INTERNATIONAL JOURNAL OF SCIENTIFIC AND UNIVERSITY RESEARCH PUBLICATION

IJSURP
International Journal Research
www.ijsurp.com

Research Paper

THE FUTURE COMMON ANCESTRY OF ALL PRESENT-DAY HUMANS

Philip M. Service || Department of Biological Sciences

At some future time, each person alive today will be either an ancestor of everyone or an ancestor of no one الإن. If the global population were unstructured by geography, race,

ace, religion and other factors, the time to future common ancestry for present-day humans would be between 33 and 66 generations, or about 1000 - 2000 years. In a structured population, migration and intermarriage are the necessary conditions for global common ancestry. Simulation of random and hierarchical migration models, shows that time to future global ancestry is generally less than triple, and often less than twice, that required for an unstructured population. The models suggest that someone alive today will become a common ancestor of the entire world population by about 5000 CE, or sooner; and that all current humans who are destined to become global common ancestors will be so by about 8000 CE, or sooner. At which time, everybody then alive will have the exact same genealogical ancestors from the present day.

Genealogy, Common Ancestry, Human الكلمات الدالة Population Structure, Random

routes, and number and frequency of migrants. However, under a wide variety of assumptions, they concluded that the MRCA of the current global human population likely lived between 2300 and 3400 years ago: or between approximately 300 and 1400 BCE. The same analysis [2] put the estimated date of identical ancestry between about 3000 and 5400 BCE.

The goal of this paper is to put a broad estimate on the time required for present-day humans to become future global common ancestors. The approach is simulation. I assume that human population structure, migration, and intermarriage on a global scale are complex and beyond accurate realistic simulation with modest computing facilities. Even if such simulation were feasible, current migration patterns will not persist indefinitely, and future patterns are unpredictable on the scale of generations or centuries. Instead, I will rely mostly on random models that generate complex population structure and migration patterns by virtue of incorporating a large number of subpopulations. Given that simulated complexity, I will show that the precise details of population structure, migration, and intermarriage are relatively unimportant for times to common ancestry. Consequently, random migration models may be a useful guide to the genealogical future of present-day humans.

1.2. Current Human Migration and Intermarriage Patterns

In order to provide context for model evaluation, I review some recent data on human migration and intermarriage. According to the Census Bureau, the estimated number of foreign-born persons living in the U.S. in 2018 was almost 45 million, or about 14% of the population; with the largest contributions coming from Latin America, Asia and Europe [3]. The 2011 National Household Survey reported that more than 20% of Canadians were foreign-born, with more than 200 geographic origins being represented. The Philippines, China, and India were the top three sources of newcomers between 2006 and 2011 [4]. In 2019, approximately 2 million Swedes were foreign-born (about 19% of the total population); they came from at least 199 different countries, including more than 22,000 who were born in the U.S. [5]. Between 2016 and 2018, more than 3.4 million persons received lawful permanent resident status in the U.S. They were born in more than 200 different countries [6]. In 2019, almost 104,000 foreign-born persons immigrated to Sweden, representing at least 180 countries. In that same year, almost 48,000 emigrated, including about 16,000 native-born Swedes [7].

In some Western countries at least, inter-ethnic and interracial marriage is occurring at a relatively high, and increasing, frequency. The 2011 National Household Survey reported that 4.6% of all

مقدمة

, the N In a random-mating (unstructured) population of constant size most recent genealogical common ancestor (MRCA) will have lived generations previously, or very nearly so [1]. As one N $_2$log progressively considers earlier generations, before the MRCA, the number of common ancestors of the current population increases. Eventually, a previous generation is reached in which everyone then alive is either a common ancestor of the present population, or is an ancestor of no one. In other words, everyone currently alive has identical ancestors in that previous generation, which is referred to as the most recent identical ancestry (MRIA) generation [1]. Identical ancestry of present-day individuals is then true for all generations before the MRIA generation.

Rather than looking backward in time to consider the past common-ancestry of present-day individuals, this paper adopts a forward looking perspective. It follows that in a random mating constant , one or more present-day individuals will Npopulation of size become common ancestors of the entire population approximately generations hence. Furthermore, all currently living individuals N $_2$log who are destined to become common ancestors will have done so within about twice that time. Viewed this way, it can be seen that the genealogical fate of present-day humans is binary: each of us is destined either for extinction or immortality. If your fate is extinction, at some future time, and forever after, you will not be an ancestor of anyone, anywhere—you will have no descendants. If your fate is genealogical immortality, at some future time and forever after, you will be an ancestor of everyone in the world, a global common ancestor. If your destiny is global ancestry, it will not be fully realized for many generations. How many is the principal topic of this paper. Nevertheless, whichever your fate, it will most likely be determined within a few generations. Obviously, if you leave no children, you become genealogically extinct immediately. On the other hand, if you have great grandchildren, or great great grandchildren, it is probable that you are destined for eventual global common ancestry and genealogical immortality.

Human populations are structured or subdivided by geographic distance, nationality, ethnicity, race, and other factors that affect patterns of mating. In structured populations, migration and intermarriage are the necessary prerequisites for common ancestry across subpopulations. For true global common ancestry to occur, no groups or subpopulations can remain reproductively isolated from all other groups indefinitely. Rohde *et al.* considered the MRCA and [2] MRIA times for the current global human population. Their simulations necessarily required estimates of uncertain quantities—specifically past human population sizes, migration

place within subpopulations and each offspring was produced from a randomly chosen pair of adults (no pair-bonding). Gender was not assigned and sib mating was permitted (and presumably occurred at the frequency expected by chance, although the mating scheme meant that full-sibs would be uncommon). Consequently, reproductive success (number of adult offspring) was approximately Poisson-distributed with mean 2.0. Immigrants displaced randomly selected residents in destination subpopulations; and vacancies created by emigrants were replaced by "excess" reproduction within source subpopulations. There was no general requirement that migration be symmetric or even reciprocal. An immigrant had the same probability of reproductive success as natives. In other words, not all immigrants reproduced.

Each simulation started at Generation 0 and progressed until the second of two events occurred. The first was when one or more members of the Generation 0 cohort became a common ancestor of the current generation. That is referred to as the time—number of generations—to first global common ancestry (FGCA). FGCA is a forward-looking analog of the rearward-looking MRCA. The second event occurred when all Generation 0 individuals, who had any descendants, were ancestors of everyone in the current generation. That is the time to complete global common ancestry (CGCA) of Generation 0 individuals, which is an analog of MRIA. (There is a third event of some interest: the appearance of the first common descendant (FCD) of all Generation 0 members who have surviving lineages. In an unstructured population the FCD occurs about midway between the FGCA and CGCA generations. The number of common descendants increases rapidly in subsequent generations, until all members of the population are common descendants of the Generation 0 cohort. That is equivalent to identical ancestry, and thus complete common descendancy is the same as complete common ancestry. The FCD is not considered further in this paper).

In order to illustrate the effect of population subdivision and facilitate comparisons among different migration geometries, the FGCA and CGCA times are expressed relative to the times for an undivided, random-mating population of the same total size. All results are based on 100 or 50 replicates for each set of simulated parameters—total population size, number of subpopulations, and number of migrants, etc.

2.3. Random Migration

The parameters of random migration geometries were: 1) the number The k.; and 2) the one-step path coefficient, $S$of subpopulations, number of randomly-assigned one-step, one-way links between . On average, each subpopulation sent kSsubpopulation pairs was other subpopulations and received immigrants $k$emigrants directly to - subpopulations. There was no requirement that one-$k$directly from tep migration be reciprocal.

Complete global common ancestry requires that genealogical information be able to move from every subpopulation to every other subpopulation. When genealogical information can move from Subpopulation A to Subpopulation D (by one or more steps), A is Connections are one-way. The fact that A is ."connected" to D connected to D does not necessarily mean that D is connected to A.) possible one-way $1 - S(S$ subpopulations, there are SWith connections between subpopulation pairs, by one or more paths of one or more steps. The "connectivity" of a random migration) possible one-way $1 - S(S$geometry is the proportional filling of the connections. All connections must be present (at least part of the time) in order for complete global common ancestry to occur. Connections between subpopulation pairs can be made by one or -more paths, and a path can be direct (one step), or indirect (multi

married or common-law couples in Canada were in "mixed unions", defined as unions in which at least one person belonged to a "visible minority group", or in which both people belonged to different minority groups. About two-thirds of mixed unions were composed of people born in different countries [8]. The Pew Research Center found that 17% of all newlyweds in the United States in 2015 had a spouse of a different race or ethnicity. For all married people, not just newlyweds, 10% had a spouse of a different race or ethnicity [9]. A similar rate of intermarriage occurs in Britain. The 2011 census revealed that in England and Wales about 9% of people living with a partner were in an "inter-ethnic" relationship, an increase from 7% ten years earlier [10].

These statistics are from a small and selective sample, a few countries for which data are easily retrieved using search engines. Global data on migration is available from the United Nations, which estimated that in 2015, about 3.3% of the world population were immigrants [11]. Distribution across countries was highly uneven: developed countries, such as the U.S., Canada, and Sweden, generally have a higher proportion of immigrants than less developed countries. Nevertheless, 211 of 234 countries had immigrant populations of at least 0.5%. Taken as a whole, these data indicate that many people are moving between countries and that much migration takes place over long distances on a global scale; and that intermarriage is common in at least some countries.

For the purposes of this paper, migration and intermarriage are equivalent. The everyday sense of migration is movement from one geographic location to another, and migration is necessary to transfer a genealogy (a set of ancestors) from one geographic subpopulation to another. In the same way, intermarriage is necessary to transfer ancestors from one culturally defined subpopulation to another, say between racial or ethnic groups that happen to live in the same geographic location. When "migration" appears in the text, it also generally refers to intermarriage, and vice versa.

2. METHODS

2.1. General Simulation Strategy

A primary consideration was to simulate complicated population structure and migration patterns, with the goal of achieving sufficient complexity to warrant extrapolation of results to the global human population. Complexity was generated by having many subpopulations, by random assignment of a limited number of migration paths, and by temporal variation in migration. The practical was 20,000. For simulations of N,limit on total population size, was S,common ancestry, the maximum number of subpopulations, 500. Subpopulation sizes were random, with minimum subpopulation size constrained to one-half the mean. For investigation of the , I considered *per se*characteristics of random migration geometries cases with up to 2000 subpopulations. For the genealogical simulations, one-step, one-way migration paths were assigned to randomly chosen subpopulation pairs, without regard to subpopulation size. Paths were sparse, in the sense that migration between almost all subpopulation pairs required multiple steps (generations). Migrants were randomly assigned to one-step paths. Programs were written in C language, and were compiled and run with Xcode and MacOS 10.14 (Mojave), on an iMac (model 18.3) with a 4.2 GHz Intel Core i7 processor and 64 GB memory.

2.2. Genealogical Simulation

Generations were discrete (non-overlapping) and reproductive adult population size was constant within a given simulation. Mating took

depending on overall migration rates. The total number of migrants and the distribution of migrants within and between different levels of the hierarchy were also varied. The various parameter sets are described with the results (Table 2).

3. RESULTS

3.1. Undivided Random-Mating Populations

In order to validate the simulation procedure and provide a quantitative standard for relative global ancestry times in structured populations, the case of an undivided population is briefly considered. The FGCA and CGCA times for various population sizes , respectively (Table 1). N_2 and $2\log N_2$ are just slightly more than log These results are essentially the same as Chang obtained for MRCA and MRIA times by simulation [1]. However, it should be noted that the FGCA and MRCA, for example, are not strictly the

-. Future common ancestry in an undivided random-Table 1 ating population. Each table entry is a summary of 100 replicate simulations.

same. When some Generation 0 individual becomes the FGCA of the future population, say by Generation 12, it is possible that the MRCA of Generation 12 could be a descendant of the FGCA who lived perhaps in Generation 1 or 2. Similarly, the CGCA and MRIA times are not necessarily the same.

The mean time to extinction of Generation 0 lineages is about 1.55 generations; and the probability of genealogical extinction is just greater than 20%. Accordingly, the number of Generation 0 members who eventually become global common ancestors is about 80% of the population (CGCA Size in Table 1). Mean genealogical extinction time and probability of extinction are independent of population size and depend only on the mean and variance of reproductive success (number of adult offspring). On the other hand, the average time of last extinction does increase with population size, because larger populations represent larger samples of lineages.

step). For two subpopulations, A and D, the path length from A to D is the path with the fewest number of migration steps; equivalent to the fewest number of generations required to move genealogical information from A to D. The length of path D to A, if it exists, is not necessarily the same.

The general properties of random migration geometries were examined by generating 1000 replicates for various combinations of The $k.$, and one-step path coefficient, S subpopulation number, principal properties of interest are connectivity and mean path length between connected subpopulation pairs. I examined geometries with -up to 2000 subpopulations, for which the number of possible pair), is almost 4 $I - S$(S wise connections between subpopulations, million.

Complete connectivity of random migration geometries was not one-step $k S$ required for the genealogical simulations. Therefore, the paths were re-randomized after every fourth generation of a replicate in order to ensure that every subpopulation was connected to every other subpopulation at least part of the time. Also, because migrants were assigned randomly to the one-step links between subpopulation pairs, and some links could be "empty", migrants were re-randomized every other generation. Periodic alteration of migration routes and numbers of migrants moving over those routes is desirable given that we would not expect human migration patterns to persist without change over many generations.

2.4. Hierarchical Migration

There is no assurance that purely random migration geometries capture all essential features of human migration and intermarriage a patterns. The alternative is defined models that mimic explicit patterns. The disadvantage of explicit geometries is that there *priori* are innumerable possibilities, and for every possibility there are innumerable combinations of parameter values. Furthermore, without detailed knowledge of actual human patterns, there can be little confidence that any particular geometry is appropriate. Nevertheless, one class of models that have intuitive appeal is nested, or hierarchical, geometries. These seem to be a natural choice when migration is considered in a geographic context and when migration between subpopulations declines with distance.

I considered a four-level hierarchy: subpopulations are "cities", cities belong to "states", states belong to "countries", and countries belong to "continents". This approach requires four different one-step path 4. On average, each city is a source of immigrants $- k1 k$ coefficients, 2 other cities in the same country $k1$ other cities in the same state: k to 3 other cities in a different country but on k but in a different state; $3k2 > k1 > k4$ cities on different continents: k the same continent, and 4. It seems reasonable to allow direct one-step migration between $k >$ 1 is the number of $S1 - 1$, where $S1 = k$ all cities in the same state (cities in a state). Migration paths were randomly chosen. The number 1), cities in M of migrants moving between cities in the same state (2), and so on, were chosen to M different states but the same country (ensure that most migration took place between cities in the same state and least migration took place between cities on different 4. Migrants were randomly assigned $M3 > M2 > M1 > M$ continents: to migration paths. There were fewer migrants per generation than available paths, so that many paths were unpopulated for some parameter sets. Migration paths were re-randomized every four generations and migrants re-randomized over paths every two generations. Several hierarchies were simulated, but in all cases the total number of cities (subpopulations) was 192. Total population size was 20,000. City sizes were random, with minimum size set at 50. The total number of cities was chosen so that the proportion of immigrants in a city in any generation would be about 10% or less,

exchange of five individuals each generation. 100 replicates for each combination of population size and migration rate.

In principle, a one-time reciprocal exchange of single migrants may be all that is needed to achieve global common ancestry with two subpopulations. The requirements are: 1) that complete common ancestry occur within each subpopulation before migrants are exchanged, and 2) that both migrants become common ancestors of their destination subpopulations. In that case, with two subpopulations of size *N*/2, the minimum time to CGCA is approximately 4 log $N(_2$/2), or four generations less than twice the time required for an undivided *population of size N*. (To be slightly more precise, at the time of reciprocal exchange, each migrant must be a common descendant of all Generation 0 members with surviving lineages (provided that none of those lineages becomes extinct after the one-time exchange). In that case, the minimum time to CGCA could be less than 4log $N(_2$/2)).

3.3. Characteristics of Random Migration Geometries

Over the range of subpopulation numbers investigated (50 - 2000), *k* ≥ 2, connectivity is essentially constant for a given value and for —that is, connectivity is practically independent of number of *k* of -subpopulations (Figure 2(a)). This is a striking and perhaps counter intuitive result. Notably, for values of

3.2. Two Subpopulations

Before considering highly structured populations, it will be useful to investigate the simplest geometry for a divided population: the case of two equal-sized subpopulations with symmetric, reciprocal migration. Results are shown for several total population sizes and migration rates (Figure 1). There are two striking features of these simulations. First, very little migration is sufficient to ensure relatively short times to global common ancestry: even with exchange of only one migrant every fifth generation, times to global ancestry are only about 60% longer than for an undivided population. Second, relative times to global ancestry depend on number of migrants, not (the proportion of immigrants in a *per se* migration rate subpopulation). For example, considering reciprocal exchange of single migrants per generation, the relative time to FGCA is about 1.18 - 1.19 despite a twenty-fold difference in total population size (1000 to 20,000), and hence twenty-fold difference in migration rate (Figure 1(a)). Over the same range of population sizes, the CGCA times are 1.36 - 1.40 relative to those for an undivided population (Figure 1(b)). That said, more migrants per generation do reduce times to global common ancestry: for example with reciprocal exchange of five individuals per generation (Figure 1).

. Relative time (generations) to global ancestry for two Figure 1 equal-size subpopulations as a function of total population size and number of reciprocal migrants per generation: (a) FGCA time; (b) CGCA time. Time is relative to that required for an undivided population of the same total size. Migration is symmetric. "1/5 migrants per generation" means that one person moves from Subpopulation A to Subpopulation B and vice versa every fifth generation. Similarly, 1/1 means reciprocal exchange of single individuals each generation, and 5/1 means reciprocal

. (a) Proportion of subpopulation pairs that are Figure 2 connected by migration, as a function of subpopulation number subpopulations, there *S*. For *k* and the one-step path coefficient,

> 2, Figure 3); despite the fact that mean path length k numbers (for (Figure 2(b)). $k$increases with subpopulation number and constant This perhaps counter-intuitive result is probably due to the fact that the total number of migrants increased with subpopulation number. = 50, for example, the number of migrants per S = 4 and kFor = 500 there were $S$was only 200, whereas for)$M = kS$generation (2000 migrants (corresponding to global per-generation migration rates of 1% and 10%, respectively). Additionally, as the number of subpopulations increases, and subpopulation size decreases, progress toward "local" complete common ancestry will occur more quickly within subpopulations. It is noteworthy that even when the one-step is only 1.0, common ancestry times appear to be $k,$path coefficient, largely insensitive to subpopulation number provided that migration paths are re-randomized every other generation (Figure 3).

If the number of migrants is constrained to be constant regardless of number of subpopulations, then more subpopulations means longer times to common ancestry, as expected (Figure 4(a), Figure 4(b)). However, the increase in common ancestry time is relatively modest: for a 10-fold increase in subpopulation

. Relative time (generations) to common ancestry for Figure 3 random migration models when global number of migrants per = N), kSgeneration increases with subpopulation number (= 20,000. (a) First global common ancestor (FGCA); (b) Complete -global common ancestry (CGCA). Migration paths were re randomized every four generations, and migrants randomized over paths every two generations; except for simulations = 1*, when re-randomizations occurred every two and one $k$with generation(s), respectively. Each data point is the mean of 50 replicate simulations.

number, the increase in time to common ancestry is only about 100%, or less (Figure 4(a), Figure 4(b)). When subpopulation number is held constant, more migration reduces common ancestry times, also as expected (Figure 4(c)). Again, however, the effect is not strong, particularly once the number of migrants per generation is

– 1) pairs when direction of migration is specified; (b) S($S$are Mean shortest path length between all connected subpopulation pairs; (c) Longest minimum path length between all connected subpopulation pairs. Means of 1000 replicates for each .k and Scombination of

≥ 4, connectivity approaches 1.0 regardless of subpopulation k number. In other words, it is possible for genealogical information to disperse globally even when each subpopulation is directly connected by migration to only a handful of other subpopulations. Note that = 2000 subpopulations, the number of possible pairwise $S$for connections is nearly 4 million when direction is specified, but) are one-step. Even when each kS = 4, only 8000 (=$k$with subpopulation sends migrants directly to only two other = 2), about 64% of subpopulation pairs $k$subpopulations on average (-are connected—mostly by multi-step pathways. The anomalous path = 1 (Figure 2(b), Figure 2(c)) would seem to be a $k$length results for consequence of the fact that when one-step paths are very sparse, relatively many complete paths between subpopulation pairs are only one step. Thus, complete path length is biased downward.

, mean path length increases with number of kFor constant subpopulations, as might be expected, although the increase is = 4, for example, mean path length $k$generally rather modest. With increases by a factor of less than two over a forty-fold increase in ≥ 2, path length is a linear $k$subpopulation number (Figure 2(b)). For . Thus, the predicted k, with slope dependent on S $_2$function of log = 6 $k$mean path length between connected subpopulations for = 5000 is only about 4.96 steps (verified by simulation) and $S$and = 10,000 subpopulations (unverified).Sonly about 5.35 steps for

, it might be more realistic to think kRather than assume constant increases with number of subpopulations: as the global $k$that population becomes structured into more and smaller groups, each -subpopulation will be connected to more other subpopulations by one step migration. If that is so, mean path length between pairs of subpopulations may remain relatively constant with increasing = 3, k = 100 and Snumber of subpopulations. For example, with average path length between subpopulations is about 4.1 steps. Mean = S = 4), (k = 200, Spath length remains essentially unchanged for (= 7) (Figure 2(b), k = 2000, S = 6), and (k= 1000, S = 5), (k500, some data not shown). Roughly speaking, as the number of subpopulations doubles, each subpopulation must send emigrants to one additional subpopulation, and receive immigrants from one additional subpopulation, in order for mean path length between all connected subpopulation pairs to remain constant.

3.4. Time to Common Ancestry with Random Migration Geometries

Times to future common ancestry were examined for a global population size of 20,000, and 50 - 500 subpopulations. Thus, mean subpopulation sizes varied from 400 to 40. The one-step path varied between one and six. Migration occurred every $k,$coefficient, generation, and, unless otherwise noted, one-step migration paths were re-randomized after every four generations within a replicate, and migrants were re-randomized over paths every other generation. Re-randomizations ensured that all subpopulations were sources and destinations of migrants for at least some generations—a requirement for simulations to proceed to CGCA.

If the number of migrants per generation is constrained to be equal to), $M = kS$the number of one-step paths between subpopulations (times to common ancestry show little dependence on subpopulation number, and in fact appear to decline at the largest subpopulation

. A. Relative time (generations) to common ancestry and Figure 4 **mean path length as functions of subpopulation number with** = 4, migrants per **k = 20,000,** N**constant global migration rate: generation = 400; (b) Same as (a) but 1000 migrants per generation; (c) Common ancestry times as a function of number of migrants per generation with constant number of** = 100 **subpopulations.** S = 4, k = 20,000, N**subpopulations: Common ancestry data points are means of 50 replicates, path lengths are means of 1000 replicates.**

testing was not carried out; and the effect was not large and not always statistically significant in the few cases examined (data not shown). Constant subpopulation size likely explains the effect of emigration: individuals lost to emigration were replaced by "excess" reproduction. Thus, average reproductive success (number of adult offspring) would have been slightly elevated in subpopulations that were sources of large numbers of emigrants, particularly if those same subpopulations received relatively few immigrants.

3.5. Hierarchical Migration

Times to global common ancestry for hierarchical geometries appear to be somewhat longer than for "flat" geometries with a similar number of subpopulations and overall migration rate, however the differences are not large and the details can matter (Table 2). Higher migration rate tended to shorten times to common ancestry (sets 1 and 3 vs. set 2). However, reducing total migration by a factor of two did not

four or more times the number of subpopulations.

Taken together, the results shown in Figure 3 and Figure 4 reinforce the earlier observation (in the two-subpopulation case) that low numbers of migrants are generally sufficient to ensure relatively short times to global common ancestry. For example, with total population size N = 20,000 and 100 migrants per generation, the global per generation migration rate is 0.5%. For 100 subpopulations, with k = 4, fewer than one-fourth of the 400 one-step migration paths will actually have migrants moving along them during any generation (for the case of 100 migrants). Nevertheless, relative times to common ancestry are less than 3.0 (Figure 4(c)). When the global migration rate is 400 per generation (2%) and there are 200 subpopulations (Figure 4(a)), or when the global migration rate is 1000 per generation (5%) and there are 500 subpopulations (Figure 4(b)), relative times to common ancestry are only about twice that of an undivided population (with k = 4 in both cases). Increased one-step connectedness (larger values of k) reduced times to global common ancestry, as expected (Figure 3). However, it is noteworthy that even with very low connectedness, such as k = 2, relative times to common ancestry never exceeded 2.4 for the range of subpopulation numbers simulated; and for k > 2, common ancestry times were 2.0 or less (at least when number of migrants per generation was S, Figure 3). (As a practical matter, if the number of migrants per generation is much less than the number of subpopulations, global common ancestry cannot be achieved in a reasonable time—even when migrants are re-randomized over one-step paths every other generation).

The probability of eventual common ancestry was unaffected by population structure. In particular, with random reproductive success and constant population size, about 80% of Generation 0 members became global common ancestors, the same as for an unstructured population. Subpopulation size did not affect the probability of common ancestry. Generation 0 individuals in subpopulations that were sources of large numbers of emigrants had a slightly increased probability of becoming common ancestors. Systematic

common ancestry.

The results for hierarchical migration geometries do not greatly alter the conclusions derived from random models. Human migration patterns are no doubt influenced by immigration laws and national boundaries. However, it is not clear that migration within many countries is likely to be restricted in a hierarchical sense. International migrants may be more likely to move to nearby countries, or at least to countries on the same continent. But, currently, there is much intercontinental migration. In other words, to the extent that human migration is hierarchical, the hierarchy may be relatively flat: something more like set 6 than set 5 (Table 2). Or perhaps better characterized with just two levels—countries and continents.

4.2. Do We Need to Know Migration and Intermarriage Rates?

Not only is the appropriate number of human subpopulations unknown, the details of migration and intermarriage patterns between subpopulations may often be obscure. I have already argued that times to common ancestry are broadly insensitive to migration rates, and very low migration rates are sufficient to produce relative common ancestry times of 2.0 or less (Figure 3 and Figure 4). In the random migration models simulated here, the one-step path coefficient, k, represents the average number of one-step, one-way connections between subpopulations. That is, on average, each subpopulation k receives migrants directly from other subpopulations and sends migrants to k other subpopulations. It is not obvious what value k is appropriate for modeling human migration and intermarriage. It is fortunate, therefore, that at least in random migration geometries, the value of k has little effect on times to common ancestry, provided that $k \geq 4$ (Figure 3). That is perhaps because when $k \geq 4$, overall connectivity between all possible subpopulation pairs (by single or multi-step pathways) approaches 1.0, even with 2000 subpopulations (Figure 2(a)); and mean path length (migration steps) between subpopulations is generally less than 5.0 (Figure 2(b)). In other words, if random models are a useful guide, modest values of k ensure that genealogical information can move from every subpopulation to every other subpopulation, and the required number of migration steps (generations) is likely to be five or less.

Again, the hierarchical models do not substantially alter the picture. It is true that the overall one-step path coefficients, k_{total}, were greater than for the purely random geometries (Table 2). That is justified on the grounds that one-step migration would be expected between all cities in the same state, and perhaps between all cities in the same country (although k was never set at that high level). For a similar reason, overall migration rates were higher in the hierarchical models, but most of that migration took place within countries. More to the point, within the group of hierarchical simulations, substantial differences in migration rates did not have large effects on relative times to common ancestry.

4.3. Relaxing Model Assumptions

4.3.1. Increasing or Decreasing Global Population

As already noted, in an undivided random-mating population of constant size, N, the time to FGCA is very nearly $\log N$ generations. N_2, But, log N_2 generations is just the time when an "average" Generation 0 member is expected to have N descendants. Formally, at Generation T, an "average" individual will have 2^T descendants, and $N = 2^{\log_2 N}$. In other words, the number of generations required for some member of Generation 0 to become an FGCA is the number of generations required for an "average" individual to have N living descendants.

. Parameter sets and results for hierarchical migration Table 2 geometries. Each results entry is a mean of 50 replicate simulations.

substantially alter common ancestry times when long distance migration was increased (set 4 vs. set 2). When the number of continents was greater and the number of cities per state less, times to common ancestry increased (set 5 vs. set 6). The longest relative time to FGCA (2.76) was obtained with set 5, the one with four continents and only four cities per state. For all sets, the migration rate between continents was 0.05% - 0.1% per generation, and the migration rate between countries on the same continent was 0.1% - 0.4%. These rates seem broadly consistent with what we know about current human migration patterns: 0.1% of the global human population is about 8 million. Set 3 had the greatest difference between short and long-distance migration rates ($M1/M4 = 100$) but also the highest overall migration rate (10%). For that set, the relative FGCA and CGCA times were 1.72 and 1.60, respectively. In contrast, the parameters for set 4 were the same as for set 3, except that overall migration was reduced five-fold to 2% by drastically reducing migration $M1$ between cities within the same state and country ($M2$). Nevertheless, the relative FGCA and CGCA times increased to only 2.40 and 2.17 (Table 2).

Overall connectivity was high (generally 100%) for the hierarchical models: that is, all possible subpopulation (city) pairs were connected by migration paths (Table 2). Also mean and longest minimum path lengths were relatively short. That is probably due to the fact that each city was connected by one step paths to all or most other cities in the same state and to several cities in the same country.

4. DISCUSSION

4.1. Do We Need to Know the Details of Human Population Structure?

Human populations are structured along numerous dimensions that result in patterns of assortative mating: obvious axes include nationality, geographic proximity, race, ethnicity, socio-economic class, and religion, to name a few. Furthermore, everybody belongs simultaneously to several groups that might be bases for assortative mating: for example American, Californian, Black, Protestant, politically conservative. Thus, it is not obvious how many "subpopulations" or groups are appropriate for modeling migration and intermarriage in the human population. It is important, therefore, that times to common ancestry in the models presented here are generally insensitive to subpopulation number. That is especially true for random migration geometries when global migration rate is allowed to increase with subpopulation number: in that case, common ancestry times are practically independent of subpopulation number (Figure 3).

Does migration rate increase with subpopulation number? Arguably, yes: if, for example, a geographic region is divided into more and more smaller governmental jurisdictions (countries, provinces, states, etc.), it seems reasonable that movement between jurisdictions will become more likely. Similarly, as human populations become more finely divided along culturally-defined dimensions (ethnicity, religion, class, etc.), it seems reasonable that intermarriage will be more common. In other words, whether or not an individual is classified as a "migrant" will depend to a considerable extent on the way that we define "subpopulations". The more fine-grained and numerous subpopulations become, the more likely that individuals will be migrants. If that is true, then the details of global human population structure may be relatively unimportant for times to future

common ancestry are typically less than twice those for an undivided population, and almost always less than triple (Figure 3 and Figure 4; Table 2). For purely random models, longer times to common ancestry occur when the global number of migrants per generation is about the same as, or less than, the number of subpopulations; and when each subpopulation sends or receives migrants from only one other subpopulation. For hierarchical models, longer times occur when the number of higher level subdivisions (continents) is relatively large compared to the number of lower level subdivisions (cities). The available evidence suggests that the current human population is not characterized by such extremes. Therefore, it seems likely that people alive today (equivalent to Generation 0 in my simulations) will become common ancestors of the future global population, and achieve genealogical immortality, sooner rather than later.

I will hazard a rough estimate that the time will be on the order of twice to triple that required for an unstructured, random-mating global population. In round numbers, the current human population is about 8 billion. If it were unstructured, the time until at least one present-day person became a common ancestor of the future global population 8 billion, or approximately 33, $_2$would be about log generations. The time to complete global common ancestry of all people alive today, whose genealogies do not go extinct, would be about twice that, or about 66 generations. These times are, respectively, about 1000 and 2000 years, if generation time is 30 years. If we double and triple these estimates, the suggested time for initial global common ancestry for someone alive today will be about 2000 - 3000 years; and the time to complete global common ancestry for present-day humans will be about 4000 - 6000 years. The usual caveat applies: true global common ancestry requires that no group remain indefinitely isolated from the rest of humanity.

If these times seem impossibly short, then it may help to realize that an "average" person destined for genealogical , or more than $7 \times$ ^{66}immortality will have about 2 descendants living in the year 4000 CE (assuming ^{19}10 (or ^{200}constant global population size). That number grows to 2) by year 8000 CE. (Obviously, if the future global ^{60}1.6 × 10 population is only 8 billion, it would be impossible for unique concurrently-living ^{200} or 2^{66}someone to have 2 descendants. It is more correct to say that there will be genealogical paths linking an "average" person today to his ^{66}2 or her (several billion) descendants living 66 generations hence.) Coincidentally, these proposed times to future common ancestry are very similar to the ages of the most recent common ancestor (MRCA) and most recent identical ancestor cohort (MRIA) of the current global human population [2 , 12]. Until recently, historically speaking, the global human population was much smaller than it is now. On the other hand, long-distance migration was also much more difficult, if not impossible, for much of the last six thousand years. Regardless of the exact time required, eventually everyone in the world will be a descendant of just about everybody alive today, famous or not, who has a few adult grandchildren or great grandchildren. The more adult children that you have, the more likely that you will become a global

descendants. If a similar logic applies to growing and declining populations—an FGCA will occur when an average member of descendants—it is easily shown that the FGCA $_TN$Generation 0 has time in an undivided population, growing or shrinking at a constant is the starting population size. $_0N$ generations, where $_0N$ $_2$rate, is log Similar reasoning suggests that the CGCA time will be about 2 generations. Assuming that population subdivision does not $_0N$ $_2$log alter this line of reasoning, the present results apply to growing and declining populations, with the number of generations to common ancestry depending on starting population size.

4.3.2. Effect of Inbreeding

Inbreeding decreases the number of distinct ancestors in previous generations [12], and also decreases the number of distinct descendants in future generations. Thus the general effect of inbreeding is to increase times to common ancestry. As already noted, inbreeding was permitted in these simulations. However, because mating was promiscuous, full-sibs were likely to be rare, as were full first-cousins, and so on. Levels of inbreeding, therefore, were probably low, and may have been less than typical for human populations [13]. Thus the GCA times in these simulations may be underestimates, although probably by not more than ten percent, or less [12].

4.3.3. Non-Poisson Distribution of Reproductive Success

Reproductive success—number of adult offspring—in these simulations is a Poisson-distributed random variable with mean and variance 2.0. As a result, approximately 13.5% of individuals in every generation have no children. If the proportion of childless people were increased, the average number of offspring among those individuals who do have children would increase (assuming population size remains constant). Consequently, descendant (and ancestor) trees would grow more rapidly and GCA times would decrease. The opposite happens if fewer people are childless. In the extreme case, everybody has two children and variance in reproductive success is zero. In that case, times to complete common ancestry can increase by as much as 38% (at least in an unstructured population) [14]. A global survey of mean and variance of number of adult offspring in human populations is beyond the scope of this paper. However, a 2014 survey indicated that 16.7% of women aged 45 - 50 in the United States were childless [15]. That suggests that Poisson-distributed reproduction in these simulations is a reasonable approximation to human populations.

استنتاج

HOW LONG UNTIL PRESENT-DAY HUMANS BECOME GLOBAL COMMON ANCESTORS?

On a global scale, current human migration appears substantial and universal (at the level of countries). There is, of course, no assurance that current patterns and rates of international migration will persist indefinitely. It is fortunate, therefore, that the models presented here argue strongly that the details of population structure and migration are simply not very important over a wide range of values for migration rates, number of subpopulations, average number of one-step connections, and hierarchical arrangement of subpopulations. Except for extreme cases—very low migrant numbers or very ≤ 2)—the times to global $k$small one-step path coefficients (

https://doi.org/10.1080/08898480903251520

United States Census Bureau (2014) Fertility of Women in the United States. Table 6, Completed Fertility for Women 40 to 50 Years Old by Selected Characteristics: June 2014. https://www.census.gov/data/tables/2014/demo/fertility/women-fertility.html

common ancestor. Your chances might be increased somewhat if you and your near-term descendants are members of groups or subpopulations with higher than average population growth and emigration rates. Little else likely matters.

ACKNOWLEDGEMENTS

ref_str

1. **Chang, J.T. (1999)** Recent Common Ancestors of All Present-Day Individuals. Advances in Applied Probability, 31, 1002-1026. https://doi.org/10.1239/aap/1029552561

2. **Rohde, D.L.T., Olson, S. and Chang, J.T. (2004)** Modelling the Recent Common Ancestry of All Living Humans. Nature, 431, 562-566. https://doi.org/10.1038/nature02842

3. **United States Census Bureau (2018)** Selected Social Characteristics in the United States. Table ID: DP02. https://data.census.gov/cedsci/table?q=Table%20ID%3A%20DP02&g=0400000US16&tid=ACSDP1Y2018.DP02&hidePreview=false

4. **Statistics Canada (2011)** Immigration and Ethnocultural Diversity in Canada. National Household Survey, Catalogue No. 99-010-X2011001. https://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-010-x/99-010-x2011001-eng.pdf

5. **Statistics Sweden (https://www.scb.se/) (2020)** Population by Country of Birth, Age and Sex. Year 2000-2020. http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__BE__BE0101__BE0101E/FodelselandArK/

6. **United States Department of Homeland Security (2018)** Table 3. Persons Obtaining Lawful Permanent Resident Status by Region and Country of Birth: Fiscal Years 2016 to 2018. https://www.dhs.gov/immigration-statistics/yearbook/2018/table3

7. **Statistics Sweden (https://www.scb.se/) (2020)** Immigrations and Emigrations by Country of Birth and Sex. Year 2000-2020. http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__BE__BE0101__BE0101J/ImmiEmiFod/

8. **Statistics Canada (2011)** Mixed Unions in Canada. National Household Survey, Catalogue No. 99-010-X2011003. https://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-010-x/99-010-x2011003_3-eng.pdf

9. **Livingston, G. and Brown, A. (2017)** Intermarriage in the U. S. 50 Years after Loving v. Virginia. Pew Research Center, Washington DC. https://www.pewsocialtrends.org/2017/05/18/intermarriage-in-the-us-50-years-after-loving-v-virginia/

10. **Office for National Statistics (2014)** What Does the 2011 Census Tell Us about Inter-Ethnic Relationships? https://webarchive.nationalarchives.gov.uk/20150505074206/http://www.ons.gov.uk/ons/dcp171776_369571.pdf

11. **United Nations, Department of Economic and Social Affairs, Population Division (2015)** International Migrant Stock 2015. https://www.un.org/en/development/desa/population/migration/data/estimates2/estimates15.asp

12. **Lachance, J. (2009)** Inbreeding, Pedigree Size, and the Most Recent Common Ancestor of Humanity. Journal of Theoretical Biology, 261, 238-247. https://doi.org/10.1016/j.jtbi.2009.08.006

13. **Freire-Maia, N. (1982)** Inbreeding: Inbreeding Levels in Different Countries. Social Biology, 29, 69-81. https://doi.org/10.1080/19485565.1982.9988479

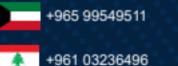14. **Campbell, R.B. (2009)** Time since Common Pedigree Ancestors with Two Progeny per Individual. Mathematical Population Studies, 16, 248-265.