

Prolusione

del Socio nazionale ALBERTO PIAZZA

Professore ordinario di Genetica umana nell'Università di Torino
Presidente della Human Genetics Foundation (HuGeF)



I profili genetici degli Italiani

Illustri ospiti, caro Presidente, cari Consoci, Signore e Signori,

È con vero piacere che parlo in questa circostanza in una sala così prestigiosa. L'invito onora con me un tema che mi è stato sempre molto caro per due ragioni: innanzitutto l'orgoglio di appartenere ad una nazione che, malgrado tutti i suoi limiti, le sue contraddizioni laceranti, le sue ambiguità, è ancora ricca di opere, donne e uomini straordinari; quindi la volontà di condividere con voi un lavoro iniziato molti anni fa, che aveva ed ha tuttora l'ambizione di coinvolgere discipline diverse e perciò di contribuire alla missione principale di una Accademia illustre come questa: varcare i confini dei propri campi di studio e ritrovarsi insieme in geografie più ampie, che possano suscitare stimoli ed interessi reciproci inattesi, moltiplicando così le occasioni di generare idee nuove.

1. Nozioni introduttive

Se il termine 'genetica', seguendo la definizione usata originalmente nel 1905 dal biologo inglese William Bateson, indica «la scienza che studia l'eredità e la variazione cercando di scoprire le leggi che governano le somiglianze e le differenze negli individui che sono in rapporto di discendenza», la *genetica delle popolazioni umane* è tradizionalmente quella parte della genetica che studia i meccanismi evolutivi i quali regolano nel tempo e nello spazio i cambiamenti biologici delle popolazioni.

Oggi le tecniche moderne di analisi del DNA rendono possibile lo studio dell'evoluzione delle popolazioni umane a livello del singolo individuo, ma i dati fino a ieri utilizzati consistono in caratteri ereditari di cui sia stato possibile determinare le regole di trasmissione da una generazione all'altra. Di questi medesimi caratteri ereditari vengono calcolate le frequenze percentuali in campioni di individui di dimensioni variabili, e perciò variamente rappresentative di popolazioni più estese, ma comunemente collocate nella stessa area geografica da cui sono tratti.

La definizione dell'oggetto di studio, la *popolazione*, è pertanto più operativa che rigorosa, e spesso, il riferimento linguistico è più risolutivo: si parla, ad esempio, di popolazione basca, e non di popolazione francese o spagnola, anche se i Baschi vivono sia in Francia sia in Spagna.

L'informazione genetica è presente nel nucleo di ciascuna cellula di un organismo. È contenuta in strutture chiamate *cromosomi* di cui nella nostra specie di *Homo sapiens* esistono 23 coppie per cellula. Ogni coppia è costituita da un cromosoma trasmesso dal padre e da un cromosoma trasmesso dalla madre, morfologicamente indistinguibili. I cromosomi sono costituiti da acido desossiribonucleico (DNA), e i *geni* si possono operativamente definire come segmenti di DNA deputati alla sintesi di *prodotti genici*. Oggi possiamo studiare direttamente il DNA; ma fino a poco tempo fa dovevamo limitare il nostro studio al *fenotipo*, cioè ad un prodotto genico che si poteva manifestare nell'aspetto esterno dell'organismo (come il colore azzurro degli occhi o la presenza del gruppo sanguigno Rh positivo, ecc.).

Gli individui della nostra specie hanno circa 25.000 geni, ma le unità chimiche elementari, i nucleotidi, che compongono il DNA di quei 25.000 geni si allineano in una sequenza di 3,5 miliardi di elementi: se tali elementi costituissero le lettere del nostro alfabeto avremmo una ricchezza lessicale di ben 7024 Divine Commedie in ognuna delle nostre cellule.

2. *Come le frequenze dei geni cambiano nel tempo e si distribuiscono nello spazio*

Si consideri ad esempio il gene che determina il gruppo sanguigno Rh positivo o negativo in Europa, rappresentato nella figura 1.

Ogni individuo può essere positivo o negativo a seconda del risultato di un test. Se, per esempio, si considera un campione sufficientemente numeroso degli abitanti della Sardegna, si trova che la frequenza degli individui Rh negativi è il 4%; ma lo stesso tipo di indagine compiuto sugli abitanti del Piemonte (per rimanere in Italia) rivelerebbe una frequenza del 15%; mentre nella popolazione basca, che vive nella regione dei Pirenei tra Francia e Spagna, si troverebbe la frequenza più alta in Europa: circa 30 Baschi su 100 hanno il gruppo sanguigno Rh negativo. Il valore assoluto della differenza di frequenza dello stesso gene in due popolazioni misura una *distanza genetica* fra di esse, ma è più significativo prendere in esame geni diversi e calcolarne una distanza genetica media.

Oggi si può misurare la variabilità genetica tra gli individui, ma fino a quando la tecnologia non ci ha permesso di prendere in esame tutta la sequenza del DNA di un individuo e per tutte le possibili coppie di individui poterne valutare le differenze, sono state le distanze genetiche medie fra tutte le coppie di popolazioni, a costituire il «documento» di base sul quale hanno lavorato gli studiosi di evoluzione biologica della nostra specie.

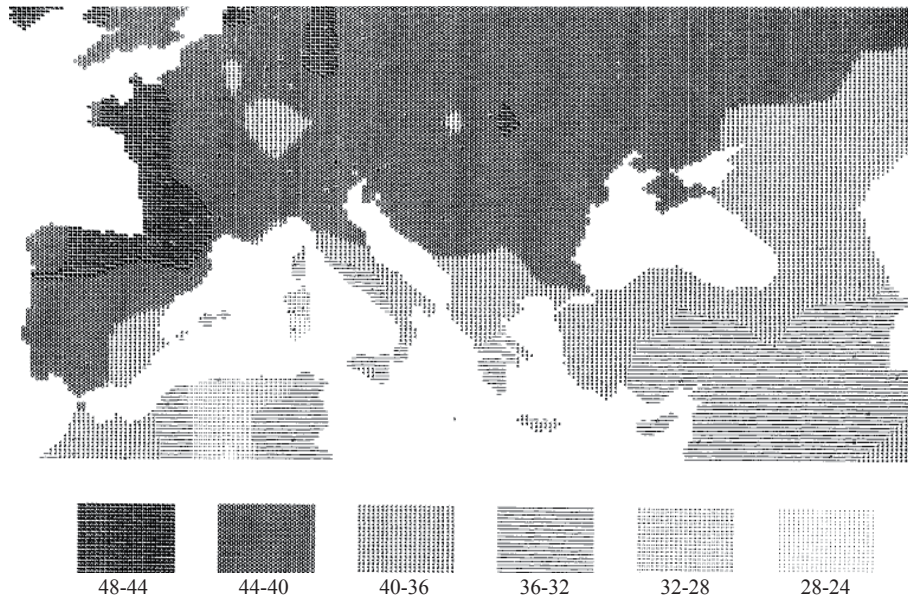


Fig. 1. Distribuzione del gene Rh negativo in Europa. Tratteggi differenti indicano le frequenze del gene come descritto.

I geni e loro frequenze variano nello spazio e nel tempo a causa di quattro meccanismi che non mi è possibile spiegare in dettaglio: 1) la *deriva genetica*, che dipende dalle dimensioni delle popolazioni; 2) la *mutazione*, che costituisce il meccanismo fondamentale del cambiamento genetico generando nuovi alleli e, nel caso in cui venga duplicato per accidente un intero gene, anche nuovi geni. Senza mutazioni non vi sarebbe evoluzione. Tuttavia, si tratta di un evento molto raro e la sua possibilità di diffondersi in una popolazione è minima; 3) la *migrazione* ha un effetto analogo a quello della mutazione, in quanto l'introduzione di nuovi individui immigrati può introdurre geni nuovi, o comunque altera la distribuzione di frequenze degli alleli nella popolazione che li riceve. L'effetto genetico della migrazione dipende allora in modo critico dal rapporto fra i numeri di individui delle due popolazioni: si tratta di un dato demografico che, purtroppo, dalle fonti archeologiche e storiche si può ricavare raramente e indirettamente, e per soli ordini di grandezza; 4) la *selezione naturale* è il meccanismo per cui i fenotipi aumentano o diminuiscono la loro frequenza a seconda della maggiore o minore capacità di adattamento

all'ambiente naturale degli individui che li portano. Fu Darwin, con la sua intuizione rivoluzionaria, a porre a fondamento dell'evoluzione delle specie il meccanismo dell'adattamento biologico per selezione naturale. L'ambiente seleziona i tipi più «adatti», cioè quelli che manifestano una aumentata fertilità o una probabilità maggiore di sopravvivenza all'età fertile dell'individuo che le porta.

3. *La distribuzione geografica dei polimorfismi genetici: perché può interessare discipline diverse dalla genetica?*

Rispetto ai classici documenti archeologici o storici, il documento «genetico» ha un vantaggio e uno svantaggio. Lo svantaggio è quello di riferirsi alle popolazioni oggi esistenti, e perciò di rappresentare il risultato sincronico di un processo evolutivo la cui sequenza diacronica va ricostruita a ritroso, senza possibilità di verifica sperimentale delle varie fasi del processo. Il vantaggio consiste in una proprietà importante del cambiamento biologico: la sua estrema lentezza. Il processo fondamentale di cambiamento biologico di tutti gli organismi viventi (la mutazione) consiste in un cambiamento casuale del messaggio chimico costituito dal DNA che forma un gene. La mutazione avviene in un individuo ed è trasmissibile ai suoi discendenti, ma siccome ogni individuo ha un numero limitato di discendenti immediati, devono trascorrere molte generazioni prima che si diffonda in un numero di individui tale che si possa facilmente osservare all'interno di una popolazione. I cambiamenti genetici possono quindi fornirci informazioni di eventi più antichi, di cui non esiste memoria storica.

4. *La storia genetica dell'Italia pre-DNA*

Un polimorfismo genetico è un carattere che si trasmette da una generazione a quella successiva secondo le regole di Mendel. Il primo polimorfismo genetico è stato il sistema ABO, scoperto nel 1900 da Landsteiner senza conoscere le sequenze specifiche di DNA che sintetizzavano i tre alleli del sistema A, B e O. Nel successivo cinquantennio, sono stati descritti numerosi altri marcatori genetici controllati o prodotti da geni, ma non sono essi stessi dei geni, perché non sono tratti di sequenze di DNA: per distinguere questa classe di marcatori genetici dai polimorfismi più recenti costituiti direttamente da segmenti o intere sequenze di DNA («polimorfismi del DNA»), li si è denominati polimorfismi «classici», mentre sarebbe più appropriato definirli poli-

morfismi pre-DNA. Ci si attende che i polimorfismi pre-DNA e i polimorfismi del DNA raccontino fondamentalmente la stessa storia, perché i primi sono prodotti dai secondi; ma il potere risolutivo e discriminante della sequenza di DNA, come vedremo, fornisce informazioni molto più dettagliate, a livello addirittura dell'individuo e non più delle comunità di individui.

L'elenco dei polimorfismi pre-DNA, sui quali si è basata la prima descrizione della struttura genetica dell'Italia, si riferisce a risultati riportati in letteratura dagli anni Sessanta agli anni Novanta, operando una selezione di quei campioni di individui di cui si sia accertata la stabilità della residenza per almeno due generazioni. Abbiamo considerato più di 50 polimorfismi genetici pre-DNA: non ci è bastato l'esame di un solo polimorfismo, quale per esempio l'Rh, perché la sua frequenza, esattamente come per un carattere antropometrico (per esempio il colore della pelle) può essere stata influenzata da fattori ambientali quali il

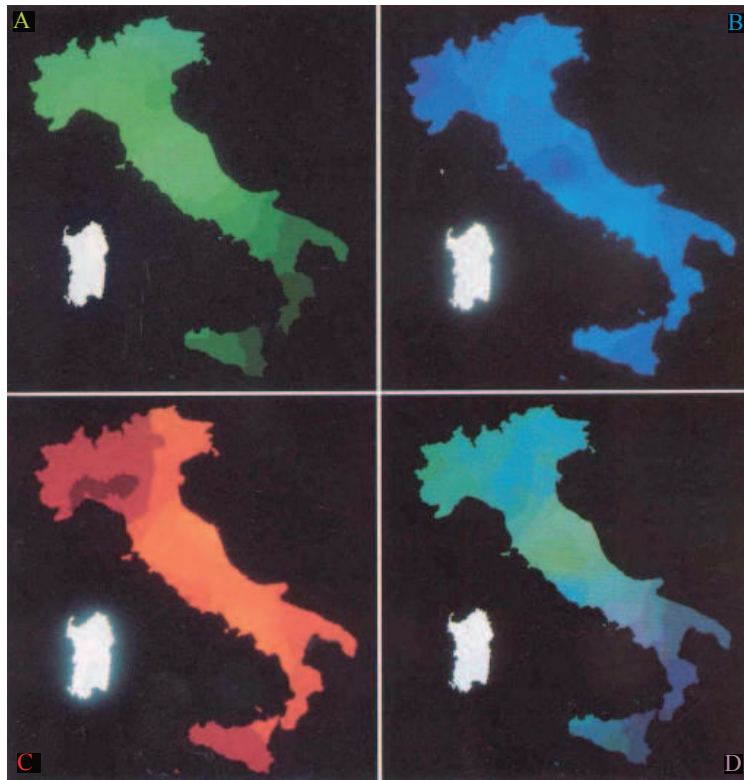


Fig. 2 a, b, c, d.

clima, il diffondersi di una malattia, ecc. In realtà, a noi interessa la storia di un altro genere di variabilità genetica: quella che risulta dalla successione di insediamenti di popolazioni diverse, dalla loro mescolanza, dalle loro vicende demografiche. Tale storia è comune a tutti i geni che portiamo, e si dimostrerà tanto più degna di fede quanto maggiore è il numero di geni osservati.

Quella cui dobbiamo riferirci è dunque una storia di più geni, per evitare il rischio di descrivere una variabilità dovuta all'ambiente invece che alla successione di eventi storici. Senza entrare in dettagli troppo tecnici, abbiamo usato l'artificio di rappresentare l'effetto simultaneo della variabilità di molti geni (nel caso dell'Italia più di 50) con l'intensità di un colore, per esempio il verde, così che a verdi più o meno intensi corrispondessero aree geografiche con valori più o meno elevati non di uno ma di molti geni. In questa rappresentazione illustrata nella figura 2a (verde) le relazioni genetiche tra le varie regioni d'Italia risultano in un'immagine geografica monocromatica, una vera e propria «mappa» genetica dell'Italia, in cui aree geografiche con uguale intensità di colore sono geneticamente simili: la «distanza» genetica viene misurata dalla differenza dell'intensità del colore.

In questa operazione di sintesi non viene restituita tutta l'informazione dei dati originali, ma solo quella più rilevante. Il procedimento consiste nel sottrarre ai valori dei dati originali quelli spiegati dalla mappa genetica in verde e riprodurre i valori risultanti in un'altra mappa: il risultato è ancora una mappa dell'Italia, ma per distinguerla dalla precedente si è adottato l'accorgimento di cambiarne il colore, per esempio si è utilizzato il rosso anziché il verde. Questo associare valori di contenuto informativo decrescente a mappe geografiche di colore diverso, è un procedimento che si può applicare in modo iterativo per recuperare dai dati originali la maggiore informazione possibile. Sfruttando poi la capacità dell'occhio umano di sintetizzare tre colori (il verde, il rosso e il blu), si sono associati a questi colori i valori delle tre mappe più informative delle frequenze di tutti i marcatori genetici e, con l'ausilio di elaboratori di immagini ad alta definizione, si sovrappongono automaticamente le tre mappe corrispondenti, così da restituire in una mappa tricromica la differenziazione genetica complessiva dell'area geografica analizzata.

Questa tecnica – originariamente applicata allo studio delle popolazioni europee in collaborazione con Paolo Menozzi e Luigi L. Cavalli-Sforza – ci ha permesso di ricostruire rappresentazioni «sintetiche» della variazione di tutti i geni presi in esame. La distribuzione dei geni esaminati in Italia² è

² A. Piazza, N. Cappello, E. Olivetti, S. Rendine, *A genetic history of Italy*, «Annals of Human Genetics», 1988, 52, pp. 203-213.

sintetizzata nelle tre mappe monocromatiche illustrate nella figura 2b (blu), figura 2c (rossa): esse non contengono tuttavia l'informazione delle frequenze geniche originali, perchè si sono esaminate le tre componenti più informative (tre colori), ovvero quella informazione, valutabile nel 60-70%, che è condivisa da tutti i geni, cioè proprio quella che dovrebbe ridurre l'effetto di fattori ambientali. La composizione tricromatica delle tre mappe è riportata nella figura 2d (tricromica).

Un'osservazione comune alle tre mappe è la posizione peculiare della Sardegna, che in tutte e quattro le figure risulta non analizzata. L'esclusione non è dettata dalla mancanza di dati, né da scarso interesse per questa bellissima isola; anzi, è vero il contrario, nel senso che la struttura genetica della Sardegna è così profondamente diversa da quella delle altre regioni italiane, che il suo esame richiede una trattazione separata. Tuttavia se dovessimo inserire i dati della Sardegna genetica nelle tre mappe monocromatiche dell'Italia, questa regione risulterebbe di colore intensissimo, e l'Italia continentale di colore pallidissimo e uniforme, tanto diverse risultano le due strutture genetiche.

L'influenza della Magna Graecia

L'esame della figura 2a indica molto chiaramente un gradiente delle frequenze geniche italiane lungo l'asse Nord-Sud. Molti eventi possono aver contribuito a questa differenziazione: ultime in ordine di tempo, ma non di scarso rilievo, le massicce migrazioni interne avvenute nel secondo dopoguerra. Si noti tuttavia che l'unità di tempo del cambiamento della frequenza di un gene è il tempo medio di una generazione (circa 20-30 anni) e che tali cambiamenti – a meno che una popolazione sostituisca completamente un'altra con un'invasione seguita da genocidio (è il caso delle popolazioni indigene delle Americhe) – si svolgono nell'arco di tempi lunghi, spesso accompagnati da processi di crescita demografica il cui effetto diventa apprezzabile solo dopo molte generazioni. L'interpretazione che abbiamo dato alla mappa della figura è che essa riproduca l'influenza genetica della *Magna Graecia* sulla storia del nostro Paese: nella figura 3 ne è riportata la geografia (in centro) insieme alla distribuzione dei geni (a sinistra) e dei cognomi di origine greca (a destra). La validità di questa interpretazione è ovviamente legata a una stima delle dimensioni numeriche delle colonie greche in Italia: la diffusione dei geni dei colonizzatori dipende dal loro rapporto con il numero dei colonizzati. I demografi storici stimano che alla fine dell'Età del bronzo (1.000 a.C.) la densità media della popolazione greca fosse circa tre volte quella del resto dell'Europa, raggiungendo forse un milione di abitanti. Tra il 1.000 ed il 400 la Grecia dello stesso periodo triplica



Fig. 3. La geografia dell'influenza genetica della *Magna Graecia* sulla storia del nostro Paese (in centro), insieme alla distribuzione dei geni (a sinistra) e dei cognomi di origine greca (a destra. Tratta da G. Zei, A. Lisa, O. Fiorani, *I Cognomi degli Italiani*, 2008).

i suoi abitanti, che raggiungono i 3 milioni. L'Italia, che dopo la Grecia costituiva l'area geografica più densamente popolata, contava nel 400 a.C. circa 4 milioni di persone: si può calcolare che, prima della dominazione di Roma, uno ogni 10-15 abitanti della penisola italiana fosse di origine greca (Beloch³ con la revisione più recente di Brunt⁴). Il loro ordine di grandezza non contraddice l'idea di una introduzione di geni greci dal Sud dell'Italia, la cui diluizione tra i geni delle popolazioni autoctone si può ancora oggi osservare sotto forma di un graduale cambiamento genetico da Sud verso Nord.

Etruschi

La carta della figura 2b (in blu) discrimina un'area geografica che potrebbe corrispondere all'antica Etruria e che per tale ragione è stata chiamata la «componente etrusca» dei nostri geni. Non è il caso di sottolineare qui l'importanza della civiltà etrusca nella nostra storia, e valutare quanto remoto nel tempo sia stato l'insediamento di quella popolazione. L'associazione geografica tra culture villanoviane (la presenza di necropoli che testimoniano il costume funerario dell'incinerazione invece dell'inumazione dopo l'età del Bronzo, e l'uso di vasi biconici come ossari) e insediamenti etruschi ha costituito un argomento a favore della natura indigena della popolazione etrusca, come aveva originalmente suggerito Dionigi di Alicarnasso nel primo secolo avanti Cristo. Le iscrizioni più antiche delle popolazioni etrusche, incise con una variante dell'alfabeto greco, risalgono circa al 700 a.C. e – a differenza del latino e della maggior parte delle altre lingue italiche – non sembrano appartenere alla famiglia di lingue indo-europee la cui introduzione in Europa è sicuramente anteriore al 3000 a.C.. Per avere un'idea delle proporzioni numeriche, si stima che nel 474 d.C., dopo la sconfitta da parte dei Greci nella battaglia di Cuma, gli Etruschi ammontassero a circa 300.000-400.000 individui, cioè a circa il 12-16% dei 2 milioni e mezzo di persone che non abitavano nelle colonie greche.

La Sardegna

Se si vuole ricostruire la filogenesi delle popolazioni europee con un albero che dalla radice si dirama in rami successivi, subito dopo la popolazione dei

³ J. Beloch, *Die Bevölkerung der griechisch-römischen Welt*, Duncker & Humblot, Leipzig, 1886.

⁴ P.A. Brunt, *Italian Manpower 225 BC-AD 14*, Oxford University Press, Oxford, 1971.

Saami (Lapponi) viene separata la popolazione dei Sardi, seguita da quella dei Greci. I primi abitanti della Sardegna furono preneolitici: nella zona centro-settentrionale i più antichi insediamenti umani risalgono a 9120 ± 380 anni fa (si tratta della più antica datazione per un 'isola mediterranea).

Intorno all'800 a.C., i Fenici colonizzarono le coste dell'isola, soprattutto quelle meridionali; successivamente la colonizzazione continuò per opera di Cartagine, una colonia fenicia in Tunisia. Nell'ambito della conquista del Mediterraneo, i Romani occuparono anche la Sardegna, ad esclusione delle zone più interne. Alla caduta dell'Impero romano seguirono le invasioni di Vandali, Bizantini e Saraceni; la dominazione araba continuò fino al X secolo. L'isola fu poi assoggettata al controllo di Pisa, seguito da quello congiunto di Catalogna e Aragona, e infine del regno sabaudo. Tutti gli invasori degli ultimi due millenni hanno però lasciato tracce genetiche limitate.

La causa di gran lunga più importante della considerevole differenza genetica tra i Sardi e le altre popolazioni europee o africane è costituita dalla loro storia e dal loro successivo isolamento.

Dai dati relativi ai geni pre-DNA si deduce che soltanto i nordafricani non hanno contribuito in modo rilevante al *pool* genico sardo (anche se una componente berbera minore è ancora presente), mentre l'Italia e la Grecia sono state probabilmente i luoghi di origine dei primi occupanti del Neolitico. I Neolitici provenivano a loro volta dal Medio Oriente e dall'odierna Turchia, ma probabilmente durante il passaggio attraverso la Grecia e l'Italia meridionale il loro profilo genetico fu diluito dal flusso genico delle popolazioni mesolitiche locali. I moderni Libanesi sono i più diretti discendenti dei Fenici che fornirono un contributo al patrimonio genetico dei Sardi.

La figura 4 riproduce la suddivisione tradizionale della Sardegna che distingue due aree di insediamento: proto-arcaica e arcaica.

La Sardegna è stata oggetto di studi sulla differenziazione linguistica e sulla distribuzione dei cognomi. Nell'isola si parlano numerosi dialetti, tutti romanzi, ma con termini indicanti un substrato preindoeuropeo di origine non identificata⁵. Allo stato attuale delle conoscenze, risulta che la prima lingua indoeuropea parlata in Sardegna fu il latino. Se questa affermazione risultasse confermata, si potrebbe pensare che i primi immigrati neolitici in Sardegna non fossero numericamente sufficienti per imporre la loro lingua; oppure che parlassero una lingua non appartenente alla famiglia indoeuropea.

I dati genetici della Sardegna, per quanto abbondanti, non sono rappresentativi dell'intera isola, come accade invece con i dati linguistici e con i

⁵ M.L. Wagner, *Historische Lautlehre Des Sardischen*, Max Niemeyer, Halle (Saale), 1941.



Fig. 4. Aree di insediamento della Sardegna antica ed aree linguistiche. Le aree linguistiche sono le seguenti: 1) Sassarese; 2) Gallurese; 3) Logudoro nordovest; 4) Logudoro est; 5) Logudoro sud-Planargia; 6) Goceano (Alta Valle del Tirso); 7) Circondario di Bitti; 8) Nuorese; 9) Baronìa di Orosei-Siniscola; 10) Monte Ferru; 11) Fonni-Barbagia di Ollolai; 12) Campidano di Oristano nord; 13) Media Valle del Tirso; 14) Barbagia di Belvi; 15) Ogliastra; 16) Campidano di Oristano sud-Arborea-Tregenta; 17) Tregenda-Parteolla; 18) transizione tra Gerrei e Tregenda; 19) Sarrabus; 20) Sulcis-Iglesiente; 21) Campidano di Cagliari; 22) Alghero.

cognomi di cui ho avuto l'opportunità di disporre. I dati linguistici usati sono quelli dell'Atlante Linguistico della Sardegna a cura di Benvenuto Terracini e Temistocle Franceschi in base ai rilievi di Ugo Pellis, una preziosa collezione da molti anni custodita nella nostra Università e da noi elaborati^{6,7} e raggruppati mappando i lessemi di 200 «significati» relativamente stabili in 21 aree linguistiche dell'isola; i 4400 cognomi sardi sono stati ricavati da fonti diverse quali i 50.000 matrimoni tra consanguinei registrati dalle parrocchie dal 1750 al 1950^{8,9,10} e i circa 480.000 utenti del telefono.

Nella figura 5 è rappresentata una sintesi cromatica (uguale colore denota omogeneità) delle distribuzioni dei cognomi, dei geni e dei lessemi nelle 22 aree linguistiche.

In tutte le rappresentazioni si può osservare un gradiente Nord-Sud. La correlazione è maggiore fra dati linguistici e cognomi (89%), perché la loro scala temporale potrebbe essere stata essenzialmente la stessa; è minore – sebbene sempre alta – fra dati genetici e dati linguistici (68%) e fra dati genetici e cognomi (66%). I dati genetici si riferiscono a un passato più remoto, e sono meno conosciuti. Si noti nella figura 6 come la molteplicità di confini che indica una eterogeneità genetica significativa si concentri nella area archeologica definita «arcaica» della Sardegna ed è il probabile effetto della deriva genetica, un meccanismo evolutivo particolarmente efficace per popolazioni insediate da lungo tempo nella stessa regione geografica, e originate da nuclei di piccole dimensioni.

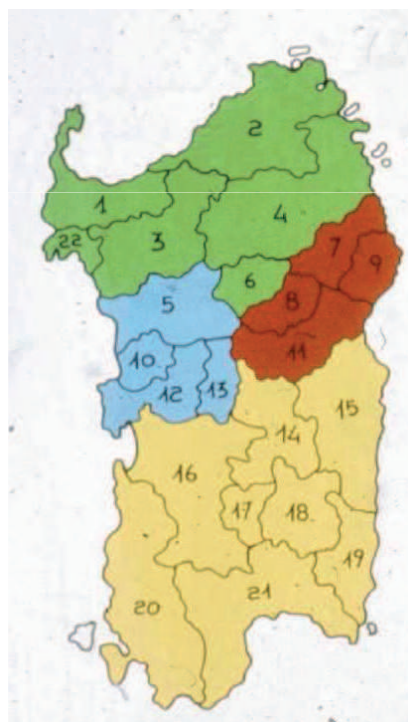
⁶ M. Contini, N. Cappello, R. Griffò, S. Rendine, A. Piazza, *Géolinguistique et géogénétique: une démarche interdisciplinaire*, «Géolinguistique», 1989, 4, pp. 129-197.

⁷ N. Cappello, S. Rendine, R. Griffò, G. Mameli, V. Succa, G. Vona, A. Piazza, *Genetic analysis of Sardinia: I. Data on 12 polymorphisms in 21 linguistic domains*, «Annals of Human Genetics», 1996, 60, pp. 125-141.

⁸ G. Zei, C.R. Guglielmino-Matessi, R. Siri, A. Moroni, L.L. Cavalli-Sforza, *Surnames as neutral alleles: Observations in Sardinia*, «Hum Biol Oceania», 1983, 55, pp. 357-365.

⁹ E.M. Wijsman, G. Zei, A. Moroni, L.L. Cavalli-Sforza, *Surnames in Sardinia. II. Computation of migration matrices from surname distributions in different periods*, «Annals of Human Genetics», 1984, 48, pp. 65-78.

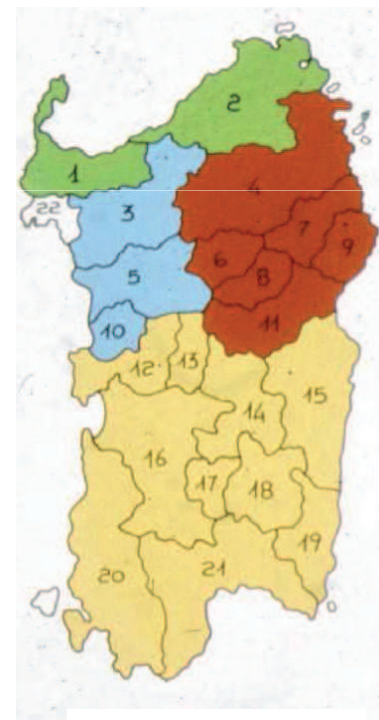
¹⁰ G. Zei, A. Piazza, A. Moroni, L.L. Cavalli-Sforza, *Surnames in Sardinia. III. The spatial distribution of surnames for testing neutrality of genes*, «Annals of Human Genetics», 1986, 50, pp. 169-180.



Cognomi



Geni



Lessemi

Fig. 5. Cognomi, geni e lessemi in 22 domini linguistici della Sardegna.

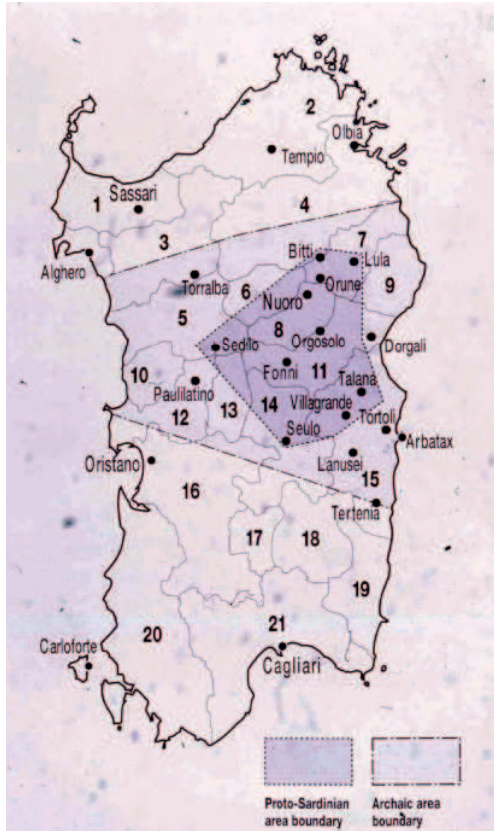


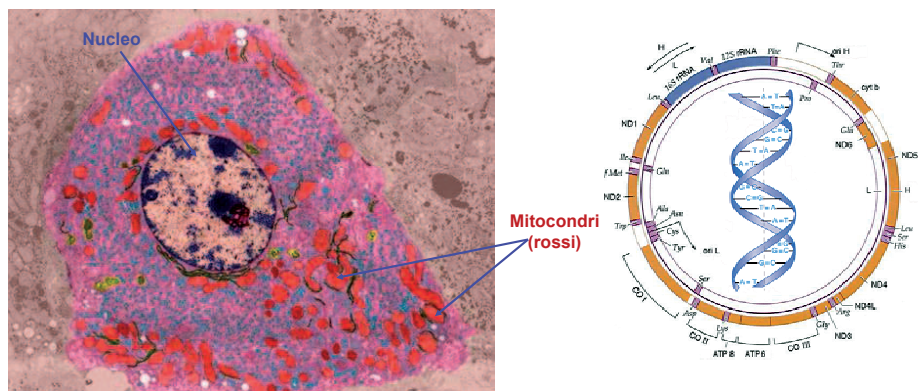
Fig. 6. Immagine genetica tricromica della Sardegna paragonata con le aree archeologiche. I numeri definiscono le aree linguistiche denominate nella didascalia della figura 4.

5. La storia genetica dell'Italia dall'analisi del DNA

5.1. Polimorfismi del DNA mitocondriale

I mitocondri sono organelli intracellulari che producono energia attraverso una serie complessa di catene respiratorie. Contengono un cromosoma circolare unico il cui DNA codifica per 13 proteine. In media, il citoplasma di una cellula umana contiene alcune migliaia di mitocondri (800-2.500), e il DNA di ogni mitocondrio ha una sequenza completamente nota di 16.569 nucleotidi.

Il DNA mitocondriale si è rivelato particolarmente utile per conoscere l'evoluzione umana e i rapporti tra le diverse popolazioni, grazie soprattutto all'alto tasso di mutazione.



Figg. 7-8. I mitocondri sono organi fuori del nucleo che producono energia e hanno un loro DNA (mtDNA) lungo 16.569 basi nucleotidiche conosciute.

5.2. DNA mitocondriale ed oltre: analisi dei campioni toscani

L'analisi del DNA mitocondriale in Italia non ha la completezza geografica delle analisi compiute saggiando i polimorfismi classici pre-DNA, ma la sua risoluzione è assai più fine, e fornisce indicazioni più particolareggiate. Il lavoro più aggiornato¹¹ riporta i dati sulle sequenze del DNA mitocondriale

¹¹A. Achilli, A. Olivieri, M. Pala, E. Metspalu, S. Fornarino, V. Battaglia, M. Accetturo, I. Kutuev, E. Khusnutdinova, E. Pennarun, N. Cerutti, C. Di Gaetano, F. Crobu, D. Palli, G. Matullo, A.S. Santachiara-Benerecetti, O. Semino, R. Villems, H.-J. Bandelt, A. Piazza, A. Torroni, *Mitochondrial DNA Variation of Modern Tuscans Supports the Near Eastern*

saggiate in Toscana (tre campioni: Murlo, Volterra, Valle del Casentino). È stato possibile identificare 39 marcatori genetici mitocondriali, le cui frequenze sono state paragonate con quelle di 10.589 soggetti distribuiti in 47 popolazioni dell'Europa e del Medio Oriente.

Di particolare interesse sono le posizioni dei campioni toscani. Mentre Volterra e la Valle del Casentino si collocano insieme alla maggior parte delle popolazioni del Mediterraneo centrale (inclusi gli altri campioni italiani), il campione di Murlo, particolarmente studiato anche sotto il profilo etno-antropologico, toponomastico e cognominale per i suoi insediamenti etruschi, rivela una chiara e inattesa affinità genetica con le popolazioni del Vicino Oriente. Una frequenza sorprendentemente alta dei marcatori mitocondriali del DNA che caratterizzano il Vicino e Medio Oriente (HV, R0a, U7, U3) colloca Murlo vicino alle popolazioni del Vicino e Medio Oriente. Quale esempio, la distribuzione geografica del marcatore U7 è illustrata nella figura 9. Questa figura dimostra per U7 quel che vale anche per gli altri 3 marcatori: tutti e quattro presentano la loro massima frequenza nel Vicino Oriente, un risultato che conferma in modo molto più risolutivo l'ipotesi formulata analizzando la mappa ottenuta con i marcatori genetici classici (figura 2b in colore blu).

La figura 10 riporta un'analisi successiva più interessante¹² perché permette di attribuire la data di origine del marcatore U7 a circa 2300 anni fa che, tenuto conto dell'errore statistico di 400 anni, non è incompatibile con una sua presenza in Toscana nel periodo della civiltà etrusca.

Ma certamente più interessante ancora è l'analisi recente di un gruppo di ricercatori spagnoli¹³ che non si limita ad esaminare il DNA dei mitocondri ma, sfruttando le tecnologie ben più raffinate disponibili oggi analizza ben 540.000 marcatori distribuiti su tutti i cromosomi dei nuclei delle nostre cellule su un campione di Toscani e campioni di controllo relativi a popolazioni europee, orientali ed africane (figura 11). Di questa analisi vorrei soprattutto richiamare la vostra attenzione sulla figura 12 per le nuove informazioni che ci offre: sono la conseguenza di una tecnologia estremamente raffinata e ricca di messaggi intellettualmente stimolanti. È rappresentato un lungo rettangolo diversamente colorato mediante 4 colori e suddiviso in 5 sotto-rettangoli cui

Origin of Etruscans, «The American Journal of Human Genetics», 2007, 80, pp. 759-768.

¹² F. Brisighelli, C. Capelli, V. Álvarez-Iglesias, V. Onofri, G. Paoli, S. Tofanelli, Á. Carracedo, V.L. Pascali, A. Salas, *The Etruscan timeline: a recent Anatolian connection*, «European Journal of Human Genetics», 2009, pp. 693-696.

¹³ J. Pardo-Seco, A. Gómez-Carballea, J. Amigo, F. Martín-Torres, A. Salas, *A Genome-Wide Study of Modern-Day Tuscans: Revisiting Herodotus's Theory on the Origin of the Etruscans*, in «PLoS ONE», 2014, 9 (9): e105920. doi:10.1371/journal.pone.0105920.

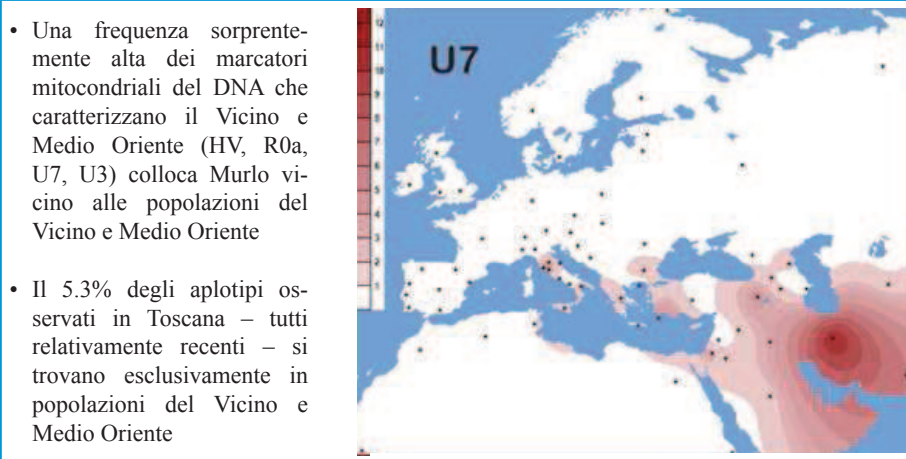


Fig. 9.

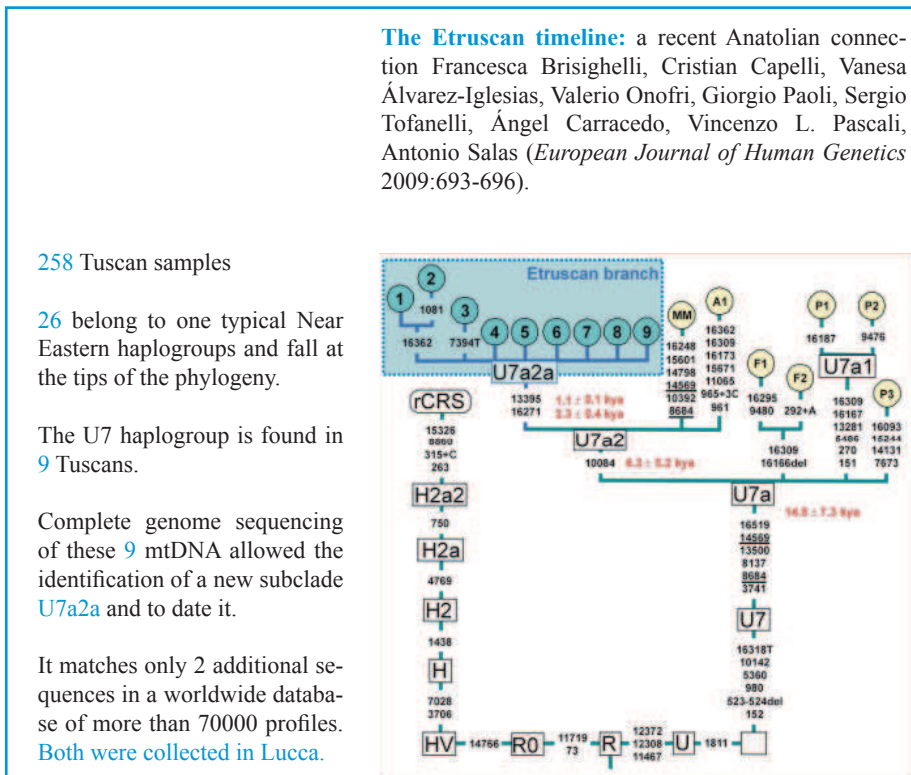



Fig. 10.

Avanzamento della tecnologia: Genome-Wide Analysis (GWA)



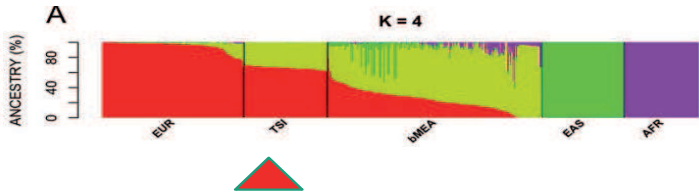
98 Toscani attuali caratterizzati da 540.000 marcatori genetici

J. Pardo-Seco, A. Gómez-Carballa, J. Amigo, F. Martínón-Torres, A. Salas (2014) *A Genome-Wide Study of Modern-Day Tuscans: Revisiting Herodotus's Theory on the Origin of the Etruscans*. PLoS ONE 9 (9): e105920. doi:10.1371/journal.pone.0105920.

A sin.: musicista etrusco, Tomba del Triclinio, Tarquinia.

Fig. 11.

Componente europea e medio-orientale di individui appartenenti a quattro diversi raggruppamenti continentali



L'età in cui individui provenienti dal Medio Oriente si sarebbero mescolati con i Toscani locali è stata stimata essere tra 3100 e 2600 anni fa.

J. Pardo-Seco, A. Gómez-Carballa, J. Amigo, F. Martínón-Torres, A. Salas (2014), *A Genome-Wide Study of Modern-Day Tuscans: Revisiting Herodotus's Theory on the Origin of the Etruscans*. PLoS ONE 9(9): e105920. doi:10.1371/journal.pone.0105920.




Fig. 12.

sono associate delle sigle: da destra AFR che sta per Africani, EAS che sta per Orientali, bMEA che sta per Medio-Orientali, TSI per Toscani ed EUR per Europei non Italiani. L'altezza del rettangolo/i misura la percentuale di ascendenza rispetto agli antenati classificati secondo i 4 colori. Per essere più comprensibile esaminiamo l'ultimo rettangolo a destra tutto viola. Possiamo considerare la base del rettangolo come l'unione di tante aste viola così vicine tra di loro che non si riescono a distinguere: ebbene queste aste viola corrispondono ad individui (nel rettangolo ne sono rappresentati un centinaio circa) e l'uniformità del colore si interpreta come uniformità del loro profilo di DNA. Dal momento che tutti quegli individui risiedono in Africa, possiamo dire che sono tutti operativamente Africani senza ascendenti non africani. Lo stesso si può affermare degli individui del lontano Oriente (colore verde) ma per gli altri gruppi di individui che risiedono in Europa (colore predominante rosso), in Medio-Oriente (colore predominante giallo), in Toscana la situazione è ben diversa: non vi è individuo il cui profilo di DNA non sia il risultato di una stratificazione delle altre componenti. Il caso dei Toscani è ovviamente quello che suscita il nostro maggiore interesse perché evidenzia in tutti gli individui esaminati un'ascendenza che è per il 70% europeo e per il 30% medio-orientali: si noti bene, questo non significa che se esaminiamo un campione di 100 individui toscani 70 di loro sono Europei e 30 sono Medio-orientali ma che tra gli ascendenti di ciascuno di loro troviamo in media 70 Europei e 30 Medio-orientali: Si tratta di un messaggio molto forte che dovremmo saper ascoltare. Se poi aggiungiamo la capacità di stimare la data in cui le due componenti si sono mescolate (tra 3,100 e 2,600 anni fa) è assai difficile non prendere almeno in considerazione la celebre ipotesi etnogenetica formulata da Erodoto. Mi permetterò di riprendere questa spinosissima questione nelle conclusioni.

5.3. Polimorfismi del DNA del cromosoma Y

Negli esseri umani, come in altri mammiferi, il sesso maschile è determinato dal cromosoma Y che, attraverso l'azione di un singolo gene, SRY (Sex determining Region Y), causa la differenziazione dei testicoli e determina la mascolinità.

La distribuzione geografica dei polimorfismi DNA del cromosoma Y è influenzata anche dal comportamento sociale degli uomini. Il 70% delle società moderne pratica la patrilocità: in caso di matrimonio tra persone che non vivono nello stesso luogo, generalmente è la donna che si sposta per seguire il marito, e non viceversa. Per conseguenza, la permanenza nel luogo di nascita

caratterizza più gli uomini che le donne, e la differenziazione locale del DNA del cromosoma Y è maggiore di quella del DNA mitocondriale, trasmesso solo per via materna.

5.3.1. *Il DNA del cromosoma Y in Italia: analisi dei campioni siciliani*

La struttura genetica della Sicilia è già stata indirettamente menzionata a proposito della componente più importante della struttura genetica dell'Italia analizzata mediante i polimorfismi pre-DNA o «classici». La si è considerata nel contesto più ampio della *Magna Graecia*, riscontrando una continuità genetica tra l'isola e le regioni meridionali d'Italia che la Grecia ha a suo tempo colonizzato.

In realtà, la documentazione archeologica e linguistica sul popolamento della Sicilia fornisce informazioni anche sui tempi immediatamente precedenti: è accertato che l'isola fu popolata nel Paleolitico Superiore e nel Neolitico¹⁴. La presenza dei Sicani (oggi associata alla *cultura di Thapsos* nell'Età del Bronzo media, 1300 a.C., e alla *cultura Pantalica dal 1250 all'850 a.C.*) è stata documentata. Nell'ultimo periodo della cultura Pantalica (prima Età del Ferro) i Sicani sono stati spinti verso l'interno e il Sud dell'isola dai Siculi (che provenivano dall'Italia continentale) da Est, e, in modo meno rilevante, dagli Elimi da Ovest, dove avevano fondato Erice e Segesta (IX-VIII secolo a.C.). Infine, a cominciare dall'VIII secolo a.C., le coste della Sicilia e dell'Italia meridionale sono state oggetto di rilevanti insediamenti da parte di colonizzatori greci, di cui rimangono numerose testimonianze storiche ed archeologiche. La colonizzazione fenicia e cartaginese ebbe luogo nello stesso periodo di tempo, ma con minor impatto, dal momento che non riuscì a sopravvivere alla potenza dei Greci eccetto che nel triangolo estremo della Sicilia occidentale, dove si sostituì agli Elimi già confinati in quell'isolato.

Una associazione puntuale tra strati genetici eterogenei e popolazioni storiche differenti della Sicilia (per esempio *Sicani, Elymi e Siculi*) è stata resa possibile dall'analisi del DNA del cromosoma Y¹⁵ il cui potere risolutivo è in grado di indicare datazioni possibili sulle differenziazioni genetiche più recenti avvenute nell'isola.

¹⁴ S. Tusa (ed.), *Prima Sicilia. Alle origini della società siciliana*, Edipoint, Palermo, 1997.

¹⁵ C. Di Gaetano, N. Cerutti, F. Crobu, C. Robino, S. Inturri, S. Gino, S. Guarrera, P.A. Underhill, V. Romano, F. Cali, M. Gasparini, G. Matullo, A. Salerno, C. Torre, A. Piazza, *Differential Phoenician and Greek migrations to Sicily is supported by genetic evidence of the Y Chromosome*, «European Journal of Human Genetics», 2009, 17, pp. 91-99.

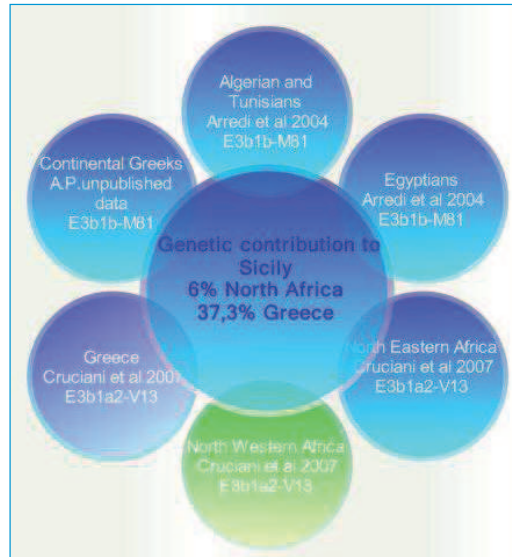


Fig. 13.

Sono stati analizzati 236 campioni distribuiti in nove aree geografiche diverse della Sicilia per una trentina di polimorfismi del DNA del cromosoma Y. Nella figura 13 sono riassunti i risultati più significativi: il contributo dei cromosomi nord-africani al *pool* genico siciliano è stato stimato in circa il 6%, mentre quello greco raggiunge la ragguardevole proporzione del 37,3%. Purtroppo, il numero dei siciliani che si distribuiscono tra le due linee filogenetiche non è sufficiente per discriminare la loro età: per tutte e due si colloca tra 3000 e 4000 anni fa, un intervallo di tempo che è in ottimo accordo con le conoscenze storiche.

Stabilire una corrispondenza biunivoca tra l'eterogeneità genetica della Sicilia osservata oggi e una presunta composizione genetica dei popoli che la occuparono prima dei Romani, può essere un esercizio pericoloso fino a quando non sarà possibile tipizzare e datare con un errore minore di quello ottenibile oggi, il DNA dei fossili pre-romani disponibili nelle aree della Sicilia e interessanti sotto il profilo archeologico. Tuttavia, l'attenzione per ulteriori ricerche si potrebbe concentrare su una ipotesi di lavoro minima:

1. il popolamento della Sicilia ha causato una differenziazione genetica lungo l'asse Ovest-Est dell'isola che ancor oggi si può identificare: i

polimorfismi genetici pre-DNA¹, i cognomi¹⁶, le isoglosse dei dialetti che si parlano all'interno dell'isola¹⁷ concordano nel mostrare questa differenziazione;

2. i polimorfismi del DNA del cromosoma Y, fornendo un quadro di riferimento temporale a tale differenziazione, identificano negli insediamenti di popolazioni pre-romane la sua probabile origine.

6. *I profili genomici degli Italiani riflettono la storia dell'Europa e del Mediterraneo*

Grazie all'impegno e all'entusiasmo di ricercatori, tutti torinesi, che lavorano allo HuGeF, la Fondazione di Ricerca che ha quali Soci la Compagnia di San Paolo, l'Università ed il Politecnico di Torino, è stato possibile, tra i numerosi progetti di ricerca nel campo della genomica umana, fare progressi significativi anche nella ricostruzione della mappa dei genomi italiani, al momento identificati in un campione pilota di 300 individui che abitano in Italia da molte generazioni. Il Progetto, al di là del suo valore conoscitivo, ha anche una ricaduta applicativa perché oggi, in una fase in cui la medicina tende sempre più ad essere una medicina di precisione indirizzata alla persona e perciò anche al suo profilo genomico, la conoscenza di profili genomici di riferimento in una geografia così eterogenea come quella italiana diventa sempre più necessaria per poter paragonare genomi di pazienti con genomi di individui sani che abitano lo stesso territorio.

Nella figura 14 sono citati i riferimenti bibliografici di un lavoro preliminare appena terminato e di prossima pubblicazione. Oltre al titolo, ai nomi di chi vi ha lavorato ed alla distribuzione regionale dei campioni esaminati, intendo sottolineare il fatto che si tratta del primo lavoro sull'Italia in cui i campioni sono stati esaminati per 2,5 milioni di marcatori del DNA: un'impresa che solo la tecnologia e la competenza sviluppate nello HuGeF sono state in grado di realizzare.

La figura 15 è una rappresentazione sintetica bidimensionale dell'eterogeneità genomica dell'Italia: ogni punto corrisponde ad un individuo, ogni colore corrisponde a una regione e la distanza tra i punti misura la distanza genomica tra gli individui. Senza approfondire, risultano evidenti sia la già

¹⁶ C.R. Guglielmino, G. Zei, L.L. Cavalli-Sforza, *Genetic and cultural transmission in Sicily as revealed by names and surnames*, «Human biology», 1991. 63(5), pp. 607-627.

¹⁷ G. Ruffino, *Sicily*, in *The Dialects of Italy*, a cura di M. Maiden e M. Parry, Routledge, London e New York, 1997, pp. 365-375.

Fig. 14. G. Fiorito, C. Di Gaetano, S. Guarrera, F. Rosa, M.F. Feldman, G. Matullo, *The Italian Genome reflects History of Europe and Mediterranean Basin* (HuGeF, Human Genetics Foundation, Torino; Dipartimento di Scienze Mediche, Università di Torino; Department of Biology, Stanford University, Stanford), «European Journal of Human Genetics», *in press*.

300 campioni dall'Italia;
1272 campioni dall'Europa,
Medio Oriente, Nord
Africa;
2,5 milioni di marcatori del
DNA.



discussa diversità della Sardegna sia il fatto che le posizioni reciproche delle varie regioni riflettono la loro geografia, anche quando non si consideri la Sardegna, la cui peculiarità rispetto all'Italia appiattisce la diversità delle altre regioni.

La successiva figura 17 pone le regioni italiane nel contesto di 35 Paesi Europei, Medio-orientali e Nord-Africani. Anche in questa rappresentazione più estesa, la geografia fisica e la geografia genomica si sovrappongono, in altre parole la diversità genomica ha nella distanza geografica il contributo di maggior peso. L'osservazione non è certo sconvolgente anche se ottenuta con il vantaggio di un numero di campioni assai ridotto rispetto a quelli necessari con i marcatori pre-DNA: ci dice semplicemente che i partner tanto più si preferiscono quanto meno distante è il loro luogo di residenza e meno diversa è la loro lingua. È da notare la curiosa somiglianza tra gli antenati degli abitanti sardi e gli antenati della *Mummia del Similaun* (anche nota come *Uomo venuto dal ghiaccio* e, informalmente, *Ötzi*) un reperto fossile ritrovato nel 1991 sulle Alpi Venoste al confine fra l'Italia (Alto Adige) e l'Austria (Tirolo) risalente a un'epoca compresa tra il 3300 e il 3100 a.C. (età del rame), un risultato che convalida un'osservazione recente analizzando la sequenza di DNA del fossile.

I profili genomici individuali e la loro comparazione in termini di differenze di DNA ci forniscono informazioni più interessanti che vorrei condividere

Fig. 15. Geografia genomica dell'Italia.

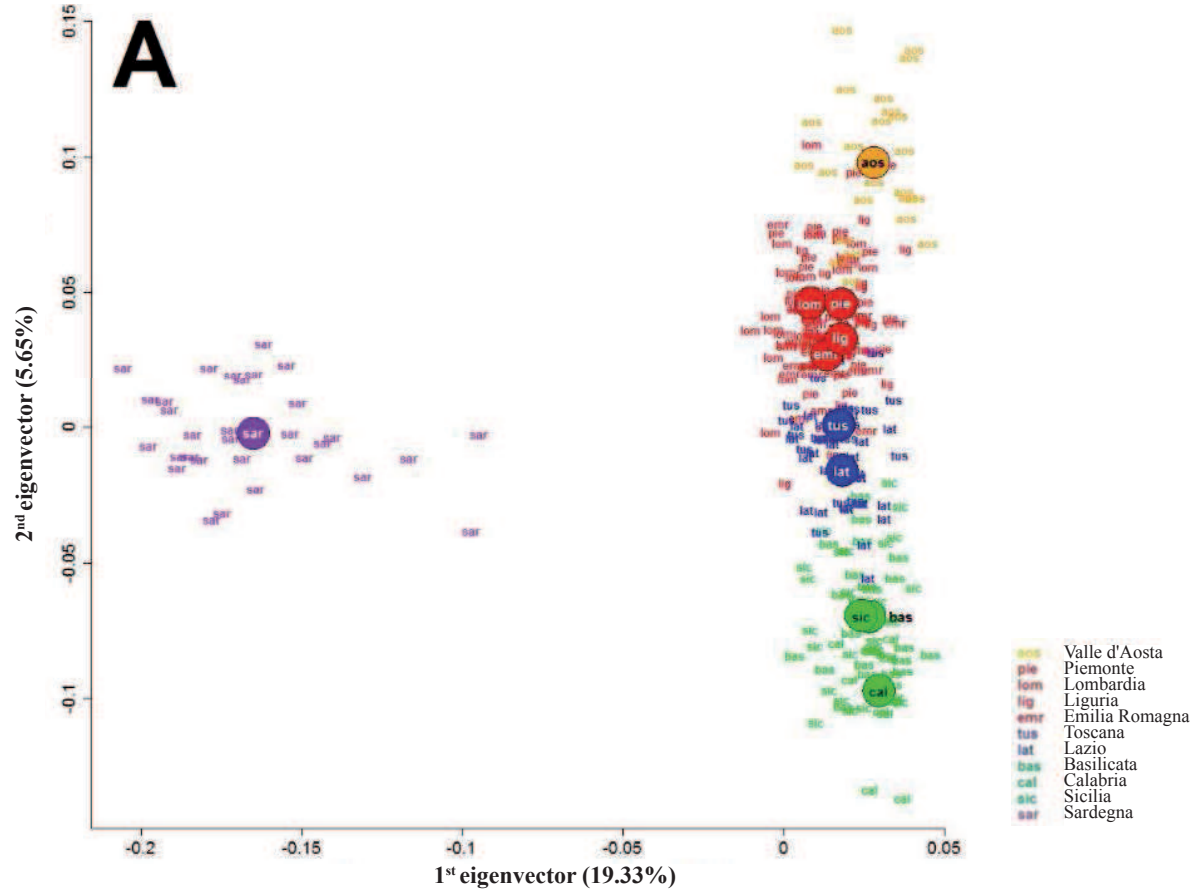
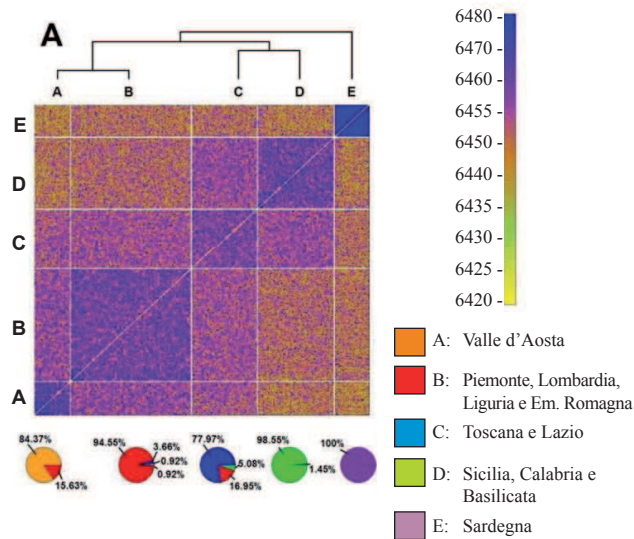


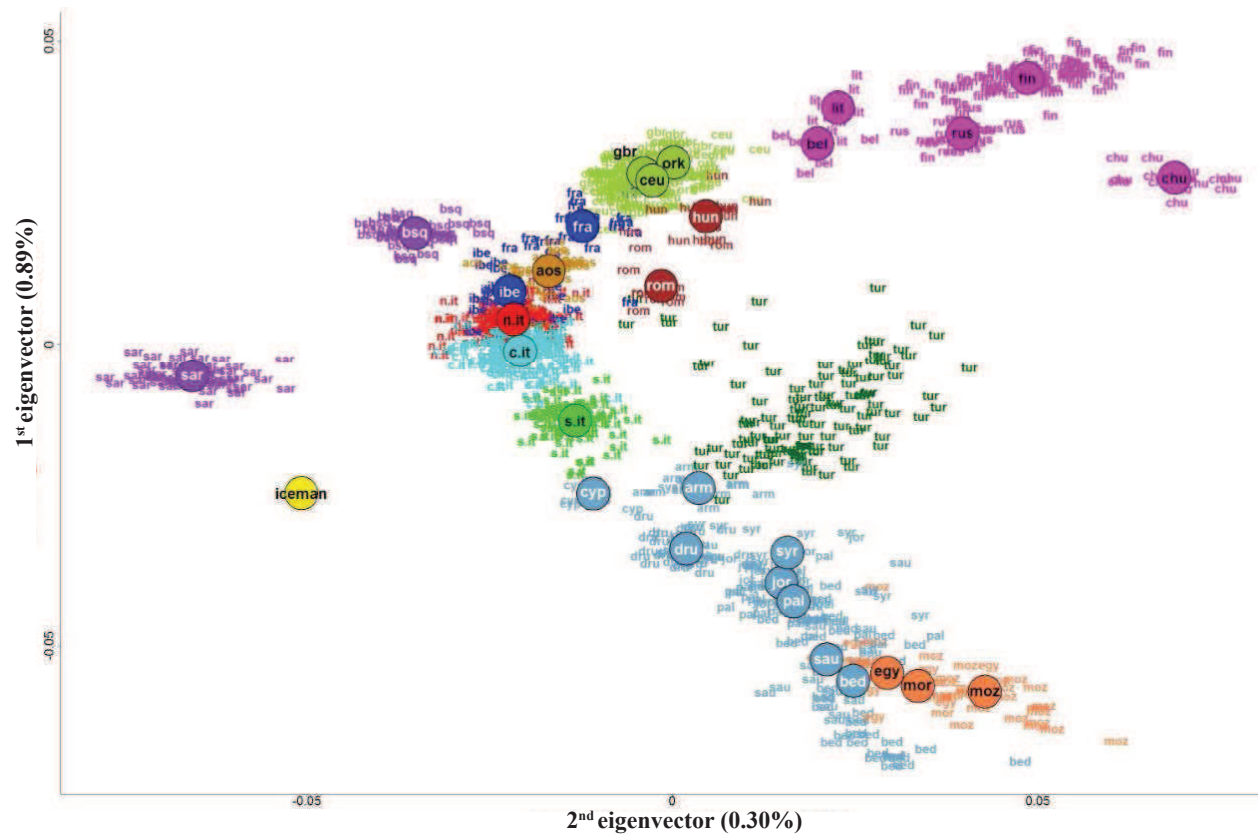
Fig. 16. La *Matrice di co-ascendenza* indica il numero di segmenti genomici ereditati dalla stessa popolazione ancestrale per ciascuna coppia di individui.



La successiva figura 17 pone le regioni italiane nel contesto di 35 Paesi Europei, Medio-orientali e Nord-Africani. Anche in questa rappresentazione più estesa, la geografia fisica e la geografia genomica si sovrappongono, in altre parole la diversità genomica ha nella distanza geografica il contributo di maggior peso. L'osservazione non è certo sconvolgente anche se ottenuta con il vantaggio di un numero di campioni assai ridotto rispetto a quelli necessari con i marcatori pre-DNA: ci dice semplicemente che i partner tanto più si preferiscono quanto meno distante è il loro luogo di residenza e meno diversa è la loro lingua. È da notare la curiosa somiglianza tra gli antenati degli abitanti sardi e gli antenati della *Mummia del Similaun* (anche nota come *Uomo venuto dal ghiaccio* e, informalmente, *Ötzi*) un reperto fossile ritrovato nel 1991 sulle Alpi Venoste al confine fra l'Italia (Alto Adige) e l'Austria (Tirolo) risalente a un'epoca compresa tra il 3300 e il 3100 a.C. (età del rame), un risultato che convalida un'osservazione recente analizzando la sequenza di DNA del fossile.

I profili genomici individuali e la loro comparazione in termini di differenze di DNA ci forniscono informazioni più interessanti che vorrei condividere commentando la figura 16 che, in modo simile alla figura 12 – il rettangolo colorato che includeva i Toscani –, usa ancora il colore per assegnare ad ogni coppia di individui la loro probabilità di un antenato comune. Il tipo di Tabella prende il nome di «matrice di co-ascendenza» (in inglese *coancestry matrix*)

Fig. 17. Geografia genomica dell'Europa.



ed è più semplice mostrare come si legge con un esempio. Le regioni italiane sono raggruppate in 3 macro-regioni (Nord, Centro, Sud), la Sardegna e la Valle d'Aosta simboleggiate nella figura da B, C, D, E ed A. A destra vi è una barra colorata dal giallo (in basso) al blu intenso (in alto) che indica in senso crescente la probabilità di un antenato comune (il blu corrisponde alla massima). Consideriamo l'ultimo rettangolo in alto a destra che riguarda individui sardi (prima riga E) di cui si cerca l'antenato comune con altri sardi (ultima colonna E): il rettangolo all'incrocio della prima riga e dell'ultima colonna è tutto blu intenso il che significa che gli antenati comuni degli individui che abitano in Sardegna sono sardi. Situazione analoga si verifica per i campioni della Valle d'Aosta (ultima riga e prima colonna), ma per tutti gli altri casi i rettangoli sono composti da puntini colorati in modo diverso tra di loro ed anche con probabilità di avere un antenato comune più bassa (colore più chiaro e tanto più chiaro quanto le regioni sono distanti, per esempio Sardegna con Valle d'Aosta, riga A con colonna E). Se ogni individuo ha una bassa probabilità di avere un antenato comune con un altro individuo residente addirittura nella stessa area geografica, questo significa che in quell'area vi è stata una immigrazione da altre aree geografiche ed una mescolanza così consistente da poter essere «ricordata» dal DNA individuale per molte generazioni.

Dunque il profilo genetico scritto nel DNA individuale testimonia non tanto l'origine della persona quanto la serie di mescolanze genomiche che si sono verificate nella sua storia passata. È naturale quindi chiedersi se abbiamo la possibilità di stimare le date di tali mescolanze. La risposta è positiva e nella nostra analisi dei profili genetici degli Italiani si sono documentate le datazioni di tre mescolanze:

1. *Circa 3000-2100 anni fa, tra individui che provenivano dal Caucaso e dal Medio Oriente, e individui che vivevano in Toscana e in Lazio.* Ripropone la questione degli Etruschi la cui presenza in Italia risale almeno all'VIII secolo a.C. per poi diluirsi e rapidamente annullarsi nel primo secolo a.C.
2. *Circa 2050-1650 anni fa, tra individui di ascendenza Nord-Europea e individui che vivevano nell'Italia Centro-Settentrionale.* L'intervallo temporale corrisponde al «Periodo delle migrazioni barbariche» con la dissoluzione finale dell'Impero romano d'Occidente (476 d.C) ed è molto verosimile che popolazioni di origine nordica abbiano contribuito alla fondazione di alcune comunità dell'Italia settentrionale: il caso dei Lombardi e degli insediamenti nella Lombardia odierna potrebbe suggerire l'eventualità di altre migrazioni nord-europee anche a causa del rapido crollo numerico delle popolazioni locali, ma modalità e meccanis-

mi sono controversi, privi come siamo di un preciso quadro demografico delle aree coinvolte.

3. *Circa 1000 anni fa, tra individui di ascendenza Nord-Africana/Medio-Orientale con individui che vivevano nell'Italia Meridionale ed in Sicilia.* L'intervallo temporale corrisponde alla conquista normanna dell'Italia meridionale, ma mi asterrò da ogni considerazione perché il campionamento delle regioni meridionali è insufficiente, mancando completamente di individui in Campania e Puglia e disponendone di pochi in Calabria. La storia dei Normanni nell'undicesimo e dodicesimo secolo, spesso mercenari al servizio degli uni e degli altri, aggiunge complessità a complessità ed il rischio di un delirio interpretativo è tanto maggiore quanto minori sono le conoscenze.

7. L'eredità genetica dell'Italia: in quale misura siamo figli del nostro passato?

Chi mi ascolta si chiederà a questo punto quali siano le conclusioni della nostra analisi. Speriamo di essere stati abbastanza convincenti nel dimostrare come l'Italia sia, dal punto di vista genetico, un mosaico di gruppi etnici ben differenziati.

Per riassumere in poche righe il risultato di ricerche durate molti anni e condensate nelle pagine precedenti, il quadro della variabilità genetica (oltre che linguistica ed onomastica) dell'Italia isola quattro scenari importanti:

1. la separazione della Sardegna dal resto del continente, anzi da tutte le altre popolazioni europee, che probabilmente rivela un'origine più antica della sua popolazione, indipendente da quella delle popolazioni italiche e con ascendenze nel Mediterraneo Medio-Orientale;
2. un gradiente di variabilità da Nord a Sud, di probabile origine pre-romana, consolidato dalle colonizzazioni greche nell'Italia meridionale;
3. la presenza di aree indipendenti di omogeneità genetica nella regione toscana e intorno alla Liguria;
4. l'eterogeneità genetica all'interno della Sicilia.

Tale quadro non solo non è stato smentito, ma è stato rafforzato e precisato dalle analisi che si sono susseguite direttamente sul DNA. Queste ultime, meno sistematiche ma con risoluzione molto maggiore delle precedenti (il campionamento non copre tutta la penisola in modo omogeneo e il numero dei campioni è talvolta inadeguato), sono tuttavia più focalizzate su problemi locali specifici e perciò acquisiscono in precisione di dettaglio a livello addi-

rittura individuale – soprattutto nell’indicare quadri di riferimento temporale – quel che possono perdere in generalità.

La risposta alla domanda in quali tempi abbiano avuto origine le differenze genetiche di cui ancor oggi osserviamo le tracce in Italia, è più difficile: la mancanza di una documentazione archeologica continua nel tempo e contigua nello spazio ci permette di datare in modo approssimato i movimenti delle popolazioni preistoriche, ma non di accertare le loro dimensioni numeriche in Italia. In generale, è nostra convinzione tuttavia che la romanizzazione dell’Italia e dell’Europa non abbia cambiato in modo sostanziale la struttura genetica delle popolazioni conquistate. La colonizzazione romana ha prodotto cambiamenti importantissimi nel sistema politico, amministrativo, commerciale, edilizio, più che sostituzioni di intere popolazioni (così come è avvenuto, ad esempio, per la colonizzazione europea delle Americhe). Analogamente, l’imposizione di un’unica lingua, il latino, non è stata omogenea, né ha prodotto cambiamenti paralleli, se ancora oggi ciascuna regione d’Italia parla un dialetto diverso, in cui troviamo molti «relitti» delle lingue usate dagli abitanti più antichi. Un profilo genetico dell’Italia, fortemente conservatore di una situazione etnica assai antica, sembra infatti trovar conforto in alcune considerazioni che la dialettologia aveva da tempo avanzato. Si è detto, per esempio, che la maggior fedeltà e somiglianza del toscano al latino rispetto ad altri dialetti, è dovuta al substrato etrusco, troppo differente dal latino per consentire ibridazioni (come è invece avvenuto nei dialetti a base osca dell’Italia meridionale). Anche le affinità linguistiche dei dialetti nord-occidentali, dall’Emilia-Romagna al Piemonte (escluso il Veneto), sarebbero da attribuirsi al comune substrato celto-ligure di quelle regioni, cioè a insediamenti che precedono la nostra era.

Il mondo italico, con la diversità delle sue culture locali che si sono più o meno chiaramente definite all’inizio dell’Età del ferro (IX-VIII secolo a.C.), sembra essersi congelato nell’immagine offerta oggi dalla sua struttura genetica a livello macro- e micro-geografico. Si tratta forse ancora di una ipotesi di lavoro, ma la «testimonianza» dell’analisi di interi genomi a livello di DNA, che oggi si avvale di strumenti di dissezione molecolare estremamente raffinati, non la smentisce.

Tuttavia il messaggio che tale analisi così sofisticata oggi ci lancia, va ben oltre: la possibilità di analizzare l’intera sequenza del DNA di un individuo così rapidamente ed economicamente ci pone domande più radicali.

Ragionando sulla questione etrusca, ma più in generale di ogni ricerca di etnogenesi, mi sono convinto che sia giunto il momento di interrogarci sulla generale validità euristica del concetto di «formazione» introdotto da

Massimo Pallottino nel suo famoso volume «Etruscologia» che ha influenzato intere generazioni di archeologi: sostituire la domanda sulla provenienza di un popolo con una formale presa d'atto della sua esistenza in quanto prodotto di un processo, sicuramente è servito ad esorcizzare il rischio razzista, in un momento storico (il libro è del 1947) in cui il mondo usciva dal dramma scatenato dalle mitologie razziali, ma di fatto ha solo spostato i termini del problema.

Come argomenta Mario Torelli in un saggio recente (2014) ed illuminante che conclude un incontro dal titolo significativo «*Origines*»: *percorsi di ricerca sulle identità etniche nell'Italia antica*¹⁸:

L'approccio di Pallottino nasceva da una fortunata trasposizione, dall'Italia contemporanea all'Etruria antica, dell'approccio della cultura nazionalista e della storiografia da questa espressa (penso a Gioacchino Volpe in particolare) alla definizione della nazione italiana. L'affermazione di Massimo d'Azeglio «fatta l'Italia, ora occorre fare gli Italiani» era una presa d'atto che il popolo italiano non esisteva *ab origine*, ma doveva essere costruito, «formato» appunto, secondo ben precisi richiami al passato, ma soprattutto con principi «moralì» rigorosi, valori perenni della borghesia. Parlare di formazione della nazione etrusca è usare una definizione, prima che astratta o generica, tautologica: la sostanza storica che invero quel termine è «come» quella nazione si è formata, che è il reale compito di chi degli Etruschi scrive la storia, senza per questo cancellare la necessità di un'analisi delle componenti remote, materiali e ideali, che hanno condotto alla nascita di quella formazione, *in primis* della genesi delle strutture socio-economiche, un fatto che interferisce profondamente nelle stesse concezioni che gli antichi avevano della propria realtà etnica. E così, la pretesa autoctonia, che tanto ha giocato nella costruzione moderna della questione dell'origine di quel popolo, non è un «fatto», ma un'opinione rivelata a Dionigi (che non per caso è l'unico a parlarcene) da un $\eta\epsilon\gamma\epsilon\mu\omicron\nu$ etrusco, forse proprio un *praetor Etruriae* della lega, che era stata appena resuscitata da Augusto [...]. Per gli antichi dichiararsi autoctoni non è altro che esprimere un diritto inalienabile sulla terra in cui si vive («da questa terra siamo nati e dunque essa ci è sempre appartenuta e sempre ci apparterrà»): l'affermazione fatta dall'aristocratico etrusco a Dionigi ha il sapore di una disperata dichiarazione delle classi proprietarie etrusche massacrate fino alle soglie dell'età imperiale, fatta allo scopo di evitare l'eventualità di futuri espropri, dopo le confische sillane prima e triumvirali poi.

Dunque non solo le razze sono state inventate ma anche la ricostruzione dell'identità etnica è spesso il frutto di una appropriazione culturale.

¹⁸M. Torelli, «Conclusioni», *Mélanges de l'École française de Rome – Antiquité*, 126-2, 2014, URL: <http://mefra.revues.org/2431>.

Permettetemi ora di lasciare gli scenari in cui porre i profili genetici degli italiani e di consultare su Internet un sito (ve ne sono molti) con lo scopo di conoscere il mio genoma di italiano: ovvero una sequenza di DNA la cui variazione da una persona all'altra costituisce la trama della nostra diversità biologica ma anche la possibile fonte di manifestazioni patologiche. In risposta ricevo l'offerta del servizio, il prezzo ed una ulteriore offerta e proposta di servizio: vuoi conoscere persone che hanno un genoma molto simile al tuo? Che significato ha questo «incontro» della genomica con *Internet* se non la nascita e lo sviluppo di un «*social network* genomico» o «genomica sociale in rete» che va oltre le funzioni sociali di un «social network» ma anche oltre le funzioni conoscitive dell'analisi del genoma di una o più persone? Perché non mi rivolgo direttamente a un laboratorio? Da che cosa trae origine?

Tale domanda, che è oggetto di un saggio di Sergio Pistoï¹⁹, va forse preceduta da un'altra, più difficile: in quale misura i concetti di identità biologica e identità culturale sono contigui?

La identità biologica va intesa *strictu sensu*, cioè senza ambigui riferimenti a nozioni lontane dalla biologia: mi riferisco al fatto che tutti i vertebrati sono sottoposti a processi di identità biologica ben definiti, i quali risiedono sostanzialmente in quel sistema essenziale per la nostra difesa biologica dal diverso da sé, che è il sistema immunitario.

Occorre aggiungere che gli odierni processi tecnologici non solo ci permettono di analizzare rapidamente il nostro profilo genomico, ma anche di scoprire le relazioni biologiche tra noi e tutti gli altri risalendo a banche dati con potenziali antenati comuni. Dunque, a monte della ricerca su Internet della propria identità genomica vi è un'ambiguità: quella di contrabbandare per una ricerca di identità biologica quella che in realtà è una ricerca di identità culturale. Alla quale si aggiunge un'altra ambiguità, spesso un inganno vero e proprio, che spero di aver illustrato dipanando la complessità dei profili genetici degli italiani: quella di voler correlare l'identità biologica con l'identità di origine.

Ci siamo mai chiesti perché sia così frequente il desiderio di andare indietro nel tempo alla ricerca delle nostre origini? Il «*social network* genomico», che sottintende un rapporto nuovo e molto più pervasivo fra individuo, scienza e tecnologia, rispetto a quello che oggi caratterizza i *social network* tradizionali quali *Facebook*, è la puntuale manifestazione di questa ricerca. La tecnologia non è più solo il mezzo con cui far pervenire in tempo reale un messaggio creato dagli utenti, ma nella genomica sociale in rete è la tecnologia stessa che diventa contenuto, fornendo la scansione del nostro DNA e delegando quindi

¹⁹ S. Pistoï, *Il Dna incontra Facebook*, Marsilio, Venezia, 2012.

ad altri il compito di incidere, almeno in parte, sulla nostra socialità, per di più coltivando in noi anche la sensazione illusoria di una socialità condivisa dal DNA comune e perciò più intima.

Tutto ciò induce cambiamenti profondi nella nostra cultura. Ricordo (anche perché oggetto di una non lontana rappresentazione scaligera) l'aria famosa di Lohengrin nell'opera omonima di Wagner: «Mai devi domandarmi – Né a palesar tentarmi – Dond'io ne venni a te – Né il nome mio qual è» e le tragiche conseguenze causate dall'aver invece posto la domanda. La condivisione di una più o meno ampia frazione di DNA con un nostro simile e la conoscenza migliore di quanto simili siano i profili genomici di entrambi con conseguente invito a esplicitarli su un sito Internet, illude doppiamente: sia perché non è la somiglianza della sequenza di DNA a render due individui più simili e amici, sia perché non garantisce né un passato (a partire dall'eventuale antenato comune, l'origine) né un destino comune.

I popoli sono costituiti di individui ed anche per le loro origini vale la stessa considerazione. È vero che il tema delle origini appare più che mai attuale in un mondo come quello contemporaneo, in cui l'identità sociale, in grave crisi, obbliga costantemente il mondo intellettuale a interrogarsi sull'origine della propria cultura, che è come dire sul proprio passato, e perciò anche sui motivi e le forme di quei processi che hanno condotto a specifiche identità. Ma è anche vero, parafrasando un passo di George Steiner²⁰, che non è il passato che ci governa se non, forse, nel senso di una biologia che cambia sempre più rapidamente nel tempo. Sono le immagini del passato, spesso fortemente strutturate e selettive, come miti, a governarci. Immagini e costruzioni simboliche del passato sono impresse, quasi alla stregua di informazioni che oggi chiameremmo epigenetiche, sulla nostra sensibilità. Ogni era storica rispecchia se stessa nella raffigurazione, e nella mitologia che attiva il proprio passato. Oggi siamo in un momento a mio parere fortunato in cui tecnologie molto innovative ci permettono di studiare l'uomo piuttosto che la popolazione alla quale appartiene o è appartenuto nel passato. In campo medico si parla appunto di medicina di precisione o «personalizzata» riflettendo una tendenza analoga. I profili genetici degli italiani – e non *il profilo*, al singolare – sono tanti quanti sono gli individui ed ognuno si costruisce il proprio futuro, ma in parte anche il proprio passato. La sfida è quella di analizzarli uno per uno per il benessere di tutti: sono anch'essi personalizzati. Spero di avervi convinto.

²⁰ G. Steiner, *Nel castello di Barbablù. Note per la ridefinizione della cultura* (1971); trad. it., Garzanti, Milano, 2011.