geórgia e Ossétia (Semino et al., 2000) possivelmente indicando uma ampla distribuição durante o Paleolítico Superior que agora está restrita a áreas marginais. Quanto ao haplogrupo I, que atinge frequências de 18% na Europa, é o único

haplogrupo confinado ao continente Europeu para o qual é consensual uma origem e diferenciação anterior ao LGM, no Paleolítico Superior (Semino e tal., 2000).

Foram também analisados sete microssatélites (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392 e DYS393) nos haplogrupos mais comuns da amostra populacional do distrito de Coimbra (R1b1b2-M269, J-M304 e E1b1b1-M35) com o objectivo de averiguar o grau de diversidade haplotípica acumulada e identificar haplótipos associados a grupos populacionais específicos.

No haplogrupo J1-M267 foi encontrado um haplótipo descrito por Zalloua *et al.* (2008) como um sinal de colonização Fenícia (PCS1+). Zalloua *et al.* (2008) encontrou este haplótipo fortemente associado ao haplogrupo J2 mas também aos haplogrupos J1 e E1b1b1. Também no haplogrupo E1b1b1a-M78 foi encontrado um haplótipo que corresponde sinal de colonização Grego (GCS1+), tal como descrito por Zalloua *et al.* (2008). O facto de os 2 indivíduos que apresentam quer o haplótipo PCS1+ quer o haplótipo GCS1+ serem naturais da região litoral do distrito (Concelho de Mira) poderá significar a presença dos Fenícios e Gregos nesta zona litoral associada à difusão marítima destas antigas civilizações, não sendo contudo de excluir a possibilidade de estes haplótipos se poderem ter originado de outras formas.

Em conclusão, o estudo efectuado com marcadores do cromossoma Y numa amostra populacional do distrito de Coimbra permitiu identificar os haplogrupos principais que têm vindo a ser descritos nas populações Europeias em geral e nas populações da Península Ibérica em particular. O haplogrupo R1b1b2, associado aos eventos migratórios ocorridos na pré-História Europeia, apresenta-se como o mais frequente na população estudada, em conformidade com o que se encontra descrito para as populações da Europa Ocidental. Foram estimadas frequências semelhantes às de outros estudos no país para os haplogrupos E1b1b1-M35 e J-M304, ambos associados à expansão dêmica da agricultura no Neolítico, mas também a eventos históricos mais recentes. É de destacar a presença no distrito dos sub-haplogrupos E1b1b1a-M78, previamente associado à difusão Grega, e E1b1b1b-M8, associado à presença muçulmana na Península entre os séculos VIII e XIII. Os sub-haplogrupos J1-M267 e J*(xJ1) (que inclui o sub-haplogrupo J2), foram identificados em proporções díspares na amostragem populacional do distrito de Coimbra. A identificação de um sinal de colonização Grega no hgE1b1b1a-M78 e de um sinal de colonização Fenícia no hgJ1-

M267 poderá vir a valorizar os eventos migratórios destes povos na região litoral do distrito.

Bibliografia

Adams SM, Bosch E, Balaesque PL et al. (2008) The genetic legacy of religious diversity and intolerance: paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *Am J Hum Genet*; **83**(6):725-36.

Alonso S, Flores C, Cabrera V et al. (2005) The place of the Basques in the European Y-chromosome diversity landscape. *Eur J Hum Genet*; **13** :1293-1302.

Balaesque P, Bowden GR, Adams SM et al. (2010) A predominantly neolithic origin for European paternal lineages. *PLoS Biol*; **8**:1-9.

Bandelt H-J, Forster P, Röhl A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*; **16**:37-48

Beleza S, Gusmão L, Lopes A et al. (2006) Micro-phylogeographic and demographic history of Portuguese male lineages. *Ann Hum Genet*; **70**:181-94.

Brion M, Quintans B, Zarrabeitia M et al. (2004) Micro-geographical differentiation in Northern Iberia revealed by Y-chromosomal DNA analysis. *Gene*; **329**, 17–25.

Cinniöglu C, King R, Kivisild T et al. (2004) Excavating Y-chromosome haplotype strata in Anatolia. *Hum Genet*; **114**(2):127-48.

Cruciani F, Santolamazza P, Shen P et al. (2002) A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am J Hum Genet*; **70**(5):1197-214

Excoffier LG, Schneider S. (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*; **1**:47-50.

Gonçalves R, Freitas A, Branco M et al. (2005) Y-chromosome lineages from Portugal, Madeira and Açores record elements of Sephardim and Berber ancestry. *Ann Hum Genet*; **69**:443-454.

Nebel A, Landau-Tasserion E, Filon D (2002) Genetic evidence for the expansion of Arabian tribes into the Southern Levant and North Africa. *Am J Hum Genet*; **70**, 1594–1596.

Neto D, Montiel R, Bettencourt C et al. (2007) The African contribution to the present-day population of the Azores Islands (Portugal): analysis of the Y chromosome haplogroup E. *Am J Hum Bio*; **19**(6):854-60.

Nogueiro I, Manco L, Gomes V, Amorim A, Gusmão L (2010) Phylogeographic analysis of paternal lineages in NE Portuguese Jewish communities. *Am J Phys Anthropol*; 141:373-81.

Pereira L & Ribeiro F (2009) *O Património Genético Português*. Gradiva Publicações, 202 pp.

Saraiva JH (1992) *História concisa de Portugal*. Publicações Europa-América 15ª ed. 375 pp.

Semino O, Magri C, Benuzzi G et al. (2004) Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. *Am J Hum Genet*; **74**: 1023–1034.

Semino O, Passarino G, Oefner PJ et al. (2000) The genetic legacy of paleolithic Homo sapiens in extant Europeans: a Y chromosome perspective. *Science* **290**:1155–1159.

Underhill PA, Lin AA, Mehdi SQ et al. (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res*; **7**:996.

Zalloua PA, Platt DE, El Sibai M et al. (2008) Identifying genetic traces of historical expansions: phoenician footprints in the Mediterranean. *Am J Hum Genet* **83**:633–642.