

Searching for extraterrestrial intelligence in our own genome

John P. Costella

770 5th St NW Apt 1207, Washington, DC 20001-2673, United States

(September 8, 2025)

Abstract

I address the controversial claim of the discovery of intelligent signals in the genetic code by proposing a far simpler test: is the value of π simply encoded anywhere in our genome? I propose a test of this hypothesis that is statistically robust against a moderate number of single-nucleotide polymorphisms and sequencing errors. I apply the proposed method to the reference human genome and to the raw high-depth whole genome sequencing data for two particular humans. I find no evidence for this intelligent signal.

1. Introduction and motivation

The 2013 announcement by shCherbak and Makukov [1] of their discovery of intelligent signals in the genetic code was immediately criticized as numerology by Myers [2], and later in passing by Prothero and Callahan [3] and in slightly more detail by Neukamm [4]. It otherwise did not get much attention until it was featured by Giorgio Tsoukalos in a 2018 episode of *Ancient Aliens* [5], arguably mainstreaming it, spurring Myers to create a video rebuttal [6].

One way to avoid accusations of numerical cherry-picking is to design as simple a scenario as possible up-front, with suitable tests of statistical significance built in, and then to measure it and report the results, regardless of the outcome. I believe that the experiment I propose below, grounded in historical precedent, fulfills these criteria. Note that I wrote almost all of this paper before performing the actual experiment, although a suspicious reader could argue that only my wife Sally and son Jack could really attest to that fact.

2. My proposal, and its historical context

The fundamental hypothesis under contention is that some form of intelligence manipulated our genome in the past, and left us a signal in there that we would only be able to discover once we achieved the technological ability to sequence our own genome. This seems to have been first proposed by George Marx in 1979 [7], extending the 1973 “directed panspermia” hypothesis of Francis Crick and Leslie Orgel [8]. Marx argued that an extraterrestrial intelligence could have left us a message in our genomic nucleotide sequence—particularly the non-coding parts—or alternatively in the “genetic code” (the mapping of the 64 three-nucleotide codons to amino acids and stop signals), but noted that the latter would have a severely limited scope for transmitting much information. shCherbak and Makukov cited Marx’s two options, but—as Myers emphasizes—only analyzed the more limited genetic code option.

Marx, in turn, acknowledged valuable discussions with Arthur C. Clarke, whose 1968 screenplay with Stanley Kubrick, *2001: A Space Odyssey* [9], posits an alien intelligence modifying early humans to make us smarter, and leaving us a signal—but on the Moon, rather than in our own genome. Clarke’s concurrent novel of the same name [10], based on the screenplay, makes it explicit that the imagined experiment on early humans involved genetic modification. In turn, *2001* was itself inspired by several of Clarke’s own short stories from the 1950s [9].

Now, if we take seriously Marx’s first hypothesis—of encoding a message in the nucleotide sequence of our genome—then the extraterrestrial intelligence would have faced the familiar quandary of leaving us a recognizable message having had no prior communication with us at all. In the context of the search for radio signals from extraterrestrial civilizations, Pyotr Makovetsky proposed in 1973 that a recognizable and unmistakable use of the universal and irrational constant π would be the most logical way to indicate the presence of an intelligent signal [11], an idea that was tested experimentally by Blair *et al.* in 1990 [12], and popularized by the 1997 movie *Contact* [13], even though the idea wasn’t actually used or even mentioned at all in the 1985 novel by Carl Sagan [14] on which the movie was based.

I argue that Makovetsky’s logic applies equally to Marx’s hypothesis of a signal encoded in our genome: the simplest and least misinterpretable way to leave us a recognizable message would be to somehow encode π into the nucleotide sequence.

Now, it’s possible that the extraterrestrial intelligence would have considered the fundamental constant to be what we call 2π , or maybe even $\pi/2$, but those are arguably the only two really viable alternatives. So how could they have done it?

It’s possible that they could have thought about encoding it using the codons that represent amino acids and stop signals. The 64 different codons could map to the 64 different six-bit integers, and they could have represented π in binary form, encoding each six-bit chunk of it in turn by the corresponding codon. But there are $64! \approx 10^{89}$ different ways to map codons to 6-bit integers, and it would be impossible for anyone in this Universe to check all of those possibilities for a signal—and the intelligence would know that.

A more feasible scheme would have been to ignore codons and simply map the four nucleotides A, C, G, and T themselves to the four dibits (two-bit integers 0b00, 0b01, 0b10, and 0b11), and encode π two bits at a time. There are only $4! = 24$ different permutations representing the possible mappings, so the analysis is completely viable. A genome can be read in two possible directions, so there are 48 possible ways that the intelligence could have encoded π in this arguably simplest way.

Encoding π in binary brings with it the great advantage that the alternative choices 2π and $\pi/2$ have identical bits, just with the binary point (the equivalent of the decimal point, but for binary numbers) shifted to the right or to the left by one bit. Thus, assuming that the intelligence left out the binary point completely, then π , 2π , and $\pi/2$ all look identical as binary bitstreams.

But that last assumption should make us worry: what if, instead of omitting the binary point, the intelligence decided to *leave out completely* the integer part of π (or 2π or $\pi/2$), namely, 3 (or 6 or 1), and encode only the bits of the *fractional* part (after the binary point)? If we were to look for a sequence starting with the bits 11001... we would miss such a message completely.

A solution to this possible problem is to make our analytical method *robust against wrong bits in general*. My proposal is the following. Consider the first b bits of π to be a “needle” integer, and the bitstream corresponding to the nucleotide sequence of our genome, encoded in one of the 48 possible ways, as a “haystack” of h enqueued bits. We start by dequeuing the first

b bits of the haystack into a “trial harness” integer. We then compute the bitwise XOR of the needle integer and the trial harness integer. The result will be a 0 bit wherever the corresponding needle and trial harness bits are equal, and a 1 bit wherever they are unequal. We now count up the number of “matching bits” contained in this result, *i.e.*, the number of 0 bits, which in general I will call the “match count,” M . The value m_1 that we get for this first XOR trial, M_1 , will be an integer between 0 and b . The discovery of a statistically improbably high value of m_1 would signify the presence of π encoded right at the start of the genomic sequence. Note that this works even if some of the haystack bits are actually wrong, provided that the number of correct bits is still sufficiently high to be statistically improbable.

But we don’t want to just look for π at the start of the haystack: we want to know if it appears *anywhere* in the haystack. Imagine that we now roll off the first bit of the haystack from the trial harness integer, and roll the next bit of the haystack onto its other end. We repeat the XOR trial, and again compute the match count, m_2 , for this second trial, M_2 . We iterate these rolling XOR trials M_i until we have loaded the last bit of the haystack into the trial harness.

To look for π , we could simply inspect each m_i in turn until we find one that we compute (by some suitable statistical test) to be statistically improbable. But we can be more sophisticated than that: why not analyze the *statistical properties* of *all* of the m_i , and look for any statistically significant deviations of those statistics from the null hypothesis that π is no more special than any other random 64-bit needle? That would also allow us to detect matches of m bits that would not, *taken individually*, be statistically improbable, but might—if they occurred frequently enough throughout the genome—be collectively statistically improbable. In other words, we would be giving ourselves the greatest possible chance of finding the hypothesized encoding of π .

To pursue this more sophisticated path, imagine that we have initialized a set of $b + 1$ integers F_m to track the frequency of each of the possible values m of the M_i :

$$F_m \equiv \sum_i m_i. \quad (1)$$

If we denote by n the total number of XOR trials, a moment’s thought shows that

$$n \equiv \sum_i \equiv \begin{cases} h - (b - 1) & \text{if } h \geq b, \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

because we don’t perform any trials while loading the first $b - 1$ bits into the trial harness, and of course we don’t perform any trials at all if there are actually less than b bits in the haystack. The total frequency will of course be n :

$$\sum_{m=0}^b F_m = n. \quad (3)$$

Now, consider a simple toy scenario in which both the needle and the haystack are made up of bits that are all independent fair coins (henceforth, for brevity, “random bits”). The XOR of each bit of the needle integer with its corresponding bit in the trial harness integer would then be a Bernoulli trial with probability $\frac{1}{2}$. The match count M_i for any XOR trial i would then be binomially distributed: $M_i \sim B(b, \frac{1}{2})$, giving us the probability $p_m \equiv \Pr(M_i = m)$ that any given XOR operation would yield m matching bits.

Because for this toy scenario the haystack consists of random bits, the complete set of n XOR operations is *itself* a Bernoulli trial for each of the $b+1$ values of F_{mn} , by which I am denoting the value of F_m after n trials, and so F_{mn} will itself be binomially distributed with probability p_m , namely, $F_{mn} \sim B(n, p_m)$.

If we were to take this toy scenario as our null hypothesis, we could simply compute appropriate confidence intervals for each of the F_{mn} using the binomial distribution, and compare the empirical f_{mn} to those confidence intervals.

But this toy model is not, of course, sufficient for analyzing a real genomic sequence: we know that there is significant repetition in the human genome, and the sequence of bases is not random. I show in Sec. 4 below that we can adjust our statistical modeling to cater for these issues. A beneficial by-product is that it allows the method to make use of the raw data from high-depth whole genome sequencing that samples each position on the genome *many* times (and hence contains a much higher amount of repetition) by sequencing many “reads” of relatively short fragments of DNA (dozens to hundreds of bases each). If we label the fragments by the index k , then the “read” of the k -th fragment represents a single haystack, as described above, where now h_k is twice the number of bases in fragment k (since each base is encoded by a dibit). Let us denote by K the number of fragment haystacks, which I will collectively refer to as the “hayfield.” The total number of bits in the hayfield, H , is just

$$H \equiv \sum_{k=1}^K h_k. \quad (4)$$

We process each of the K haystacks using the algorithm above. From (2), the total number of XOR trials for the hayfield will be

$$N \equiv \sum_{k=1}^K n_k. \quad (5)$$

We can simply sum all the frequencies F_{mn_k} over all of the K haystacks in the hayfield:

$$F_{mN} \equiv \sum_{k=1}^K F_{mn_k}. \quad (6)$$

The generalization of (3) is then just

$$\sum_{m=0}^b F_{mN} = N. \quad (7)$$

We can then feed these frequencies F_{mN} into the statistical method that I will propose in Sec. 4.

Before that, however, I will first make the analysis more concrete by putting some numbers to it for the particular experimental tests that I performed of the human genome.

3. Experimental setup

One source of suitable raw data for the human genome is clearly the current GRCh38.p14 reference genome [15]. Its FASTA file is conveniently easy to process, but in addition to the

specific bases A, C, G, and T, it also contains IUPAC base codes representing the alleles of single-nucleotide polymorphisms (where two, three, or all four of the bases are possible at that position), as well as large blocks of N (“any”) base codes as placeholders for unknown or uncharacterized regions. It would make little sense to try to encode a message in regions containing the former, and the latter regions of the file contain no information at all. To maintain the fidelity of the data used for the experiment, I decided that I would end the current fragment at any of these ambiguous or unknown base codes, in addition to the 704 places in the file at which a new (labeled) segment is started. I preprocessed the data in this way into a custom file format that was most convenient for the remaining processing steps, containing 1,927 fragments with a total of 3,136,819,154 bases.

I foresaw that having the reference human genome as my sole source of empirical data could validly be criticized, if I didn’t end up finding a signal, precisely because of these ambiguous, unknown, and uncharacterized regions in the data. To head off such potential criticism, I decided that I should also analyze the actual raw genomic sequence fragments of at least one individual human. To that end, in September 2024 my wife Sally and I each purchased 100× Whole Genome Sequencing from Nebula Genomics. We received our data about six weeks later. Sally’s two paired-end FASTQ files each contained 1,560,558,378 reads of raw fragments with each read length being 150 bases, yielding a total of 468,167,513,400 bases sequenced. Each of my files contained 1,424,404,286 reads of 150 bases each, yielding a total of 427,321,285,800 bases sequenced. As the haploid human genome contains about 3.2 billion base pairs, those numbers would imply around 134–146× coverage, *i.e.*, each position on the haploid genome was sampled, on average, 146 times for Sally, and 134 times for me. Nebula advised that their computed average alignment coverage, which corrects for things like duplicate reads, was 139.05 for Sally and 126.54 for me, *i.e.*, about 5% lower in each case.

The Nebula FASTQ files did not contain any ambiguous base codes, but they did occasionally contain unknown N codes, and, even more occasionally, significant blocks of them (dozens to scores of consecutive bases). Again, to maintain the fidelity of the data, I ended the current fragment at any such unknown base code. I then combined the resulting fragments from the two paired-end FASTQ files for each of us into my custom file format, yielding 468,161,682,911 known bases in 3,125,343,986 fragments for Sally, and 427,315,431,908 bases in 2,853,209,738 fragments for me.

For the analytical method proposed in Sec. 2 I chose to use a needle size of $b = 64$ bits. Not only is this most convenient from a programming perspective—most computers and languages these days use 64-bit integers natively—but it is also large enough that the probability that a random needle would appear in a particular dataset is negligible unless that dataset were tens of exabytes in size, which won’t be the case for any data that we will be considering here. (Which is also the reason why we have not had to move to 128-bit microprocessors.)

4. Statistical framework

Let’s now return to the method I proposed in Sec. 2 with these concrete numbers in mind. The probability distribution of any particular XOR trial M_i of a random 64-bit needle with a trial harness of 64 haystack bits is binomial, $M_i \sim B(64, \frac{1}{2})$, so that

$$p_m \equiv \Pr(M_i = m) = \binom{64}{m} \frac{1}{2^{64}}. \quad (8)$$

For convenience, I show the approximate numerical values of this distribution in Table 1. Around 47% of the time M will fall between 30 and 34; around 74% between 28 and 36; and around 97% between 24 and 40. The odds of M being less than 6 or greater than 58 is about one in a trillion.

Let us now consider just the first two XOR trials, M_1 and M_2 , under two extreme scenarios.

For the first extreme scenario I return to the toy scenario of Sec. 2, where the haystack bits are random and independent. Each F_m is then the result of two Bernoulli trials with probability p_m , and so $F_{m2} \sim B(2, p_m)$. I show some of the more interesting numerical values of $\Pr(F_{m2} = f_{m2})$ in Table 2.

For the second extreme scenario, consider the case when the second trial harness integer is *guaranteed to be identical* to the first, such as would happen if the first 65 bits of the haystack were guaranteed to be all 0s or all 1s. We obviously then have the identity $M_2 \equiv M_1$; *i.e.*, not just that M_2 has the same distribution as M_1 , as with the first extreme scenario above, but that M_2 is always numerically and identically *equal* to M_1 . Clearly, we would now have $F_{m2}/2 \sim B(1, p_m)$, which for convenience I am going to write as $F_{m2} \sim 2B(1, p_m)$. I show the equivalent of Table 2 in Table 3. Clearly the distribution is “lumpier” than that of Table 2.

We can compare these two extreme scenarios by looking at the mean and variance of the F_{m2} for each scenario. For concreteness, I show these numerically in Table 4. It’s evident that $E[F_{m2}]$ for any particular m is independent of the scenario, but $\text{Var}(F_{m2})$ is twice as large in the identical scenario than it is in the independent scenario. Of course, this is obvious from the binomial distribution $X \sim B(n, p)$, for which $E[X] = np$ and $\text{Var}(X) = npq$, with the standard definition $q \equiv 1 - p$, so that $E[2B(1, p)] = 2p$ is equal to $E[B(2, p)] = 2p$, but $\text{Var}(2B(1, p)) = 4pq$ is double that of $\text{Var}(B(2, p)) = 2pq$.

A moment’s thought shows that these properties extend to all N of the XOR trials as we run the hayfield through our trial harness against our given random needle: $E[F_{mN}]$ will always be Np_m but $\text{Var}(F_{mN})$ could be anywhere between Np_mq_m and $N^2p_mq_m$ depending on how independent the bits of the hayfield are of each other. For these two homogeneous extreme scenarios, the value of the variance is clearly related to the Shannon information content of the hayfield in some sort of reciprocal way. This relationship is less straightforward for scenarios between these extremes, because the Shannon information will be numerically dominated by the less probable 64-bit tokens, but the variance of our XOR trials will be numerically dominated by the most repetitive ones; in other words, the reciprocal nature of the relationship means that what is additive for one is not additive for the other.

Regardless of where the variance falls between these two extremes, we could *estimate* it by running the entire hayfield through the trial harness Q times, each time with a different random needle R_j ; each R_j yields one set of observations f_{mNj} of the F_{mN} . But it is not immediately obvious what distribution we should then use to make use of this single statistic for a real hayfield containing repetition.

A better and more general approach is to simply *collect the full set* of the f_{mNj} , and use them to form an empirical estimate of the *distribution* of each F_{mN} . Because our datasets are large, Q will not be large (it can take on the order of eight processor-hours to run even a single needle across one encoding of a Nebula dataset), so it is more than feasible to keep and store all of the empirical frequencies. The flip-side concern of this, however, is whether such a relatively small value of Q would give us sufficient empirical resolution of each frequency distribution to be able to determine whether any given frequency we find for our π needle is statistically improbable. But since Q is relatively small, this would only be a problem if any frequency we found for the π

m	$p_m \equiv \Pr(M=m)$
32	0.099
31 or 33	0.096
30 or 34	0.088
29 or 35	0.075
28 or 36	0.061
27 or 37	0.046
26 or 38	0.033
25 or 39	0.022
24 or 40	0.014
23 or 41	8.0×10^{-3}
22 or 42	4.4×10^{-3}
21 or 43	2.2×10^{-3}
20 or 44	1.1×10^{-3}
19 or 45	4.7×10^{-4}
18 or 46	2.0×10^{-4}
17 or 47	7.5×10^{-5}
16 or 48	2.6×10^{-5}
15 or 49	8.6×10^{-6}
14 or 50	2.6×10^{-6}
13 or 51	7.1×10^{-7}
12 or 52	1.8×10^{-7}
11 or 53	4.0×10^{-8}
10 or 54	8.2×10^{-9}
9 or 55	1.5×10^{-9}
8 or 56	2.4×10^{-10}
7 or 57	3.4×10^{-11}
6 or 58	4.1×10^{-12}
5 or 59	4.1×10^{-13}
4 or 60	3.4×10^{-14}
3 or 61	2.3×10^{-15}
2 or 62	1.1×10^{-16}
1 or 63	3.5×10^{-18}
0 or 64	5.4×10^{-20}

Table 1: The approximate numerical values of the binomial distribution of M for a single XOR of a random 64-bit needle with any arbitrary trial harness of 64 haystack bits.

$m \backslash f_{m2}$	0	1	2
32	0.811	0.179	0.010
31 or 33	0.817	0.174	0.009
30 or 34	0.832	0.160	0.008
29 or 35	0.855	0.139	0.006
28 or 36	0.882	0.114	0.004
27 or 37	0.910	0.088	0.002
26 or 38	0.936	0.063	0.001
25 or 39	0.957	0.043	0.000
24 or 40	0.973	0.027	0.000
23 or 41	0.984	0.016	0.000
22 or 42	0.991	0.009	0.000

Table 2: Some of the more interesting numerical values of the probabilities $\Pr(F_{m2} = f_{m2})$ for the frequencies F_{m2} after two XORs of a random 64-bit needle with two independent 64-bit trial harness integers.

$m \backslash f_{m2}$	0	1	2
32	0.901	0.000	0.099
31 or 33	0.904	0.000	0.096
30 or 34	0.912	0.000	0.088
29 or 35	0.925	0.000	0.075
28 or 36	0.939	0.000	0.061
27 or 37	0.954	0.000	0.046
26 or 38	0.967	0.000	0.033
25 or 39	0.978	0.000	0.022
24 or 40	0.986	0.000	0.014
23 or 41	0.992	0.000	0.008
22 or 42	0.996	0.000	0.004

Table 3: The equivalent of Table 2 for $\Pr(F_{m2} = f_{m2})$ for the case of the random needle being XORed with the same 64-bit trial harness integer twice.

m	$E[F_{m2}]$		$\text{Var}(F_{m2})$	
	Independent	Identical	Independent	Identical
32	0.199	0.199	0.179	0.358
31 or 33	0.193	0.193	0.174	0.348
30 or 34	0.176	0.176	0.160	0.320
29 or 35	0.151	0.151	0.139	0.278
28 or 36	0.121	0.121	0.114	0.228
27 or 37	0.092	0.092	0.088	0.175
26 or 38	0.065	0.065	0.063	0.126
25 or 39	0.043	0.043	0.043	0.085
24 or 40	0.027	0.027	0.027	0.054
23 or 41	0.016	0.016	0.016	0.032
22 or 42	0.009	0.009	0.009	0.017

Table 4: Numerical values of the mean and variance of the F_{m2} shown in Tables 2 and 3.

needle happened to be smaller than or larger than *any* corresponding frequency we found for the set of random needles. In such a case we could inspect the numerical results and form a rough estimate of the statistical significance of such an extreme value based on the observed distribution for random needles. I show how this works in Sec. 5 below by “cheating” and choosing a needle from within the genome file itself.

Now, I noted above that there are 48 possible ways that the genome could be encoded for a binary message. Half of these come from reading the genome in the reverse direction. For estimating the distributions of the F_{mN} , the bit-reverse of any given random needle is also a possible random needle, so we don’t need to explicitly cater for both. For every one of the remaining 24 mappings of nucleotides to dibits there is another mapping that has all of the bits in the dibits inverted (*i.e.*, $0 \leftrightarrow 1$). Again, when estimating distributions with random needles, the bitwise-inverse of any given random needle is also a possible random needle, so we don’t need to consider both. Thus we are left with 12 mappings to analyze. For definiteness, we can select these 12 mappings as those for which the base A maps only to 0b00 and 0b01, whereas the other three bases can map to any of the four dibits.

Intuitively, we might expect the distribution of each F_{mN} to be the same for each of the 12 possible mappings, because we are “throwing” random needles at it (which will henceforth be my shorthand terminology for the XOR trial harness process described above.) But I don’t believe that this is guaranteed, because of the way that we are rolling the bits through the trial harness: there are potential correlations between adjacent bits that could, in bulk, yield different distributions, depending on how the different bases are mapped to dibits. To be conservative, my program computes and stores the frequencies for each of the 12 mappings separately.

When searching a genome for the particular needle π , we must of course consider the genome and its reverse to be two separate mappings. Rather than actually reversing the genome, we can search for π in the reversed genome by simply reversing the 64 bits of our π needle into a “reverse needle,” and search for both needles. But we *can* achieve efficiencies for the 12 pairs of bit-inverted mappings: inverting all the bits of the trial harness for its XOR with a particular needle simply implies $m \rightarrow 64 - m$ for the number of matching bits. Thus we only need to consider the 12 mappings, provided that we look for both unexpectedly high *and* unexpectedly low values

of m in the results; the latter would indicate that π was encoded using one of the 12 mappings that we didn’t actually test.

5. Experimental results

My codebase for this project [16] contains a number of programs; I will only describe some of them here. One, `throw_needles_at_genome`, analyzes one of my custom-format genome files by throwing random needles at it as described in Sec. 4, collecting and saving the empirical frequencies f_{mNjv} for each random needle j for each of the twelve mappings v to a data file. The program can be run any number of times; each time it processes the corresponding genome file it adds the new data to the corresponding data file, so that an arbitrary amount of empirical statistical data can be obtained by rerunning the program as many times as one wants.

Although not required for the experimental process, the program `genome_stats` reads one of these data files and outputs some summary statistics for each value of m and v to a CSV file. The mean and variance can be compared to those obtained in Sec. 4 (supplied in the program output), as a sanity check, but these quantities are not actually used in the final analysis.

For playing with the data, the program `genome_file_head` creates a genome file containing just the first n bases of a given genome file, analogous to the standard Unix utility `head`. These smaller “head” subsets (as I’ll refer to them) can be used to explore how the programs work without the longer run times required for the full genome files. For example, I just used it to create million-base heads for each of the three genome files, which I called `ref-1m.genome.gz`, `sally-1m.genome.gz`, and `john-1m.genome.gz`.

I then ran `throw_needles_at_genome` against each of these million-base heads to collect statistics about their distributions for random needles. These files are small enough that it’s easy to throw thousands of random needles against each of them in minutes.

I then hit each million-base head with *yet another* random needle. Creating a random 64-bit integer is not difficult, but for convenience I include a program `random_needle` that does exactly that. I ran it just now, and it gave me the random needle `33a2887f1bc5706d`. I fed this needle into the program `find_needle_in_genome`, which searches for all 24 variations of a given needle (the 12 dibit mappings of the haystack, for both the needle and its reverse) in each of the million-base heads. (If you are trying these analyses yourself, you will be able to replicate the results for the reference genome, but not for Sally’s or mine, which I probably should not publish—but you *can* use the programs on *your own* genome, if you wish.)

Once that was done, I ran the program `needle_significance`, which joins these results of `find_needle_in_genome` for a given needle and genome file with those of `analyze_genome` for a given genome file, computes the quantile of each observed frequency, and writes out the results to both a CSV file and (for my own convenience in writing this paper) a $\text{\LaTeX}$ file. The results of the latter are shown in Tables 5, 6, and 7. The header of each column shows the permutation of the mapping of bases to dibits relative to the default alphabetical mapping; for example, “1320” stands for $0 \rightarrow 1$, $1 \rightarrow 3$, $2 \rightarrow 2$, and $3 \rightarrow 0$, so that $A \rightarrow 0b01$, $C \rightarrow 0b11$, $G \rightarrow 0b10$, and $T \rightarrow 0b00$. Note that for small and large m , all of the frequencies are zero, and so all the quantiles “freeze out” at 0.5, *i.e.*, the median. Away from these extremes, the quantiles are all somewhere between 0 and 1, which tells us that they are within the range of frequencies observed when I threw random needles against the corresponding genome file using `analyze_genome`. This is as we would expect: it would be unlikely (but not impossible) to obtain a frequency less than

<i>m</i>	forward needle 33a2887f1bc5706d												reverse needle b60ea3d8fe1145cc											
	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320
0	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
1	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
2	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
3	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
4	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
5	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
6	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
7	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
8	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
9	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499
10	0.495	0.496	0.496	0.496	0.495	0.497	0.495	0.496	0.496	0.496	0.496	0.494	0.497	0.495	0.496	0.496	0.495	0.497	0.495	0.496	0.496	0.496	0.494	0.497
11	0.480	0.479	0.481	0.484	0.483	0.480	0.484	0.481	0.480	0.481	0.480	0.478	0.480	0.479	0.481	0.484	0.483	0.480	0.484	0.481	0.480	0.481	0.480	0.478
12	0.422	0.415	0.418	0.418	0.422	0.886	0.418	0.417	0.416	0.426	0.415	0.421	0.422	0.415	0.418	0.418	0.422	0.949	0.418	0.417	0.416	0.426	0.415	0.421
13	0.615	0.265	0.869	0.778	0.259	0.267	0.251	0.250	0.266	0.611	0.255	0.250	0.257	0.265	0.253	0.255	0.259	0.267	0.762	0.250	0.266	0.955	0.255	0.250
14	0.168	0.318	0.174	0.448	0.460	0.060	0.183	0.284	0.191	0.057	0.294	0.309	0.424	0.056	0.430	0.897	0.326	0.060	0.942	0.284	0.191	0.184	0.166	0.309
15	0.504	0.191	0.074	0.632	0.363	0.004	0.342	0.021	0.246	0.184	0.201	0.176	0.326	0.531	0.569	0.424	0.244	0.492	0.775	0.701	0.536	0.184	0.501	0.528
16	0.551	0.019	0.015	0.314	0.704	0.107	0.096	0.227	0.053	0.249	0.359	0.021	0.735	0.084	0.415	0.338	0.430	0.000	0.813	0.385	0.724	0.094	0.748	0.027
17	0.486	0.169	0.202	0.411	0.323	0.175	0.625	0.033	0.018	0.404	0.163	0.172	0.385	0.014	0.258	0.443	0.117	0.070	0.301	0.298	0.058	0.082	0.118	0.064
18	0.337	0.054	0.070	0.473	0.112	0.124	0.462	0.095	0.094	0.159	0.178	0.001	0.487	0.008	0.123	0.391	0.447	0.082	0.692	0.189	0.385	0.268	0.178	0.015
19	0.470	0.007	0.054	0.730	0.413	0.069	0.516	0.363	0.125	0.232	0.329	0.003	0.430	0.010	0.176	0.524	0.262	0.094	0.581	0.140	0.201	0.279	0.157	0.018
20	0.526	0.042	0.171	0.547	0.274	0.111	0.490	0.328	0.102	0.377	0.350	0.085	0.602	0.049	0.390	0.529	0.290	0.091	0.278	0.286	0.369	0.270	0.360	0.025
21	0.430	0.066	0.104	0.536	0.366	0.225	0.390	0.235	0.216	0.227	0.387	0.119	0.514	0.173	0.286	0.605	0.382	0.230	0.398	0.218	0.090	0.234	0.083	0.022
22	0.342	0.162	0.160	0.457	0.436	0.160	0.517	0.172	0.205	0.205	0.450	0.034	0.474	0.191	0.351	0.510	0.482	0.213	0.322	0.255	0.197	0.365	0.347	0.018
23	0.402	0.197	0.244	0.553	0.495	0.172	0.489	0.185	0.116	0.249	0.292	0.083	0.515	0.167	0.386	0.545	0.361	0.342	0.425	0.369	0.127	0.264	0.247	0.077
24	0.325	0.257	0.478	0.427	0.500	0.358	0.438	0.241	0.130	0.274	0.287	0.076	0.372	0.128	0.260	0.519	0.394	0.288	0.507	0.220	0.117	0.359	0.198	0.050
25	0.483	0.238	0.510	0.510	0.503	0.413	0.490	0.327	0.147	0.395	0.416	0.183	0.416	0.269	0.463	0.496	0.492	0.359	0.519	0.306	0.199	0.379	0.321	0.214
26	0.475	0.368	0.510	0.579	0.465	0.430	0.450	0.328	0.219	0.377	0.366	0.153	0.500	0.396	0.507	0.501	0.478	0.400	0.430	0.549	0.085	0.349	0.475	0.258
27	0.540	0.481	0.521	0.524	0.444	0.456	0.414	0.478	0.251	0.427	0.293	0.362	0.497	0.456	0.498	0.494	0.461	0.504	0.496	0.114	0.286	0.464	0.607	0.414
28	0.421	0.480	0.573	0.495	0.448	0.488	0.423	0.566	0.430	0.314	0.411	0.488	0.399	0.488	0.440	0.489	0.541	0.500	0.504	0.127	0.455	0.379	0.514	0.489
29	0.390	0.558	0.608	0.446	0.535	0.541	0.443	0.784	0.535	0.539	0.397	0.460	0.429	0.581	0.476	0.539	0.520	0.534	0.344	0.520	0.585	0.473	0.704	0.534
30	0.538	0.780	0.659	0.566	0.514	0.582	0.590	0.582	0.639	0.625	0.549	0.776	0.419	0.828	0.721	0.526	0.595	0.625	0.549	0.573	0.626	0.550	0.595	0.730
31	0.689	0.854	0.699	0.466	0.596	0.856	0.450	0.643	0.790	0.696	0.410	0.774	0.554	0.818	0.716	0.533	0.706	0.805	0.430	0.726	0.762	0.592	0.555	0.770
32	0.961	0.995	0.964	0.944	0.963	0.987	0.931	0.941	0.984	0.968	0.940	0.993	0.952	0.997	0.971	0.951	0.948	0.985	0.930	0.956	0.992	0.959	0.932	0.994
33	0.686	0.856	0.606	0.568	0.759	0.867	0.483	0.561	0.778	0.664	0.666	0.776	0.762	0.841	0.662	0.551	0.697	0.807	0.504	0.640	0.750	0.694	0.583	0.759
34	0.656	0.716	0.501	0.539	0.572	0.602	0.589	0.593	0.759	0.579	0.698	0.768	0.680	0.673	0.561	0.494	0.484	0.677	0.641	0.541	0.656	0.553	0.622	0.783
35	0.542	0.557	0.418	0.444	0.563	0.490	0.415	0.275	0.545	0.562	0.767	0.552	0.598	0.532	0.431	0.404	0.528	0.538	0.386	0.552	0.576	0.650	0.553	0.449
36	0.467	0.463	0.424	0.466	0.484	0.430	0.459	0.561	0.350	0.380	0.678	0.426	0.467	0.479	0.349	0.522	0.490	0.372	0.390	0.150	0.501	0.447	0.293	0.438
37	0.477	0.333	0.343	0.502	0.466	0.431	0.437	0.247	0.330	0.356	0.477	0.300	0.528	0.411	0.358	0.465	0.475	0.473	0.508	0.471	0.437	0.473	0.233	0.302
38	0.343	0.383	0.343	0.518	0.474	0.404	0.441	0.221	0.294	0.322	0.277	0.207	0.536	0.356	0.397	0.460	0.463	0.398	0.453	0.162	0.294	0.375	0.379	0.254
39	0.406	0.319	0.396	0.531	0.415	0.348	0.504	0.432	0.251	0.342	0.248	0.089	0.451	0.312	0.505	0.519	0.467	0.348	0.525	0.116	0.248	0.305	0.205	0.110
40	0.376	0.213	0.434	0.475	0.384	0.348	0.475	0.328	0.115	0.280	0.217	0.099	0.396	0.148	0.454	0.573	0.393	0.315	0.486	0.392	0.080	0.241	0.504	0.092
41	0.425	0.211	0.395	0.512	0.379	0.285	0.565	0.509	0.134	0.277	0.201	0.043	0.308	0.246	0.436	0.504	0.442	0.347	0.407	0.327	0.031	0.115	0.362	0.105
42	0.353	0.150	0.495	0.498	0.373	0.233	0.557	0.152	0.058	0.282	0.336	0.090	0.326	0.147	0.522	0.538	0.382	0.219	0.458	0.299	0.024	0.178	0.428	0.071
43	0.403	0.141	0.699	0.517	0.449	0.289	0.503	0.243	0.067	0.201	0.277	0.042	0.208	0.069	0.310	0.526	0.399	0.115	0.552	0.428	0.086	0.158	0.166	0.052
44	0.443	0.111	0.534	0.484	0.292	0.304	0.206	0.197	0.057	0.259	0.224	0.077	0.120	0.210	0.487	0.459	0.455	0.251	0.434	0.469	0.044	0.160	0.378	0.036
45	0.646	0.098	0.656	0.398	0.414	0.260	0.309	0.288	0.031	0.248	0.329	0.129	0.085	0.184	0.262	0.601	0.417	0.360	0.416	0.258	0.100	0.315	0.223	0.083
46	0.295	0.076	0.648	0.606	0.320	0.144	0.454	0.618	0.186	0.264	0.070	0.182	0.172	0.090	0.315	0.464	0.148	0.216	0.509	0.640	0.141	0.235	0.262	0.073
47	0.144	0.2																						

m	forward needle 33a2887f1bc5706d												reverse needle b60ea3d8fe1145cc											
	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320
0	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
1	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
2	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
3	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
4	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
5	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
6	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
7	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
8	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
9	0.500	0.499	0.499	0.498	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.500	0.499	0.499	0.498	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499
10	0.494	0.493	0.494	0.493	0.493	0.494	0.492	0.494	0.494	0.494	0.493	0.493	0.494	0.493	0.494	0.493	0.493	0.494	0.492	0.494	0.494	0.494	0.493	0.493
11	0.472	0.471	0.475	0.470	0.472	0.467	0.471	0.475	0.470	0.475	0.472	0.470	0.472	0.471	0.475	0.470	0.472	0.467	0.471	0.475	0.470	0.475	0.472	0.470
12	0.394	0.393	0.387	0.392	0.392	0.391	0.866	0.867	0.394	0.393	0.389	0.391	0.394	0.393	0.387	0.392	0.870	0.867	0.389	0.389	0.394	0.972	0.389	0.391
13	0.801	0.580	0.555	0.210	0.211	0.210	0.209	0.196	0.213	0.802	0.198	0.571	0.197	0.212	0.555	0.804	0.576	0.210	0.909	0.965	0.580	0.581	0.555	0.204
14	0.250	0.131	0.110	0.809	0.032	0.301	0.713	0.244	0.604	0.851	0.422	0.297	0.024	0.131	0.417	0.302	0.306	0.137	0.852	0.244	0.714	0.481	0.023	0.138
15	0.202	0.041	0.263	0.335	0.611	0.051	0.762	0.195	0.419	0.555	0.409	0.008	0.409	0.257	0.670	0.257	0.429	0.142	0.403	0.415	0.350	0.610	0.678	0.128
16	0.255	0.079	0.171	0.536	0.110	0.103	0.672	0.007	0.536	0.245	0.240	0.185	0.226	0.279	0.070	0.561	0.389	0.324	0.909	0.473	0.510	0.416	0.091	0.071
17	0.538	0.055	0.286	0.567	0.279	0.067	0.443	0.237	0.300	0.107	0.256	0.308	0.512	0.188	0.156	0.577	0.157	0.232	0.661	0.427	0.300	0.537	0.094	0.149
18	0.524	0.100	0.148	0.582	0.414	0.219	0.581	0.122	0.104	0.306	0.411	0.233	0.576	0.097	0.345	0.558	0.142	0.405	0.722	0.299	0.488	0.390	0.168	0.121
19	0.507	0.151	0.313	0.579	0.465	0.277	0.532	0.144	0.328	0.429	0.200	0.156	0.472	0.121	0.266	0.579	0.462	0.373	0.509	0.266	0.388	0.440	0.142	0.205
20	0.534	0.180	0.332	0.635	0.344	0.447	0.576	0.320	0.383	0.404	0.352	0.238	0.467	0.257	0.369	0.575	0.458	0.298	0.604	0.263	0.414	0.438	0.172	0.310
21	0.536	0.265	0.307	0.581	0.433	0.335	0.617	0.274	0.422	0.375	0.350	0.330	0.492	0.206	0.539	0.609	0.450	0.371	0.571	0.402	0.366	0.528	0.260	0.286
22	0.531	0.257	0.325	0.547	0.484	0.355	0.601	0.258	0.400	0.373	0.334	0.365	0.465	0.311	0.417	0.543	0.475	0.435	0.531	0.287	0.397	0.486	0.311	0.300
23	0.595	0.305	0.454	0.538	0.474	0.351	0.536	0.277	0.431	0.447	0.372	0.327	0.460	0.337	0.367	0.565	0.405	0.417	0.532	0.306	0.439	0.481	0.317	0.344
24	0.512	0.405	0.502	0.527	0.471	0.418	0.538	0.360	0.386	0.484	0.348	0.381	0.486	0.381	0.327	0.544	0.434	0.456	0.506	0.448	0.421	0.462	0.339	0.376
25	0.451	0.415	0.442	0.528	0.475	0.429	0.544	0.426	0.442	0.487	0.418	0.417	0.457	0.408	0.298	0.501	0.492	0.449	0.510	0.331	0.443	0.467	0.437	0.408
26	0.308	0.456	0.488	0.527	0.490	0.475	0.473	0.438	0.500	0.468	0.421		0.456	0.412	0.540	0.515	0.474	0.446	0.517	0.379	0.225	0.477	0.474	0.454
27	0.090	0.472	0.420	0.530	0.488	0.510	0.498	0.481	0.458	0.520	0.396	0.476	0.259	0.455	0.683	0.505	0.500	0.455	0.527	0.530	0.475	0.489	0.428	0.485
28	0.065	0.478	0.627	0.502	0.508	0.534	0.444	0.903	0.468	0.506	0.075	0.471	0.716	0.525	0.687	0.470	0.500	0.473	0.476	0.810	0.451	0.477	0.677	0.462
29	0.386	0.558	0.624	0.479	0.510	0.557	0.451	0.725	0.491	0.524	0.489	0.585	0.483	0.552	0.561	0.491	0.495	0.509	0.465	0.393	0.507	0.492	0.620	0.536
30	0.655	0.652	0.629	0.489	0.542	0.595	0.478	0.613	0.574	0.556	0.697	0.643	0.464	0.646	0.713	0.522	0.544	0.588	0.498	0.568	0.548	0.583	0.655	0.622
31	0.519	0.865	0.646	0.518	0.605	0.769	0.464	0.630	0.717	0.660	0.564	0.848	0.429	0.903	0.650	0.513	0.722	0.725	0.485	0.625	0.672	0.618	0.630	0.877
32	0.629	0.961	0.593	0.570	0.739	0.886	0.587	0.627	0.848	0.861	0.729	0.980	0.528	0.937	0.560	0.608	0.778	0.930	0.551	0.708	0.905	0.785	0.640	0.974
33	0.613	0.890	0.514	0.543	0.663	0.750	0.558	0.520	0.748	0.658	0.634	0.816	0.537	0.852	0.519	0.543	0.738	0.761	0.500	0.664	0.787	0.622	0.634	0.867
34	0.723	0.627	0.488	0.478	0.561	0.566	0.520	0.613	0.593	0.561	0.643	0.613	0.652	0.642	0.449	0.478	0.609	0.633	0.473	0.594	0.589	0.580	0.608	0.639
35	0.564	0.543	0.288	0.480	0.562	0.494	0.508	0.588	0.550	0.494	0.820	0.537	0.787	0.562	0.272	0.471	0.502	0.548	0.481	0.670	0.523	0.538	0.573	0.550
36	0.440	0.462	0.131	0.495	0.534	0.475	0.511	0.727	0.522	0.453	0.694	0.474	0.921	0.502	0.554	0.474	0.471	0.458	0.498	0.674	0.525	0.525	0.535	0.457
37	0.553	0.440	0.394	0.493	0.488	0.448	0.511	0.297	0.504	0.448	0.460	0.434	0.567	0.435	0.334	0.505	0.467	0.457	0.516	0.326	0.504	0.516	0.378	0.432
38	0.483	0.403	0.554	0.501	0.458	0.426	0.501	0.337	0.482	0.448	0.370	0.384	0.490	0.403	0.372	0.488	0.431	0.425	0.514	0.337	0.492	0.456	0.443	0.402
39	0.512	0.357	0.468	0.504	0.456	0.410	0.523	0.415	0.460	0.443	0.293	0.380	0.551	0.364	0.460	0.536	0.447	0.435	0.539	0.366	0.453	0.462	0.308	0.358
40	0.372	0.318	0.449	0.489	0.407	0.390	0.559	0.358	0.423	0.427	0.318	0.358	0.359	0.308	0.457	0.500	0.438	0.382	0.534	0.360	0.420	0.413	0.355	0.341
41	0.449	0.332	0.457	0.562	0.429	0.422	0.542	0.438	0.408	0.470	0.344	0.332	0.377	0.307	0.500	0.548	0.446	0.413	0.602	0.367	0.373	0.412	0.366	0.367
42	0.353	0.312	0.528	0.553	0.331	0.475	0.590	0.355	0.431	0.489	0.283	0.353	0.314	0.308	0.579	0.582	0.444	0.389	0.568	0.373	0.380	0.384	0.378	0.349
43	0.435	0.293	0.594	0.529	0.388	0.380	0.647	0.328	0.368	0.499	0.335	0.323	0.415	0.319	0.608	0.551	0.457	0.362	0.594	0.317	0.370	0.362	0.428	0.307
44	0.381	0.279	0.585	0.516	0.346	0.408	0.620	0.375	0.348	0.381	0.209	0.252	0.307	0.354	0.513	0.589	0.498	0.312	0.623	0.304	0.377	0.363	0.341	0.336
45	0.279	0.210	0.465	0.501	0.321	0.493	0.623	0.356	0.240	0.435	0.279	0.174	0.272	0.251	0.580	0.539	0.373	0.433	0.603	0.208	0.322	0.323	0.175	0.273
46	0.149	0.210	0.536	0.666	0.262	0.447	0.427	0.148	0.181	0.430	0.292	0.083	0.234	0.201	0.487	0.482	0.330	0.468	0.561	0.158	0.185	0.407	0.209	0.347
47	0.202	0.162																						

<i>m</i>	forward needle 33a2887f1bc5706d												reverse needle b60ea3d8fe1145cc											
	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320
0	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
1	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
2	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
3	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
4	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
5	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
6	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
7	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
8	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
9	0.498	0.498	0.499	0.499	0.498	0.500	0.498	0.498	0.499	0.499	0.499	0.499	0.498	0.498	0.499	0.499	0.498	0.500	0.498	0.498	0.499	0.499	0.499	0.499
10	0.493	0.492	0.493	0.493	0.493	0.494	0.491	0.495	0.494	0.494	0.494	0.493	0.493	0.492	0.493	0.493	0.494	0.491	0.495	0.494	0.494	0.494	0.494	0.493
11	0.471	0.473	0.468	0.470	0.472	0.472	0.474	0.470	0.472	0.970	0.472	0.471	0.471	0.473	0.468	0.470	0.472	0.472	0.474	0.470	0.472	0.473	0.472	0.471
12	0.864	0.391	0.389	0.393	0.392	0.869	0.396	0.384	0.397	0.873	0.383	0.393	0.388	0.870	0.389	0.393	0.867	0.391	0.396	0.384	0.397	0.972	0.383	0.393
13	0.808	0.206	0.196	0.573	0.579	0.575	0.576	0.200	0.213	0.580	0.197	0.204	0.570	0.206	0.915	0.211	0.217	0.575	0.952	0.200	0.213	0.580	0.559	0.204
14	0.422	0.800	0.108	0.031	0.035	0.468	0.454	0.023	0.034	0.894	0.106	0.455	0.117	0.278	0.255	0.136	0.467	0.136	0.962	0.406	0.719	0.465	0.106	0.136
15	0.613	0.135	0.145	0.476	0.144	0.051	0.691	0.203	0.095	0.280	0.088	0.250	0.342	0.338	0.556	0.536	0.478	0.094	0.528	0.203	0.482	0.091	0.273	0.467
16	0.390	0.103	0.318	0.496	0.295	0.034	0.480	0.111	0.360	0.328	0.601	0.271	0.576	0.250	0.228	0.362	0.268	0.204	0.849	0.053	0.501	0.463	0.032	0.019
17	0.538	0.134	0.326	0.612	0.419	0.155	0.499	0.093	0.415	0.298	0.345	0.273	0.153	0.037	0.245	0.514	0.327	0.203	0.619	0.234	0.190	0.427	0.143	0.200
18	0.683	0.175	0.275	0.587	0.421	0.289	0.561	0.111	0.307	0.333	0.176	0.193	0.455	0.208	0.325	0.646	0.170	0.233	0.606	0.360	0.414	0.371	0.246	0.138
19	0.521	0.082	0.276	0.578	0.273	0.316	0.536	0.097	0.400	0.315	0.259	0.214	0.573	0.194	0.342	0.575	0.392	0.365	0.608	0.384	0.382	0.421	0.192	0.270
20	0.587	0.190	0.242	0.533	0.400	0.362	0.565	0.247	0.326	0.413	0.360	0.196	0.565	0.258	0.363	0.565	0.412	0.375	0.619	0.253	0.386	0.378	0.186	0.215
21	0.546	0.254	0.377	0.547	0.425	0.397	0.600	0.211	0.381	0.390	0.372	0.324	0.659	0.243	0.511	0.562	0.370	0.447	0.546	0.395	0.460	0.442	0.290	0.214
22	0.712	0.330	0.452	0.549	0.469	0.396	0.584	0.345	0.393	0.425	0.402	0.294	0.518	0.281	0.528	0.535	0.439	0.419	0.578	0.304	0.449	0.431	0.258	0.314
23	0.495	0.358	0.468	0.536	0.501	0.434	0.567	0.339	0.421	0.432	0.350	0.369	0.527	0.364	0.515	0.546	0.434	0.437	0.526	0.377	0.404	0.461	0.234	0.346
24	0.587	0.394	0.554	0.541	0.456	0.409	0.536	0.405	0.432	0.450	0.431	0.397	0.397	0.379	0.406	0.521	0.426	0.418	0.527	0.428	0.429	0.449	0.397	0.384
25	0.534	0.425	0.478	0.522	0.457	0.425	0.519	0.414	0.449	0.446	0.343	0.411	0.564	0.411	0.383	0.510	0.468	0.420	0.521	0.376	0.422	0.461	0.476	0.414
26	0.307	0.428	0.541	0.497	0.465	0.470	0.502	0.610	0.454	0.467	0.353	0.429	0.469	0.408	0.553	0.501	0.472	0.447	0.503	0.458	0.460	0.480	0.522	0.418
27	0.295	0.458	0.486	0.512	0.483	0.480	0.481	0.544	0.473	0.491	0.341	0.444	0.461	0.468	0.457	0.505	0.483	0.461	0.523	0.337	0.454	0.476	0.432	0.468
28	0.091	0.498	0.940	0.520	0.498	0.540	0.466	0.495	0.474	0.493	0.447	0.469	0.196	0.501	0.491	0.461	0.468	0.494	0.487	0.626	0.486	0.509	0.767	0.476
29	0.359	0.525	0.742	0.470	0.515	0.546	0.481	0.599	0.522	0.516	0.555	0.548	0.260	0.534	0.630	0.489	0.511	0.503	0.513	0.686	0.524	0.491	0.419	0.546
30	0.571	0.626	0.566	0.508	0.550	0.552	0.499	0.642	0.599	0.580	0.716	0.646	0.510	0.636	0.654	0.498	0.586	0.531	0.492	0.605	0.571	0.533	0.663	0.622
31	0.478	0.838	0.541	0.438	0.607	0.706	0.479	0.546	0.637	0.650	0.583	0.903	0.486	0.844	0.580	0.499	0.628	0.810	0.486	0.592	0.706	0.603	0.690	0.820
32	0.567	0.961	0.579	0.537	0.743	0.885	0.692	0.623	0.860	0.836	0.708	0.878	0.593	0.957	0.596	0.646	0.761	0.916	0.566	0.681	0.868	0.721	0.631	0.961
33	0.535	0.877	0.375	0.550	0.721	0.786	0.562	0.558	0.787	0.746	0.636	0.808	0.609	0.850	0.554	0.503	0.775	0.769	0.497	0.604	0.759	0.697	0.587	0.881
34	0.527	0.666	0.375	0.507	0.587	0.573	0.492	0.578	0.572	0.542	0.617	0.663	0.583	0.665	0.395	0.519	0.598	0.576	0.504	0.624	0.567	0.569	0.614	0.636
35	0.723	0.542	0.491	0.490	0.513	0.516	0.484	0.691	0.536	0.525	0.843	0.518	0.298	0.553	0.340	0.494	0.509	0.497	0.466	0.491	0.516	0.582	0.565	0.494
36	0.822	0.465	0.096	0.482	0.510	0.469	0.507	0.518	0.477	0.508	0.522	0.522	0.619	0.479	0.131	0.465	0.480	0.536	0.488	0.616	0.494	0.536	0.129	0.488
37	0.468	0.441	0.354	0.526	0.508	0.458	0.489	0.433	0.479	0.486	0.273	0.446	0.595	0.433	0.175	0.490	0.495	0.484	0.497	0.431	0.516	0.488	0.421	0.461
38	0.529	0.409	0.492	0.511	0.481	0.445	0.504	0.409	0.442	0.438	0.234	0.425	0.628	0.417	0.429	0.499	0.451	0.437	0.506	0.408	0.481	0.487	0.408	0.399
39	0.471	0.390	0.534	0.524	0.457	0.418	0.521	0.438	0.445	0.421	0.355	0.392	0.498	0.375	0.502	0.537	0.416	0.443	0.486	0.339	0.441	0.449	0.460	0.409
40	0.514	0.366	0.465	0.491	0.435	0.406	0.521	0.401	0.431	0.477	0.372	0.352	0.464	0.376	0.544	0.548	0.448	0.423	0.533	0.366	0.410	0.433	0.427	0.376
41	0.441	0.344	0.601	0.538	0.418	0.445	0.530	0.373	0.420	0.464	0.406	0.350	0.399	0.347	0.527	0.520	0.437	0.420	0.539	0.405	0.390	0.441	0.405	0.348
42	0.434	0.312	0.521	0.551	0.386	0.442	0.634	0.399	0.411	0.428	0.340	0.335	0.477	0.346	0.593	0.577	0.475	0.406	0.573	0.327	0.381	0.389	0.412	0.409
43	0.419	0.307	0.568	0.491	0.369	0.444	0.608	0.227	0.407	0.434	0.405	0.292	0.362	0.361	0.648	0.604	0.535	0.412	0.607	0.368	0.365	0.370	0.313	0.355
44	0.266	0.291	0.729	0.565	0.304	0.385	0.592	0.343	0.354	0.459	0.272	0.339	0.347	0.238	0.679	0.594	0.491	0.386	0.668	0.153	0.252	0.325	0.397	0.403
45	0.360	0.201	0.662	0.544	0.318	0.508	0.594	0.308	0.210	0.378	0.327	0.266	0.282	0.300	0.597	0.514	0.421	0.374	0.622	0.348	0.284	0.301	0.109	0.359
46	0.139	0.314	0.522	0.656	0.194	0.498	0.620	0.279	0.157	0.421	0.175	0.304	0.482	0.339	0.641	0.479	0.305	0.421	0.573	0.262	0.167	0.334	0.135	0.258
47	0.255	0.1																						

or greater than any that we observed during our analyses of these small genome files, given that we threw thousands of random needles at each of them.

So let’s now look at what happens when we *do* observe such an extreme frequency. The program `cheat_needle` allows us to “cheat” by randomly choosing a 64-bit needle from *within* a specified genome file, from a randomly-chosen starting base in the file, randomly choosing one of the 24 mappings of bases to dibits, randomly starting half-way through the dibit corresponding to that start base, and randomly reversing the needle. I ran it just now on the million-base head of the reference genome, and it gave me the needle `028c9aa092c20003`, using the 0132 mapping in the forward direction. Using the `find_needle_in_genome` program for this needle on that genome file and `needle_significance` on the result yields Table 8. Note the quantiles denoted **0** or **1**: these are *exactly* 0 or 1, and hence represent measured frequencies that are, respectively, less than or greater than any observed during our analysis of the genome file. Note that one of them is for $m = 64$ for the forward needle for the 0132 mapping: these are precisely the choices that `cheat_needle` happened to randomly use when it selected the needle `028c9aa092c20003`.

Whenever `needle_significance` finds any such “extreme” (0 or 1) quantiles, it also populates a CSV file with one row for every such quantile, with both the actual observed frequency as well as some univariate summary statistics of the empirical distribution from the random-needle analysis. For convenience, it also outputs a L^AT_EX file of a table of up to 40 of the most statistically remarkable results, determined simply by how far the observed frequency falls outside the observed range, normalized to that range, with higher values of m being preferred for tie-breaking. Those results are shown in Table 9. The first row is for our perfect match $m = 64$ result. We see that the needle was actually matched *twice* in the haystack: not just from where we plucked it from, but from some other location as well. This is not surprising, because the human genome has significant repetition. In this case, none of our random needles found a similarly perfect match—it would be astounding if any did—and so all of the corresponding frequencies were zero, so that the range is actually zero. In such cases the higher difference of observed frequency is preferred as a tie-breaker, followed by higher values of m . Note that the needle was also matched with $m = 62$ once, $m = 59$ twice, and $m = 58$ once; again, this reflects repetition in the genome.

The entries for $m = 46$ are more curious, but they have a relatively mundane explanation. The reference genome data file starts with 10,000 N (unknown) bases, and it is straightforward to establish that it contains many more sequences of consecutive N bases: of the first 12,500 rows of FASTA data of 80 bases each (*i.e.*, the first million bases that constitute the head file that I created), 2,001 of the rows contain at least 32 consecutive N bases. Although not all of those rows have all 80 bases as N, most of them do, which would yield around 160,000 N bases. Since my default mapping of these unknown N bases is to A, which in the default mapping is mapped to `0b00`, these blocks of N bases create stretches of consecutive 0 bits that add up to around 320,000 bits in total. Now, it so happens that 46 of the 64 bits in our selected needle `028c9aa092c20003` are 0. For any of the 12 mappings that keeps A mapped to `0b00` (*i.e.*, all those starting with 0 in Table 9), these stretches of 0 bits in the haystack will match these 46 0 bits in our needle. Thus we end up with around 320,000 matches of $m = 46$, plus a thousand or two more from other matches within the million-base head. So even this completely manufactured artifact in the genome file, created by my mapping the N “buffer blocks” in the reference genome file to 0 bits in my `GenomeFile` format, can, with enough care, be analyzed and made sense of using the method I have described above.

<i>m</i>	forward needle 028c9aa092c20003													reverse needle c000434905593140																							
	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320	
0	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	
1	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	
2	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	
3	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	
4	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	
5	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	
6	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	
7	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	
8	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	
9	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	
10	0.495	0.496	0.496	0.496	0.495	0.497	0.495	0.496	0.496	0.496	0.496	0.494	0.497	0.495	0.496	0.496	0.998	0.497	0.495	0.496	0.496	0.496	0.496	0.496	0.496	0.496	0.496	0.496	0.496	0.496	0.496	0.496	0.496	0.496	0.496	0.497	
11	0.971	0.996	0.481	0.484	0.483	0.972	0.484	0.481	0.988	0.481	0.480	0.478	0.480	0.479	0.481	0.484	1.000	0.480	0.978	0.481	0.480	0.481	0.480	0.481	0.987	0.480	0.481	0.987	0.480	0.481	0.987	0.480	0.481	0.987	0.480	0.481	
12	0.950	0.415	0.880	0.418	0.995	0.978	0.880	0.417	0.998	0.426	0.415	0.421	0.989	0.997	0.996	0.878	0.990	0.996	1.000	0.947	0.996	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	
13	0.986	0.992	0.998	0.613	0.259	0.873	0.997	0.991	0.998	0.955	0.959	0.996	0.998	0.986	0.869	0.963	0.997	0.997	0.994	0.963	0.996	0.995	0.990	0.984	0.996	0.995	0.990	0.984	0.996	0.995	0.990	0.984	0.996	0.995	0.990	0.984	
14	0.978	0.955	0.969	0.986	0.949	0.995	1.000	0.998	0.998	0.993	0.872	0.977	0.993	0.998	0.977	0.967	0.989	0.990	0.998	0.983	1.000	0.995	0.989	0.955	0.997	0.987	0.998	0.978	0.955	0.969	0.986	0.949	0.995	1.000	0.998	0.998	0.955
15	0.997	0.987	0.998	0.979	0.983	0.983	1	0.999	1	1.000	0.989	0.986	0.991	0.965	0.998	0.978	0.996	0.997	1	0.999	1.000	0.999	0.998	0.974	0.999	0.987	0.999	0.968	0.990	0.987	1	0.999	1	0.999	0.998	0.974	
16	0.999	0.987	0.999	0.968	0.990	0.987	1	0.999	1	0.999	1.000	0.983	0.999	0.995	1.000	0.987	0.990	0.988	1	1	1	0.998	0.995	0.997	0.999	0.987	0.999	0.968	0.990	0.987	1	0.999	1	0.998	0.995	0.997	
17	0.998	0.984	0.999	0.953	0.980	0.992	1	1.000	1	0.996	0.999	0.992	0.999	0.985	0.998	0.982	0.990	0.995	1.000	0.999	1	0.998	1.000	0.994	0.999	0.984	0.999	0.953	0.980	0.992	1	0.999	1	0.998	0.995	0.992	
18	0.997	0.985	0.999	0.968	0.981	0.990	1	1.000	1	0.993	0.998	0.991	0.998	0.985	0.999	0.975	0.982	0.985	1	0.998	1.000	0.998	1.000	0.996	0.999	0.985	0.999	0.968	0.981	0.990	0.985	0.999	0.998	0.995	0.992		
19	0.994	0.973	0.994	0.960	0.981	0.985	0.999	0.997	0.999	0.991	0.998	0.994	0.994	0.965	0.994	0.955	0.963	0.976	0.998	0.996	0.999	0.995	0.999	0.992	0.999	0.984	0.999	0.960	0.981	0.990	0.985	0.999	0.998	0.995	0.992		
20	0.990	0.971	0.992	0.945	0.967	0.976	0.998	0.997	0.998	0.989	0.997	0.990	0.994	0.965	0.994	0.955	0.963	0.976	0.998	0.996	0.999	0.995	0.999	0.992	0.999	0.984	0.999	0.960	0.981	0.990	0.985	0.999	0.998	0.995	0.992		
21	0.984	0.957	0.987	0.914	0.966	0.972	0.994	0.992	0.994	0.980	0.994	0.984	0.992	0.962	0.993	0.921	0.967	0.967	0.994	0.992	0.994	0.982	0.994	0.978	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943		
22	0.981	0.927	0.978	0.901	0.931	0.938	0.992	0.990	0.992	0.977	0.992	0.980	0.983	0.929	0.983	0.877	0.917	0.949	0.992	0.990	0.992	0.975	0.992	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943		
23	0.967	0.840	0.965	0.795	0.897	0.914	0.982	0.981	0.982	0.958	0.981	0.961	0.966	0.883	0.963	0.770	0.909	0.915	0.982	0.980	0.982	0.957	0.981	0.948	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943		
24	0.936	0.736	0.921	0.675	0.788	0.819	1	1	1	0.996	1	0.998	0.922	0.759	0.917	0.672	0.807	0.806	1	1	1	0.996	1	0.998	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943		
25	0.816	0.516	0.824	0.384	0.564	0.641	0.957	0.952	0.957	0.858	0.957	0.866	0.810	0.455	0.842	0.427	0.630	0.670	0.957	0.955	0.957	0.854	0.957	0.861	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943		
26	0.572	0.168	0.526	0.157	0.229	0.248	0.937	0.931	0.937	0.706	0.937	0.667	0.429	0.196	0.460	0.155	0.200	0.220	0.937	0.934	0.937	0.668	0.936	0.753	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943		
27	0.015	0.005	0.008	0.002	0.017	0.014	0.903	0.860	0.904	0.273	0.903	0.147	0.013	0.005	0.002	0.007	0.006	0.008	0.903	0.900	0.904	0.196	0.903	0.178	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943		
28	0	0.000	0	0.000	0.000	0.000	0.861	0.012	0.855	0.001	0.737	0	0	0	0	0.000	0.000	0	0.819	0.023	0.863	0.001	0.249	0.001	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943		
29	0	0	0	0	0	0	0.191	0.001	0.220	0	0.004	0	0	0	0	0	0	0	0.168	0.002	0.148	0.000	0.004	0	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943		
30	0	0	0.000	0	0	0	0.016	0.001	0.009	0	0.001	0	0	0	0	0	0	0	0.015	0.001	0.009	0	0.002	0	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943		
31	0	0	0	0	0	0	0.001	0.001	0.000	0	0.002	0	0	0	0	0	0	0	0.002	0.000	0.000	0	0.000	0.000	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995	0.943		
32	0	0	0	0	0	0	0.001	0.000	0	0	0.001	0.000	0	0	0	0	0	0	0.001	0.000	0	0	0.002	0.000	0.986	0.995	0.943	0.976	0.986	0.995	0.943	0.976	0.986	0.995			

m	direction	map	frequency	minimum	Q1	median	mean	Q3	maximum
64	forward	0132	2	0	0	0	0	0	0
59	forward	0312	2	0	0	0	0	0	0
62	forward	0132	1	0	0	0	0	0	0
58	forward	0312	1	0	0	0	0	0	0
46	forward	0123	321,051	152	272	314	328	364	1254
46	reverse	0123	321,040	152	272	314	328	364	1254
46	forward	0213	321,149	147	271	314	327	365	1312
46	reverse	0213	321,045	147	271	314	327	365	1312
46	forward	0231	321,833	127	254	303	328	369	1656
46	reverse	0231	321,774	127	254	303	328	369	1656
46	forward	0312	322,005	140	251	297	328	366	1760
46	reverse	0312	321,948	140	251	297	328	366	1760
46	forward	0132	321,906	120	253	298	327	363	1750
46	reverse	0132	321,741	120	253	298	327	363	1750
46	reverse	0321	322,053	128	253	294	328	366	1917
46	forward	0321	322,037	128	253	294	328	366	1917
51	reverse	0321	45	0	0	0	1.2	2	25
48	reverse	0312	579	7	29	39	44.5	53	338
50	forward	1230	86	0	2	3	4.34	6	50
48	forward	0312	576	7	29	39	44.5	53	338
49	forward	0312	227	0	8	12	14.6	18	140
49	forward	0231	223	0	8	13	14.7	18	143
49	reverse	0312	216	0	8	12	14.6	18	140
53	reverse	0231	6	0	0	0	0.0577	0	4
47	reverse	0312	1210	36	91	113	126	143	826
48	forward	0132	519	5	30	39	44	52	359
50	forward	0312	100	0	1	3	4.33	6	69
48	reverse	0321	597	9	29	38	44.2	52	417
47	reverse	0321	1162	43	92	112	126	142	822
49	reverse	0321	222	0	8	12	14.4	18	156
52	reverse	0312	17	0	0	0	0.28	0	12
48	forward	0321	584	9	29	38	44.2	52	417
47	forward	0321	1123	43	92	112	126	142	822
48	reverse	1230	435	8	30	39	44.3	52	319
16	reverse	1203	420	7	30	39	44.2	52	309
16	forward	1203	414	7	30	39	44.2	52	309
50	reverse	0312	92	0	1	3	4.33	6	69
48	reverse	0213	341	8	32	41	44.3	53	259
48	forward	0123	356	7	32	41	44.5	53	271
49	reverse	1320	138	0	8	13	14.4	18	108
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 9: The 40 most “remarkable” results of Table 8, as described in the text. [This data is out of date, and needs to be recomputed.]

The remaining entries in Table 9 are more “regular”: they represent frequencies that fall outside the range observed with the thousands of random needles I threw at the genome file, but not as exceptionally so, being at least of the same order of magnitude. If we obtained those sorts of results with the π needle, rather than our cheat needle 028c9aa092c20003 plucked from the file itself, then some statistical work would need to be done to assess whether any of them is statistically significant, given the overall parameters of the experiment. I’ll return to this sort of question shortly.

Now, since our analysis of a needle plucked from the reference genome was somewhat marred by the large buffer blocks of N codes, let’s switch to the million-base heads of the raw sequencing data for Sally and me. These files *do* contain the occasional code for ambiguous bases, but not large blocks of them.

First, I just used `cheat_needle` to choose a random cheat needle from within Sally’s million-base head: it gave me a77eb923aeabc0ad, using mapping 2013 in the forward direction. Running `find_needle_in_genome` and `needle_significance` yielded Tables 10 and 11. These results are about as vanilla as you can get for a cheat needle: there is just a single extreme quantile, a perfect “anti-match” ($m = 0$) for the mapping 1320 in the forward direction. This tells us that we would have had a perfect match for the ones’ complement mapping 2013, which is precisely the mapping that `cheat_needle` used when plucking the needle.

Likewise, I just did the same for the million-base head of my own genome file. The cheat needle db54e97d5645be4f was plucked using the 3012 mapping in the reverse direction. The results are shown in Tables 12 and 13. Again, there is a single perfect anti-match for mapping 0321 in the reverse direction, which means a perfect match for the ones’ complement mapping 3012, again in agreement with `cheat_needle`’s random decisions.

I believe that the above examples establish that my programs at least appear to be working correctly. Let us therefore cut to the chase and see what happens when we look for the bits of π in the three full genome files. In hexadecimal, $\pi = 0x3.243f6a8885a308d3\dots$, which in binary is 0b11.0010010000111111011010101000100010000101101000110000100011010011. If we take out the binary point a rewrite the first 64 bits in hex, we get 0xc90fdaa22168c234. This is needle we are looking for.

The reference human genome is the easiest of the three to analyze, in practical terms, because it is 110–120 times smaller than each of our Nebula high-depth files.

[Not completed. But π is not in there. You’ll just have to trust me on that, or wait for me to finish the codebase.]

6. Discussion

My analysis above establishes that an intelligence did not insert the value of π into the human genome, at least under the simplest scenario that its bits were encoded directly using the four base nucleotides sufficiently frequently throughout the genome that the message would survive millennia of mutations. Let me now discuss the ramifications of this null result.

The simplest explanation is that no intelligence manipulated our genome in the past.

If this were a standard scientific investigation performed a century ago, there would be little more to say. But this is clearly a nonstandard investigation, and “science” has lost all credibility in recent decades through political perversion, so some further discussion is warranted.

One possibility is that an intelligence manipulated our genome, but did not deem it necessary

m	forward needle a77eb923aeabc0ad												reverse needle b503d575c49d7ee5											
	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320
0	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
1	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
2	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
3	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
4	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
5	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
6	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
7	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
8	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
9	0.500	0.499	0.499	0.498	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.500	0.499	0.499	0.498	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499
10	0.494	0.493	0.494	0.493	0.493	0.494	0.492	0.494	0.494	0.494	0.493	0.493	0.494	0.493	0.494	0.493	0.493	0.494	0.492	0.994	0.993	0.494	0.493	0.493
11	0.472	0.968	0.475	0.470	0.472	0.467	0.471	0.475	0.470	0.973	0.472	0.470	0.472	0.968	0.475	0.967	0.472	0.996	0.471	0.475	0.470	0.475	0.472	0.968
12	0.394	0.997	0.866	0.869	0.392	0.391	0.389	0.389	0.394	0.872	0.868	0.867	0.394	0.393	0.866	0.392	0.972	0.391	0.389	0.389	0.394	0.997	0.994	0.867
13	0.557	0.580	0.796	0.804	0.957	0.975	0.800	0.990	0.213	0.906	0.798	0.977	0.197	0.978	0.910	0.907	0.809	0.808	0.209	0.965	0.213	0.986	0.982	0.955
14	0.250	0.929	0.908	0.990	0.962	0.954	0.795	0.997	0.917	0.938	0.960	0.899	0.250	0.965	0.956	0.990	0.470	0.981	0.289	0.958	0.983	0.978	0.996	0.962
15	0.876	0.989	0.613	0.972	0.973	0.951	0.691	0.998	0.867	0.983	0.958	0.983	0.269	0.984	0.823	0.981	0.973	0.963	0.837	0.946	0.921	0.952	0.991	0.956
16	0.848	0.974	0.717	0.982	0.980	0.974	0.634	0.995	0.846	0.943	0.980	0.994	0.730	0.981	0.785	0.982	0.971	0.983	0.828	0.909	0.920	0.980	0.997	0.975
17	0.888	0.979	0.765	0.977	0.977	0.982	0.771	0.996	0.886	0.975	0.984	0.976	0.602	0.979	0.654	0.983	0.982	0.985	0.874	0.983	0.878	0.979	0.996	0.961
18	0.845	0.970	0.810	0.968	0.974	0.976	0.722	0.993	0.836	0.981	0.985	0.982	0.842	0.980	0.797	0.985	0.980	0.983	0.807	0.986	0.900	0.976	0.994	0.972
19	0.873	0.979	0.905	0.977	0.977	0.975	0.750	0.993	0.861	0.975	0.981	0.980	0.819	0.980	0.844	0.981	0.978	0.982	0.778	0.987	0.885	0.974	0.993	0.971
20	0.815	0.975	0.894	0.971	0.973	0.976	0.755	0.993	0.791	0.972	0.985	0.977	0.841	0.980	0.879	0.981	0.977	0.981	0.807	0.991	0.835	0.972	0.988	0.975
21	0.880	0.974	0.870	0.967	0.968	0.972	0.730	0.995	0.783	0.971	0.990	0.974	0.871	0.976	0.863	0.969	0.973	0.978	0.740	0.993	0.807	0.966	0.992	0.972
22	0.881	0.972	0.905	0.966	0.967	0.972	0.680	0.995	0.744	0.969	0.993	0.969	0.881	0.973	0.870	0.969	0.971	0.975	0.685	0.993	0.742	0.964	0.994	0.965
23	0.889	0.962	0.913	0.960	0.962	0.968	0.640	0.992	0.683	0.966	0.993	0.963	0.847	0.969	0.866	0.963	0.966	0.970	0.683	0.992	0.700	0.958	0.991	0.961
24	0.883	0.950	0.916	0.955	0.956	0.962	0.581	0.994	0.619	0.958	0.994	0.959	0.881	0.952	0.878	0.954	0.962	0.960	0.561	0.995	0.614	0.954	0.994	0.953
25	0.889	0.940	0.847	0.940	0.944	0.949	0.448	0.994	0.508	0.937	0.993	0.942	0.814	0.943	0.878	0.937	0.949	0.947	0.461	0.994	0.506	0.946	0.995	0.939
26	0.898	0.912	0.893	0.916	0.933	0.938	0.344	0.991	0.421	0.925	0.994	0.910	0.870	0.917	0.892	0.907	0.928	0.928	0.345	0.997	0.390	0.917	0.988	0.918
27	0.664	0.873	0.644	0.875	0.885	0.900	0.272	0.996	0.309	0.875	0.983	0.875	0.900	0.873	0.787	0.862	0.891	0.870	0.257	0.995	0.289	0.878	0.973	0.881
28	0.280	0.809	0.041	0.790	0.815	0.823	0.191	0.308	0.186	0.797	0.071	0.755	0.527	0.759	0.485	0.748	0.814	0.793	0.177	0.531	0.185	0.814	0.119	0.810
29	0.127	0.646	0.067	0.667	0.674	0.591	0.128	0.005	0.122	0.655	0.005	0.622	0.142	0.642	0.222	0.653	0.636	0.591	0.108	0.013	0.115	0.652	0.003	0.647
30	0.081	0.393	0.079	0.398	0.350	0.295	0.069	0.006	0.057	0.362	0.010	0.354	0.126	0.378	0.088	0.402	0.325	0.324	0.062	0.013	0.049	0.381	0.005	0.356
31	0.099	0.130	0.100	0.138	0.089	0.086	0.040	0.006	0.029	0.114	0.010	0.120	0.128	0.127	0.090	0.122	0.090	0.093	0.030	0.006	0.026	0.120	0.006	0.132
32	0.128	0.026	0.124	0.029	0.021	0.020	0.046	0.007	0.027	0.030	0.008	0.037	0.125	0.025	0.125	0.037	0.024	0.024	0.033	0.007	0.026	0.034	0.006	0.024
33	0.105	0.035	0.126	0.032	0.027	0.022	0.129	0.007	0.099	0.031	0.006	0.034	0.101	0.031	0.128	0.038	0.026	0.024	0.152	0.007	0.097	0.026	0.007	0.038
34	0.147	0.057	0.174	0.062	0.045	0.046	0.375	0.012	0.337	0.060	0.006	0.060	0.140	0.064	0.093	0.069	0.048	0.046	0.410	0.007	0.351	0.051	0.007	0.064
35	0.104	0.105	0.307	0.101	0.090	0.089	0.651	0.007	0.629	0.096	0.009	0.126	0.060	0.118	0.094	0.126	0.104	0.099	0.686	0.006	0.646	0.107	0.016	0.097
36	0.775	0.181	0.718	0.181	0.177	0.183	0.773	0.411	0.789	0.185	0.080	0.190	0.440	0.185	0.753	0.183	0.178	0.201	0.796	0.006	0.810	0.189	0.846	0.182
37	0.946	0.291	0.917	0.263	0.307	0.313	0.873	0.994	0.873	0.285	0.975	0.296	0.845	0.301	0.914	0.268	0.303	0.310	0.875	0.942	0.885	0.276	0.999	0.283
38	0.920	0.385	0.897	0.367	0.430	0.448	0.901	0.993	0.917	0.385	0.991	0.383	0.864	0.369	0.934	0.371	0.410	0.460	0.918	0.981	0.926	0.390	0.998	0.385
39	0.917	0.499	0.893	0.472	0.562	0.546	0.931	0.991	0.942	0.525	0.991	0.472	0.903	0.480	0.874	0.488	0.551	0.574	0.938	0.992	0.948	0.513	0.994	0.482
40	0.887	0.579	0.901	0.575	0.669	0.659	0.950	0.991	0.955	0.597	0.995	0.539	0.920	0.573	0.909	0.579	0.659	0.661	0.955	0.993	0.960	0.595	0.993	0.581
41	0.884	0.649	0.889	0.662	0.737	0.737	0.960	0.990	0.961	0.682	0.995	0.638	0.932	0.604	0.900	0.619	0.715	0.732	0.959	0.994	0.966	0.696	0.993	0.652
42	0.900	0.712	0.926	0.709	0.773	0.785	0.970	0.989	0.973	0.738	0.995	0.680	0.911	0.717	0.909	0.679	0.781	0.765	0.967	0.993	0.972	0.758	0.992	0.696
43	0.880	0.750	0.891	0.757	0.812	0.828	0.975	0.988	0.977	0.751	0.994	0.730	0.931	0.687	0.912	0.744	0.783	0.797	0.962	0.995	0.972	0.783	0.988	0.741
44	0.907	0.799	0.894	0.783	0.841	0.847	0.975	0.991	0.981	0.797	0.995	0.736	0.873	0.777	0.921	0.791	0.817	0.866	0.973	0.994	0.976	0.837	0.987	0.794
45	0.906	0.841	0.880	0.818	0.867	0.882	0.978	0.988	0.980	0.835	0.992	0.739	0.950	0.798	0.895	0.812	0.836	0.870	0.974	0.991	0.975	0.815	0.990	0.774
46	0.925	0.802	0.686	0.733	0.893	0.914	0.982	0.988	0.981	0.803	0.993	0.733	0.855	0.776	0.870	0.719	0.854	0.888	0.969	0.987	0.971	0.861	0.989	0.860
47	0.789	0.838																						

m	forward needle db54e97d5645be4f												reverse needle f27da26abe972adb											
	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320	0123	0132	0213	0231	0312	0321	1023	1032	1203	1230	1302	1320
0	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
1	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
2	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
3	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
4	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
5	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
6	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
7	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
8	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
9	0.498	0.498	0.499	0.499	0.498	0.500	0.498	0.498	0.499	0.499	0.499	0.499	0.498	0.498	0.499	0.498	0.498	0.498	0.498	0.499	0.499	0.499	0.499	0.499
10	0.493	0.492	0.493	0.493	0.493	0.494	0.491	0.495	0.494	0.494	0.494	0.493	0.493	0.992	0.493	0.493	0.493	0.494	0.491	0.495	0.494	0.494	0.494	0.493
11	0.471	0.473	0.468	0.470	0.472	0.472	0.474	0.470	0.472	0.473	0.472	0.471	0.471	0.473	0.468	0.470	0.472	0.472	0.474	0.470	0.472	0.473	0.969	0.471
12	0.388	0.870	0.865	0.393	0.392	0.869	0.396	0.384	0.397	0.395	0.972	0.973	0.388	0.998	0.389	0.393	0.968	0.973	0.875	0.863	0.873	0.873	0.383	0.871
13	0.206	0.802	0.556	0.801	0.579	0.954	0.576	0.561	0.579	0.984	0.961	0.953	0.808	0.905	0.196	0.951	0.955	0.811	0.212	0.989	0.807	0.973	0.804	0.800
14	0.577	0.983	0.579	0.896	0.796	0.915	0.030	0.799	0.145	0.803	0.965	0.954	0.260	0.992	0.579	0.973	0.946	0.977	0.454	0.980	0.482	0.917	0.800	0.989
15	0.795	0.981	0.556	0.873	0.947	0.955	0.465	0.919	0.761	0.956	0.981	0.984	0.414	0.971	0.674	0.968	0.880	0.943	0.589	0.979	0.653	0.977	0.806	0.955
16	0.333	0.983	0.171	0.943	0.965	0.969	0.145	0.809	0.472	0.913	0.983	0.989	0.576	0.972	0.407	0.976	0.946	0.977	0.201	0.956	0.603	0.965	0.946	0.991
17	0.427	0.978	0.515	0.960	0.974	0.962	0.545	0.929	0.415	0.958	0.927	0.991	0.494	0.989	0.610	0.966	0.940	0.971	0.093	0.973	0.655	0.973	0.923	0.980
18	0.549	0.979	0.656	0.971	0.965	0.969	0.143	0.962	0.581	0.958	0.967	0.987	0.464	0.986	0.520	0.959	0.945	0.964	0.281	0.973	0.640	0.981	0.950	0.981
19	0.607	0.978	0.688	0.963	0.969	0.968	0.223	0.962	0.631	0.962	0.964	0.984	0.581	0.982	0.538	0.972	0.961	0.966	0.332	0.979	0.562	0.975	0.951	0.978
20	0.504	0.980	0.723	0.959	0.965	0.970	0.175	0.959	0.567	0.960	0.971	0.983	0.482	0.980	0.636	0.963	0.958	0.968	0.254	0.972	0.520	0.963	0.943	0.980
21	0.582	0.975	0.718	0.953	0.962	0.973	0.216	0.971	0.588	0.964	0.967	0.980	0.645	0.976	0.757	0.960	0.959	0.969	0.261	0.979	0.548	0.964	0.963	0.974
22	0.625	0.974	0.693	0.956	0.958	0.971	0.201	0.965	0.465	0.967	0.973	0.977	0.632	0.976	0.695	0.959	0.965	0.967	0.225	0.982	0.497	0.962	0.967	0.977
23	0.712	0.974	0.613	0.958	0.957	0.967	0.201	0.973	0.377	0.962	0.973	0.976	0.686	0.973	0.650	0.952	0.960	0.967	0.221	0.973	0.394	0.960	0.969	0.976
24	0.718	0.972	0.599	0.954	0.954	0.966	0.191	0.982	0.346	0.958	0.974	0.972	0.714	0.971	0.663	0.957	0.960	0.967	0.187	0.973	0.334	0.957	0.972	0.973
25	0.739	0.967	0.693	0.949	0.958	0.960	0.136	0.978	0.278	0.959	0.966	0.965	0.743	0.965	0.722	0.948	0.958	0.966	0.151	0.965	0.279	0.959	0.967	0.969
26	0.833	0.961	0.781	0.943	0.947	0.958	0.113	0.988	0.204	0.945	0.973	0.955	0.677	0.956	0.612	0.941	0.942	0.962	0.107	0.988	0.206	0.944	0.982	0.962
27	0.791	0.945	0.854	0.931	0.944	0.947	0.101	0.994	0.140	0.932	0.879	0.942	0.763	0.942	0.743	0.926	0.941	0.947	0.102	0.981	0.123	0.945	0.979	0.946
28	0.912	0.909	0.784	0.912	0.922	0.915	0.090	0.963	0.081	0.911	0.131	0.886	0.722	0.885	0.950	0.915	0.916	0.919	0.088	0.602	0.086	0.914	0.514	0.907
29	0.386	0.787	0.506	0.851	0.887	0.840	0.078	0.046	0.055	0.852	0.032	0.770	0.635	0.767	0.595	0.852	0.855	0.832	0.078	0.059	0.057	0.881	0.024	0.762
30	0.395	0.518	0.367	0.721	0.676	0.580	0.064	0.034	0.036	0.625	0.034	0.499	0.442	0.582	0.339	0.742	0.654	0.563	0.064	0.032	0.040	0.702	0.027	0.526
31	0.299	0.198	0.282	0.408	0.296	0.191	0.058	0.024	0.030	0.321	0.021	0.174	0.274	0.176	0.271	0.427	0.285	0.212	0.049	0.026	0.028	0.290	0.020	0.204
32	0.294	0.031	0.390	0.107	0.065	0.044	0.117	0.018	0.038	0.072	0.017	0.031	0.336	0.029	0.342	0.099	0.083	0.051	0.126	0.020	0.030	0.063	0.017	0.030
33	0.229	0.029	0.241	0.069	0.035	0.032	0.369	0.022	0.200	0.044	0.020	0.028	0.302	0.028	0.265	0.060	0.045	0.033	0.424	0.029	0.220	0.038	0.018	0.023
34	0.293	0.040	0.254	0.059	0.041	0.035	0.715	0.026	0.616	0.055	0.034	0.037	0.282	0.039	0.229	0.061	0.042	0.036	0.713	0.036	0.610	0.038	0.027	0.037
35	0.375	0.052	0.321	0.075	0.055	0.055	0.858	0.022	0.834	0.056	0.085	0.070	0.211	0.062	0.352	0.081	0.062	0.052	0.847	0.029	0.844	0.056	0.049	0.052
36	0.251	0.088	0.666	0.087	0.085	0.085	0.913	0.731	0.912	0.084	0.857	0.105	0.362	0.104	0.875	0.082	0.085	0.082	0.912	0.410	0.904	0.092	0.732	0.083
37	0.630	0.123	0.712	0.095	0.123	0.147	0.918	0.912	0.947	0.111	0.993	0.150	0.477	0.161	0.591	0.100	0.105	0.136	0.933	0.941	0.944	0.111	0.981	0.138
38	0.654	0.183	0.583	0.125	0.172	0.186	0.940	0.945	0.959	0.151	0.985	0.235	0.715	0.198	0.710	0.122	0.145	0.175	0.941	0.973	0.958	0.168	0.984	0.178
39	0.601	0.260	0.642	0.130	0.192	0.275	0.949	0.967	0.966	0.213	0.980	0.269	0.570	0.265	0.683	0.126	0.205	0.242	0.947	0.960	0.965	0.204	0.984	0.284
40	0.682	0.323	0.673	0.141	0.265	0.349	0.957	0.975	0.967	0.259	0.966	0.333	0.663	0.326	0.696	0.174	0.273	0.336	0.958	0.966	0.967	0.286	0.980	0.371
41	0.723	0.462	0.664	0.204	0.327	0.438	0.959	0.972	0.968	0.307	0.971	0.376	0.647	0.359	0.692	0.205	0.317	0.475	0.953	0.953	0.971	0.278	0.978	0.450
42	0.672	0.483	0.582	0.224	0.340	0.475	0.957	0.967	0.968	0.408	0.973	0.492	0.678	0.476	0.611	0.179	0.439	0.497	0.960	0.974	0.970	0.376	0.965	0.517
43	0.639	0.547	0.588	0.251	0.389	0.477	0.961	0.973	0.973	0.446	0.963	0.501	0.763	0.489	0.507	0.192	0.488	0.544	0.959	0.965	0.972	0.375	0.983	0.564
44	0.625	0.581	0.528	0.384	0.437	0.503	0.965	0.970	0.971	0.505	0.966	0.554	0.759	0.506	0.447	0.203	0.471	0.639	0.964	0.966	0.971	0.393	0.978	0.555
45	0.550	0.614	0.451	0.326	0.484	0.554	0.968	0.968	0.972	0.535	0.986	0.600	0.710	0.538	0.401	0.208	0.415	0.585	0.961	0.976	0.964	0.473	0.963	0.661
46	0.440	0.713	0.522	0.293	0.450	0.530	0.969	0.966	0.970	0.340	0.980	0.617	0.691	0.563	0.361	0.147	0.410	0.592	0.969	0.957	0.963	0.409	0.963	0.700
47	0.255	0.690																						

to leave us a message. Another is that it left us a message, but it was either more complex, or encoded in a more obscure manner, than the simple scenario postulated above. Neither of these are possible to refute. In the former case, our only chance of detecting intelligent manipulation of our genome might be the detection of properties that could not arise naturally. In the latter, a more adventurous investigator may hypothesize and discover that more complex signal, but they must consider the multiple comparisons trap before claiming significance.

Yet another possibility is that π *is* there in our genome, but that I missed detecting it. There are at least two reasons why this might be the case. The first is that my analysis or code is flawed sufficiently badly that it caused me to miss seeing the message. To mitigate this possibility, I am making available all computer code that I have used, so that others may check it for errors. The second is that the message *is* there in our genome, but omitted from the sequencing data provided by the NIH in the reference human genome and the raw sequencing data provided to Sally and me by Nebula Genomics. It is a sad testament to the state of the world that such a scenario is today even thinkable. The only way I could see of definitively disproving it would be for a sufficiently suspicious and wealthy individual or group to fund the completely independent sequencing of full human genomes, on which data the above method could be used.

Absent such a contrary proof, I think it is reasonable to conclude that it is unlikely that an intelligence sought to communicate its presence to us by leaving us a message in our genome.

References

- [1] V. I. shCherbak and M. A. Makukov, *Icarus* **224** (2013) 228.
- [2] P. Z. Myers, freethoughtblogs.com/pharyngula (2013).
- [3] D. R. Prothero and T. D. Callahan, *UFOs, Chemtrails, and Aliens: What Science Says* (Indiana University Press, Bloomington, 2017).
- [4] M. Neukamm, pandasthumb.org (2021).
- [5] Ancient Aliens, [Series 13, Episode 13](#) (2018).
- [6] P. Z. Myers, youtube.com (2020).
- [7] G. Marx, *Acta Astronautica* **6** (1979) 221.
- [8] F. H. C. Crick and L. E. Orgel, *Icarus* **19** (1973) 341.
- [9] S. Kubrick and A. C. Clarke, *2001: A Space Odyssey* (MGM, 1968).
- [10] A. C. Clarke, *2001: A Space Odyssey* (New American Library, New York, 1968).
- [11] P. V. Makovetskii, *Sov. Astron.* **20** (1976) 123.
- [12] D. G. Blair *et al.*, seti.org (1990).
- [13] R. Zemeckis, *Contact* (Warner Bros., 1997).
- [14] C. Sagan, *Contact: A Novel* (Simon & Schuster, New York, 1985).
- [15] NIH, [Genome assembly GRCh38.p14](#) (2022).
- [16] J. P. Costella, johncostella.com/aliens#code (2025).