Modeling the recent common ancestry of all living humans

Supplementary Methods A:
Further Explanation and Derivations of Mathematical Results

Douglas L. T. Rohde, Steve Olson, Joseph T. Chang

The purpose of this note is to explain the derivations of the mathematical results in the paper. Much of the reasoning follows similar lines and uses similar techniques to those presented in full detail in the earlier paper [1]. Here we will draw freely upon results and arguments from that paper, and some of those arguments will be sketched in a less rigorous way here.

One of the principal conceptual messages of the results discussed here about the simple graph models is that a seemingly rather severe form of population subdivision can still be compatible with recent common ancestors. In these simple models, the population is divided up into subpopulations that exchange migrants very infrequently. We assume some small fixed number of migrant individuals per generation; for example, that number could be just one migrant per generation, of even smaller.

The model begins with a connected graph $\mathbb{G}$ consisting of G nodes, which we will refer to here as "islands," with a constant population size of n / G on each island. This is a discrete-time model with time measured in generations. We could choose to call an arbitrary generation $t = 0$, and then t increases by 1 whenever time proceeds forward by one generation. Each individual lives on a particular island (the individual's "home island") in a particular generation. We will use the notation $I(t, i, m)$ to refer to individual number m on island i in generation t.

Each individual has two parents in the previous generation. The two parents are chosen independently, both according to the following probabilistic process. There is a "migration probability" which we will denote by μ_n. With probability μ_n, an individual's parent is chosen from a different island from the home island of the individual, in which case the parent's island is equally likely to be any of the neighbors of the individual's home island (where neighbors are determined by the edges in the graph $\mathbb{G}$). With probability $1 - \mu_n$ an individual's parent is chosen from the same island as the individual. In either case – whether chosen from the individual's home island or a neighboring island – the parent is taken to be uniformly distributed, that is, equally likely to be any of the n / G individuals on the chosen island in the previous generation. In other words, to choose a parent of individual $I(t+1, i, m)$, we would first choose a random index m' uniformly distributed over the set $\{1, 2, \ldots, \frac{n}{G}\}$ and then with probability $1 - \mu_n$ take the parent to be $I(t, i, m')$, and with probability μ_n, choose an island j randomly from among the neighbors of i and take the parent to be $I(t, j, m')$. For each individual, two parents are chosen in this way.

We call an individual a *migrant* if at least one of that individual's parents is from an island other than the individual's home island. (It might be more natural to call the parent the migrant, but we will retain this terminology here as it has been convenient.)

Here we take the migration probability μ_n to be of order $1 / n$ by letting c be a constant and taking $\mu_n = c / n$. With the idea of modeling a strongly subdivided population, we are letting the migration rate have a very small order of magnitude. For example, we could choose c so that the expected number of migrants in the whole population is just 1 per generation. Or we could choose c one tenth as large, which would model a situation in which in a span of ten generations just one migrant is expected in the whole population.

As defined in the paper, a *common ancestor* of a given set of individuals is an individual who is an ancestor of everyone in the set. For example we will speak of a common ancestor of everyone on a particular island at a particular time, or the whole population at a particular time. We use "CA" as an abbreviation for common ancestor. T_n is the number of generations back to the most recent common ancestor (MRCA) of the population. U_n is the number of generations back to the "IA point," the most recent generation in which all current individuals have identical ancestors.

We will use standard notation related to orders of magnitude and asymptotic behavior as $n \to \infty$:

- $f(n) = o\left(g(n)\right)$ means $\dfrac{f(n)}{g(n)} \to 0$ as $n \to \infty$,
- $f(n) = O\left(g(n)\right)$ means $\dfrac{f(n)}{g(n)}$ is bounded,
- $f(n) \asymp g(n)$ means that both $f(n) = O\left(g(n)\right)$ and $g(n) = O\left(f(n)\right)$ hold, and
- $f(n) \sim g(n)$ means $\dfrac{f(n)}{g(n)} \to 1$ as $n \to \infty$.

For notational convenience we will omit writing the obvious "greatest-integer" type functions that are needed in order to round real numbers into integers, such as in the phrase "in generation $t = (1-\varepsilon)\left(D+1+\zeta\right)\log_2 n$."

Statements of Results

We are given a connected graph $\mathbb{G}$ and define the distance $d(i,j)$ to be the number of edges in a shortest path joining i and j. This definition is extended to sets of nodes by considering shortest paths joining some node of one set to some node of the other set. That is, for sets of nodes A and B we define $d(A,B) = \min\{d(i,j) : i \in A, j \in B\}$, and as a special case, for a set of nodes A, we define $d(A,j) = \min\left\{d(i,j) : i \in A\right\}$.

The *radius* of $\mathbb{G}$ is $R = \min_{i \in \mathbb{G}}\{\max_{k \in \mathbb{G}} d(i,k)\}$, and a node i is called a *center* of $\mathbb{G}$ if $\max_{k \in \mathbb{G}} d(i,k) = R$. Let $C(\mathbb{G})$ denote the collection of all centers of $\mathbb{G}$. The *diameter* of $\mathbb{G}$ is $D = \max_{i \in \mathbb{G}}\{\max_{k \in \mathbb{G}} d(i,k)\}$.

We assume throughout that $R > 0$; the case $R = 0$ (that is, $\mathbb{G}$ has just one node) was treated in [1].

For $i \in C(\mathbb{G})$, let S_i be a set of minimal size that consists of neighbors of node i and satisfies $\max_{j \in \mathbb{G}} d(\{i\} \cup S_i, j) = R - 1$. Define H_i to be the number of nodes in S_i, $\Delta_i = (H_i - 1)/H_i$, and $\Delta = \min_{i \in C(\mathbb{G})} \Delta_i$. Note $0 \le \Delta < 1$.

The results are asymptotic, with the number of islands and the graph fixed, and n tending to infinity. We use "$\lg$" to denote the base-2 logarithm.

Theorem 1: $\frac{T_n}{(R+\Delta)\lg n}$ converges in probability to 1 as $n \to \infty$.

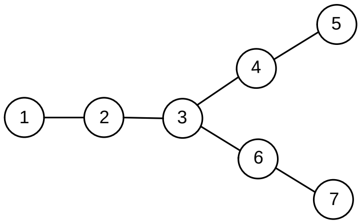
Theorem 2: Let $\zeta \approx 0.7698$ be as defined in Theorem 2 of [1]. $\frac{U_n}{(D+1+\zeta)\lg n}$ converges in probability to 1 as $n \to \infty$.

To attempt a quick and very rough explanation of the theorems in a nutshell: the main idea of Theorem 1 is that every $\lg n$ generations, the set of nodes occupied by descendants of any given individual expands to include all of its neighbors. One can imagine this as a set that expands like clockwork, with a clock that ticks once every $\lg n$ generations. At time 0 the set starts out including only one node. With each tick of the clock, the set expands to encompass all neighbors of nodes currently in the set. Applying this idea to the individuals on a center node of a graph of radius *R* gives the essence of the result: in roughly *R* ticks of the clock, or $R \lg n$ generations, this set of nodes expands to include the whole graph. Similar comments apply to give a rough explanation of Theorem 2 – at least why the diameter appears. The longest path in the graph is of length D. So after the clock ticks about D times, everyone who is destined to become a CA of the full population should have done so.

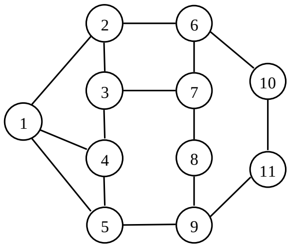
A few more remarks about the results are in order here. First, the reason that the process behaves much like a regularly ticking clock is related to the fact that the distribution of T_n is concentrated around $\lg n$, with little variability. Second, as in the case where $\mathbb{G}$ has one node [1], at the IA point about 80 percent of the population in the structured model consists of common ancestors of everyone in the population today and the lineages of the remaining 20 percent have gone extinct. The derivation will also show that as n increases MRCAs are increasingly likely to be found in center nodes of the graph -- in particular, in nodes i that minimize H_i.

Examples

In the graph shown at right, $R = 2$, and the center node 3 has $H = 3$ neighbors 2, 4, and 6, such that the set $S_3 = \{3,2,4,6\}$ lies within $R-1=1$ of each node, that is, $\max_{j \in G} d(S_3, j) = R-1 = 1$. So here, $\Delta = (H-1)/H = 2/3$, and $T_n \sim (2+\frac{2}{3})\lg n$. Since the diameter of this graph is 4, we have $U_n \sim (5+\zeta)\lg n$.



For another example, consider the 11-node graph at right. For this graph, the radius is $R = 3$ and the centers are nodes 1, 2, 5, 6, 7, 8, and 9. It turns out that each of the centers i has $H_i = 2$. For example, node 1 has $S_1 = \{2,5\}$; in fact every node is within $R-1=2$ of the set $\{2,5\}$. Node 6 has $S_6 = \{2,7\}$; every node is within 2 of the set $\{2,6,7\}$. For this graph, $H = 2$, $\Delta = \frac{1}{2}$, and $T_n \sim (3+\frac{1}{2})\lg n$. The diameter of this graph is 4; for example, $d(3,11) = 4$. Therefore, $U_n \sim (5+\zeta)\lg n$.



In a complete graph having $G > 1$ nodes, each pair of nodes is joined by an edge. Such a graph has radius $R = 1$ and diameter $D = 1$. Also, for each i the set S_i is simply $\{1, \ldots, G\} - \{i\}$, so that $H_i = G - 1$, $\Delta_i = (G-2)/(G-1)$, and $\Delta = (G-2)/(G-1)$. Thus, $T_n \sim (1 + (G-2)/(G-1)) \lg n$; for example, for $G = 2$ we have $T_n \sim \lg n$, for $G = 3$ we have $T_n \sim (1 + \frac{1}{2}) \lg n$, and for $G = 5$ we have $T_n \sim (1 + \frac{3}{4}) \lg n$. The result for U_n is $U_n \sim (2 + \zeta) \lg n$.

More on Simulations and Approximations

The above theorems give the main term of asymptotic results of the form $T_n \sim (R + \Delta) \lg(n)$ as $n \to \infty$, for example, which do not distinguish among various possible explicit forms for the lower order effects. For example, the statements $T_n \sim (R + \Delta) \lg(n) + 2.7$ and $T_n \sim (R + \Delta) \lg(n) - 8.3$, which differ by having different additive offsets, would both be consistent with the theorem. However, we can find some simple formulas that agree with the asymptotic results in the theorems, make intuitive sense, and provide rather good approximations to the simulation results. Defining G as above to be the number of nodes in the graph, if we use the approximations $T_n \sim (R + \Delta) \lg(n / G)$ and $U_n \sim (D + 1.77) \lg(n / G)$, we get the following predicted values for the simulation results in Table 1 and Table 2.

Predictions for Table 1 using the formula $T_n \sim (R + \Delta) \lg(n / G)$:

	n=1000	*2000*	*4000*	*8000*	*16000*
One node	10.0	11.0	12.0	13.0	14.0
Three fully-connected nodes	13.4	14.9	16.4	17.9	19.4
Five fully-connected nodes	14.7	16.4	18.2	19.9	21.7
Ten-node graph in Figure 1	19.9	22.9	25.9	28.9	31.9

Predictions for Table 2 using the formula $U_n \sim (D + 1.77) \lg(n / G)$:

	n=1000	*2000*	*4000*	*8000*	*16000*
One node	17.6	19.4	21.2	22.9	24.7
Three fully-connected nodes	24.8	27.6	30.4	33.1	35.9
Five fully-connected nodes	23.2	26.0	28.8	31.5	34.3
Ten-node graph in Figure 1	45.0	51.7	58.5	65.3	72.1

Comparing these predictions to the actual simulation results in Table 1 and Table 2 of the paper, we see that the agreement is quite good, with the predictions capturing the main features of the simulation results.

More Terminology, Head Starts and the Idea of Delta

To introduce convenient terminology for some ideas that were introduced in [1] and are helpful here, let us say that a given individual $I(t_1, i_1, m_1)$ is *established* on a given island i_2 in a given generation t_2 if the number of descendants of $I(t_1, i_1, m_1)$ on island i_2 in generation t_2 is greater than $\lg^2(n)$. We say $I(t_1, i_1, m_1)$ is established in generation t_2 (without reference to a particular island i_2) if $I(t_1, i_1, m_1)$ is established on some island in generation t_2. We say an individual is *in jeopardy* in a given generation if that individual is not established on any island in that generation and is also not extinct in that generation.

The idea is that an individual who is *not* in jeopardy at a given time is either already extinct, which means he has no surviving descendants at that time, or established, which means that his number of descendants is large enough to assure that it is very unlikely that the individual will become extinct in the future. That is, individuals who are not in jeopardy are either extinct or are very likely to become CA's. We say an individual is *destined* to become established, a CA, and so on, if these events will occur for that individual in the future.

Let us say that a *"head start"* has been completed for a center island i as soon as some individual on island i has migrant descendants who have become established on each of the islands in some set S_i consisting of neighbors of node i and satisfying $\max_{j \in \mathbb{G}} d(\{i\} \cup S_i, j) = R - 1$. We will express the time at which some individual from island i becomes a CA of the whole population as the sum of the time required to complete a head start for i and the additional time required to complete the process of becoming a CA after completing the head start.

Proposition H: For a center island i, the time required to complete a head start is $\sim \Delta_i \lg n$. That is, for $0 < \alpha < 1$, as $n \to \infty$, the probability that at least one individual from island i has completed a head start by time $\alpha \lg n$ converges to 0 if $\alpha < \Delta_i$ and converges to 1 if $\alpha > \Delta_i$.

To verify the proposition, consider the individuals on island i at time 0. Those individuals fall into a number of categories, defined in terms of the number of descendants those individuals have at time $\alpha \lg n$. As shown in [1], a fraction of nearly $1 - \rho \approx 0.8$ of those individuals will have become established by time $\alpha \lg n$, with their descendants having grown geometrically to reach a size of $\asymp 2^{\alpha \lg n} = n^{\alpha}$. Most of the remaining individuals will have become extinct, and a few will be in between (including that small set of people who might remain in jeopardy, depending on the value of α). Consider one of these established individuals who has a number of descendants whose order of magnitude is n^{α}. Each of these descendants has probability $\asymp 1/n$ of being a migrant to a neighboring island. So the probability that the individual has migrant descendants on each of the H_i neighboring islands in the "head start" set S_i is $\asymp (n^{\alpha}/n)^{H_i} = n^{(\alpha-1)H_i}$. Consequently, the expected number of individuals who have migrant descendants on each of the H_i neighboring islands in S_i by time $\alpha \lg n$ is $\asymp n^{1+(\alpha-1)H_i}$. Therefore, if $1 + (\alpha - 1)H_i < 0$, that is, $\alpha < \Delta_i = (H_i - 1)/H_i$, then the expected number of individuals having completed a head start by time $\alpha \lg n$ converges to 0, so that the probability that any of the n individuals on island i has completed a head start by time $\alpha \lg n$ converges to 0 as $n \to \infty$. On the other hand, if $1 + (\alpha - 1)H_i > 0$, that is, $\alpha > \Delta_i$, then the expected number of individuals who have completed a head start by time $\alpha \lg n$ grows to infinity as $n \to \infty$. From this, together with a demonstration of asymptotic pairwise independence among the events that different individuals complete a head start by time $\alpha \lg n$ (along lines similar to the proofs of Lemmas 19 and 20 of [1]), it follows that the probability that at least one individual has completed a head start by time $\alpha \lg n$ converges to 1.

Lower bound in Theorem 1

Recall we call a given individual a *migrant* if either parent of that individual lived on an island other than the given individual's home island. Note that since all of our results concern time spans that are only order $\lg n$, there are only order $\lg n$ migrants in total.

Proposition L: Let s be a positive integer and let $\varepsilon > 0$. The probability that there exists an individual at a given time t_0 who has any descendant on an island at distance s steps away from that individual's home island within $(1-\varepsilon)(s-1)\lg n$ generations approaches 0 as $n \to \infty$.

To see this, note that if there is such an individual $I(t_0, i_0, m_0)$, then there must be a path of islands $i_0 \to i_1 \to \cdots \to i_s$ and a chain $I(t_1, i_1, m_1), I(t_2, i_2, m_2), \ldots, I(t_s, i_s, m_s)$ of migrant descendants of $I(t_0, i_0, m_0)$, with each $I(t_k, i_k, m_k)$ being a descendant of $I(t_{k-1}, i_{k-1}, m_{k-1})$, and $I(t_k, i_k, m_k)$ having a parent from island i_{k-1}. We know nothing particular about t_1 except that $t_1 - t_0 > 0$ (for example, it is likely that we could find individuals at time t_0 who have a migrant child, so that we could have $t_1 - t_0 = 1$). However, for each $k \ge 2$, we claim that with probability approaching 1, none of the differences $t_k - t_{k-1}$ will be less than $(1-\varepsilon)\lg n$. In fact, with probability approaching 1, no migrant up to a time of order $\lg n$ can have any further migrant descendants within $(1-\varepsilon)\lg n$ generations. That is, the probability that there exist times $t < t' < s\lg n$ satisfying $t' - t \le (1-\varepsilon)\lg n$ and there exist migrant individuals $I(t, i, m)$ and $I(t', i', m')$ with $I(t', i', m')$ being a descendant of $I(t, i, m)$ approaches 0 as $n \to \infty$. The reason for this is that for any given individual, the probability of having a migrant descendant within $(1-\varepsilon)\lg n$ generations is of order $n^{-\varepsilon}$, so that since there are only order $\lg n$ migrants in total up to time $s\lg n$, the probability of *some* migrant having a migrant descendant within $(1-\varepsilon)\lg n$ generations is $O(n^{-\varepsilon}\lg n)$, which converges to 0.

Lower bound: For $\varepsilon > 0$, we have $P\{T_n \ge (1-\varepsilon)(R+\Delta)\lg n\} \to 1$ as $n \to \infty$.

To prove this, fix a node i_0. We want to show that the probability that some individual on i_0 becomes a CA within $(1-\varepsilon)(R+\Delta)\lg n$ generations approaches 0. We will assume i_0 is a center of the graph. This is the more involved case; we omit the similar but easier proof for a node that is not a center of the graph.

Recall we say that a head start has been completed when some individual has descendants who are established on each island in a set of islands that is within a distance of $R-1$ from every node of the graph. Let τ denote the time required for some individual living on i_0 at time 0 to have completed a head start. That is, before time τ, no individual on i_0 has completed a head start.

Let $t_0 = (1-\varepsilon)\Delta\lg n$. We have

$$P\{T_n \le (1-\varepsilon)(R+\Delta)\lg n\} \le P\{\tau \le t_0\} + P\left[\{\tau > t_0\} \cap \{T_n \le (1-\varepsilon)(R+\Delta)\lg n\}\right].$$

By Proposition H, $P\{\tau \le t_0\}$ approaches 0 as $n \to \infty$.

Next we want to show that the probability of the event $\{\tau > t_0\} \cap \{T_n \le (1-\varepsilon)(R+\Delta)\lg n\}$ converges to 0. For convenience, let $I(0, i_0, 1)$ denote whichever of the individuals on island i_0 at time 0 becomes a CA the fastest (if there is a tie, choose one arbitrarily). Let *A* denote the set of islands reached by the descendants of $I(0, i_0, 1)$ by time t_0. Let j be an island such that $d(A, j) = R$; we know we can find such a j whenever $\tau > t_0$. We observe that for both events $\{\tau > t_0 = (1-\varepsilon)\Delta\lg n\}$ and

$\{T_n \le (1-\varepsilon)(R+\Delta)\lg n\}$ to occur it must be the case that at least one of the following two events occurs:

(1) There exist $i \in A$ and m such that individual $I(t_0, i, m)$ has a descendant on island j within $(1-\varepsilon)(R-1)\lg n$ generations

(2) The remaining time, after the first migrant descendant of some individual on island i_0 reaches island j, required for some individual on i_0 to become a CA of j, is at most $(1-\varepsilon)\lg n$ generations. In other words, some set of migrants to island j "collectively" become a common ancestor of island j within $(1-\varepsilon)\lg n$ generations, in the sense that there is a set of migrants $\{I(u_1, j, m_1), I(u_2, j, m_2), \ldots, I(u_K, j, m_K)\}$ with $u_1 \le u_2 \le \cdots \le u_K$ such that the union of their sets of descendants at time $u_1 + (1-\varepsilon)\lg n$ is everyone on j. That is,

$$\bigcup_{k=1}^{K} \{\text{descendants of } I(u_k, j, m_k) \text{ on island } j \text{ at time } u_1 + (1-\varepsilon)\lg n\}$$
$$= \{I(u_1 + (1-\varepsilon)\lg n, j, m) : 1 \le m \le n/G\}.$$

By Proposition L and because $d(A, j) = R$, the probability of the event in (1) approaches 0 as $n \to \infty$.

For the event in (2) to occur, clearly at least one migrant to island j must have at least order $n/(\lg n)$ descendants within $(1-\varepsilon)\lg n$ generations, since there are a total of only order $\lg n$ migrants to island j within the time span relevant to the desired result. But the proof of Proposition 15 of [1] shows that the probability that a migrant can have order $n/(\lg n)$ descendants within $(1-\varepsilon)\lg n$ generations is very small – in fact, it is of order $o(n^{-p})$ for all p. So taking the union over the order $\lg n$ migrants still gives a probability of $o(n^{-p}\lg n)$, which approaches 0.

Upper bound in Theorem 1

Proposition U: Suppose an individual $I(0, i, m)$ from island i is established on island j in a given generation t, and let $\varepsilon > 0$. Then, with probability that approaches 1 as $n \to \infty$, by generation $t + (1+\varepsilon)\lg n$, individual $I(0, i, m)$ will be a CA of island j and will be established on each island neighboring island j.

Arguments from the upper bound in Theorem 1 of [1] applied to this setting show that with probability approaching 1, within $(1+\varepsilon/2)\lg n$ generations, $I(0, i, m)$ will become a CA of island j. And within an additional $(\varepsilon/2)\lg n$ generations [in fact in $o(\lg n)$ generations], among the descendants of $I(0, i, m)$ will also be migrants who have become established on each island neighboring j. In fact, there are $\asymp 1$ migrants each generation, and each migrant is destined to become established with a probability that approaches $1-\rho \approx 0.8$, so that it takes only $O(1)$ generations for migrants who are destined to become established to reach each island neighboring j. Furthermore, with probability approaching 1, each migrant who is destined to become established will in fact do so within a time that is only $O(\lg\lg n) = o(\lg n)$.

Upper bound: For each $\varepsilon > 0$, we have $P\left\{T_n \leq (1+\varepsilon)\left(R+\Delta\right)\lg n\right\} \to 1$ as $n \to \infty$.

Let i^* be an island achieving the minimum in the definition $\Delta = \min_{i \in C(\mathbb{G})} \Delta_i$ and let S^* denote S_{i^*}. By Proposition H, within $(1+\varepsilon)\Delta \lg n$ generations, some individual $I(0, i^*, m)$ on island i^* will become established on each island in the set $\{i^*\} \cup S^*$. Next, by induction, applying Proposition U repeatedly, we see that within an additional $k(1+\varepsilon)\lg n$ generations, individual $I(0, i^*, m)$ will have become a CA of all islands whose distance from $\{i^*\} \cup S^*$ is less than *k*, and will be established on all islands whose minimum distance to $\{i^*\} \cup S^*$ is equal to *k*. In particular, since all islands are within a distance of $R-1$ from $\{i^*\} \cup S^*$, it follows that by generation $(1+\varepsilon)\Delta \lg n + (R-1)(1+\varepsilon)\lg n$, individual $I(0, i^*, m)$ will have become established on all islands in the graph. From here, an additional $(1+\varepsilon)\lg n$ generations suffices to complete the process, making $I(0, i^*, m)$ a CA of all islands.

Lower bound in Theorem 2

Let $\zeta \approx 0.7698$ be as defined in Theorem 2 of [1].

Lower bound: For each $\varepsilon > 0$, $P\{U_n \geq (1-\varepsilon)\left(D+1+\zeta\right)\log_2 n\} \to 1$ as $n \to \infty$.

Let islands i_0 and i_1 be separated by a distance of *D* from each other, and define $t_1 = (1-\varepsilon)\left(D+1+\zeta\right)\log_2 n$. We claim that with probability approaching 1, there are individuals $I(0, i_0, m_0)$ and $I(t_1, i_1, m_1)$ such that $I(0, i_0, m_0)$ is not extinct at time t_1 and $I(0, i_0, m_0)$ is not an ancestor of $I(t_1, i_1, m_1)$. In other words, the claim is that with probability approaching 1, at time t_1, there is an individual $I(0, i_0, m_0)$ on island i_0 who is not extinct but is not yet a CA of island i_1. By [1], we know that there are many (i.e. a number that approaches infinity as $n \to \infty$) individuals living on island i_0 at time 0 who, at time $(1-\varepsilon)\zeta \lg n$, are destined to become CA's but are not yet established and also have no descendants on any island other than i_0. Let $I(0, i_0, m_0)$ be one of these individuals. Now we just need to show that with probability approaching 1, it will take more than $(1-\varepsilon)\left(D+1\right)\lg n$ generations for $I(0, i_0, m_0)$ to become a CA of island i_1. This follows from Proposition L, from the same reasoning as used earlier to establish the lower bound in Theorem 1.

Upper bound in Theorem 2

Proposition U2: The probability that an individual who is established on a given island fails to become a CA of that island within $(1+\varepsilon)\lg n$ additional generations is $o(1/n)$.

This follows from minor modifications of the analogous result in [1]. The same arguments work because this statement concerns the number of descendants of a given individual on just a single island. The only issue to check here is that the result is not changed by those few individuals per generation who may have a child on a different island; it turns out that this makes no important change in the behavior of the process of counts of descendants of a given individual on a single island.

Upper bound: For $\varepsilon > 0$, $P\{U_n \le (1+\varepsilon)(D+1+\zeta)\log_2 n\} \to 1$ as $n \to \infty$.

Let $0 < \delta < \varepsilon$.

Establishment Stage: From [1], we know that within $(1+\delta)\zeta \lg n$ generations, everyone (that is, all individuals $I(0,i,m)$ on all islands at time 0) is out of jeopardy – either extinct or established.

Let us call the individuals $I(0,i,m)$ who become established the "original established individuals." The remaining individuals from time 0 are all extinct at the end of the Establishment Stage.

To complete the proof we show that with probability approaching 1, the original established individuals will all become CA's of the full population within $(1+\delta)(D+1)\lg n + o(\lg n)$ additional generations. We have $D+1$ additional stretches of $(1+\delta)\lg n$ generations to work with.

Growth Stage 0: We wait until all established individuals have become CA's of their own islands.

Proposition U2 implies that with probability approaching 1, Growth Stage 0 requires less than $(1+\delta)\lg n$ generations. After Growth Stage 0 is completed, we begin Migration Stage 1.

Migration Stage 1: We monitor migrants and wait until we have seen, for each edge of the graph and for both directions along that edge, a migrant along that edge in that direction, with that migrant being established on the destination island.

Migration Stage 1 is readily seen to take only $o(\lg n)$ generations. In fact, the time required to collect a full set of migrants who are destined to become established on the destination island (but have not yet become established) has a distribution that is easily upper bounded by a geometric decay, so this time contributes just $O(1)$ to Migration Stage 1. Then, the additional time required for those migrants who are destined to become established actually to become established is just $O(\lg \lg n) = o(\lg n)$.

At the end of Migration Stage 1, each of the original established individuals is a CA of his home island and has become established on each island within a distance of 1 from his home island.

Growth Stage 1: We wait until the established migrants found during Migration Stage 1 all have become CA's of their islands.

Just as with Growth Stage 0, with probability approaching 1, Growth Stage 1 takes less than $(1+\delta)\lg n$ generations. At the end of Growth Stage 1, each of the original established individuals has become CA of each island within a distance of 1 from his home island. We continue to define Migration and Growth Stages in the same way, inductively.

Migration Stage k: After the end of Growth Stage $k-1$, we begin monitoring migrants and wait until we have seen, for each edge of the graph and for both directions along that edge, a migrant along that edge in that direction, with that migrant being established on the destination island.

Growth Stage k: We wait until the established migrants found during Migration Stage k have all become CA's of their islands.

As above, with probability approaching 1, Migration Stage k takes $o(\lg n)$ generations and Growth Stage k takes less than $(1+\delta)\lg n$ generations. At the end of Growth Stage k, each of the original established individuals has become CA of each island within a distance of k from his home island.

Since each island is within a distance of D from every other island, it follows that at the end of Growth Stage D, each of the original established individuals is a CA of the full population. With probability approaching 1, the total time taken for this to occur is less than $(1+\delta)\zeta\lg n$ for the Establishment Stage, plus $(1+\delta)\lg n$ for Growth Stage 0, plus $D\left((1+\delta)\lg n + o(\lg n)\right)$ for Migration Stage 1, Growth Stage 1, …, Migration Stage D, and Growth Stage D. That is, as n approaches infinity, $P\{U_n \le (1+\delta)\zeta\lg n + (1+\delta)\lg n + D\left((1+\delta)\lg n + o(\lg n)\right)\} \to 1$. Since $\delta < \varepsilon$, this implies $P\{U_n \le (1+\varepsilon)(D+1+\zeta)\lg n\} \to 1$.

Reference

[1] Chang, J. T. Recent common ancestors of all present-day individuals. *Adv. Appl. Probab.* **31**, 1002-1026, with invited discussion and author's reply, 1027-1038 (1999)