**Are synthetic genomes parts of a genetic lineage?**

Gunnar Babcock, Oct. 6th 2019

This is the author’s final draft of a paper forthcoming *The British Journal for the Philosophy of Science*

**Abstract**

Biologists are nearing the creation of the first fully synthetic eukaryotic genome. Does this mean that we still soon be able to create genomes that are parts of an existing genetic lineage? If so, it might be possible to bring back extinct species. But do genomes that are synthetically assembled, no matter how similar they are to native genomes, really belong to the genetic lineage on which they were modeled? This paper will argue that they are situated within the same genetic lineage. To see why requires closely examining whether material overlap between parents and offspring is a necessary feature of biological reproduction. The processes used to create synthetic genomes shows that these processes are a form of scaffolded reproduction because they use external machinery and take ownership of the material parts used to create synthetic genomes. Closely examining these processes also reveals, surprisingly, that ‘synthetic reproduction’ can take place between entities that don’t participate in the same biological lineages.

*1 Introduction*

*2 The argument for lineage-less genomes*

*3**Synthetic eukaryotic chromosomes and material overlap*

*4 Biological reproduction, material and information*

*5 Synthetic reproductive processes and their implications*

*Appendix*

**1. Introduction**

Biologists are nearing the creation of the first multi-chromosomal, eukaryotic synthetic genome. If synthetic genomes copied from *Saccharomyces cerevisiae* became functional in organisms, would those organisms be members of the species *S. cerevisiae*? Researchers don’t seem to be sure what to say as they begin to face this problem. For example, (Zhang, *et al.* [2017]) have altered parts of the genome in a species of yeast to resemble those from a different species of yeast, in what they call ‘species morphing’. Thus, synthetic biologists are beginning to bump up against what philosophers might call the species problem. But what if researchers don’t alter any of the parts of the genome and merely synthetize it? Would organisms that have synthetic genomes share the same genetic lineage as the organisms on which they are modeled? Relying on some of the prominent notions of species and reproduction there may be reasons to think that they aren’t members of the same species, as Piotrowska ([2018]) argues regarding some of the methods used in synthetic biology. This paper examines the relationship between synthetically derived genomes and their native counterparts. It will argue that synthetic genomes should be considered parts of the genetic lineages on which they are modeled. It will also draw the conclusion that reproduction can take place between entities that don’t share a genetic lineage, thereby casting doubt on whether reproduction, by itself, is sufficient for generating cohesion between parts (or organisms) of a species.

The paper begins with an argument that concludes that synthetic genomes cannot be parts of the genetic lineage on which they were modeled. It then considers some of the implications this argument carries. Section two sketches the process used by biologists to create eukaryotic, synthetic chromosomes. This will make it clear why synthetic genomes have a very different causal genesis than the genetic parts from other members of the species *S. cerevisiae.* The third section turns to a debate between Godfrey-Smith ([2009]) and Griesemer ([2000], [2014]) as to whether material overlap is a feature of biological reproduction. The final section presents an account of the processes used to create synthetic genomes as a form of biological reproduction, though it is in many ways vastly unlike all other forms of reproduction with which we’re familiar. This form of synthetic reproduction also reveals that genetic lineages can be quite permissive when they are freed from their material constrains.

**2. The argument for lineage-less genomes**

Genetic lineages are created by the replication of genes. For our purposes, the ‘replication of genes’ is any instance where one strand of DNA replicates itself to generate another, similar strand of DNA. This notion of gene replication (a type of reproduction) is sufficient to create a cohesion property that binds genes together across generations of replication. And, such genetic lineages are spatiotemporally restricted or historical entities. This can be seen by combining the observations that species are spatiotemporal entities, following Hull ([1976], [1978]) and Ghiselin ([1974]) and that the condition binds a species together into a single entity is, minimally, a shared lineage (Ereshefsky [1992]). Biological lineages, like species, are spatiotemporally restricted in this way. This means even if a genome possesses a near perfect resemblance to other genes in a particular lineage, that alone is insufficient for being a part of that lineage. There must be a certain kind of causal link, something like gene replication, between a genome and the lineage of which it’s a part.

Assuming this is the right metaphysical account of genetic lineages, advances being made in synthetic biology lead us to a serious question about the status of synthetic genomes. Of course, the progress made in synthetic biology is philosophically notable for other reasons too. One of these other reasons is that as we get closer to synthesizing eukaryotic genomes we are fast approaching the ability to understand and alter the genes of multi-chromosomal organisms to precisely affect their phenotypical traits. Many philosophers, like Preston ([2008]), consider the ethical implications synthetic microorganisms might have because the reasons for synthesizing genomes are ‘not to preserve properties of the existing bacteria with modified behavior. It is to create an entirely new organism with DNA constructed in its entirety according to human plan’ (Preston [2008], p.33). However, while possessing the ability to tailor the traits of an organism is an ethically intriguing issue, there is another more fundamental question that lurks in this possibility. Are synthetic genomes parts of the genetic lineage from which they were copied?

Piotrowska thinks that they can’t be part of the same spatiotemporal lineage without material overlap during reproduction, because ‘without material overlap, then, synthetic reproduction of DNA breaks the spatiotemporal relation required for biological reproduction and species membership a method that relies on it [synthetic replicas] will not be in a position to resurrect members of *M. primigenius*’ (Piotrowska [2018], p.12). If this is true, it means that an organism that has a completely synthetic genome wouldn’t be a member of the species from which the genome was copied. And while Piotrowska argues that certain, semi-conservative methods of replicating DNA, like Polymerase Chain Reaction (PCR) [[1]](#footnote-1), retain a material connection to a genetic lineage she thinks that other methods that generate DNA from ‘scratch’ fail this test. But is it right to think that a perfectly identical, synthetic genome created from scratch really fails to have enough of a causal connection to the genetic lineage from which it was copied to be a part of that lineage?

To answer this question I look to a debate between Griesemer ([2000a], [2000b], [2014]) and Godfrey-Smith ([2009]) over whether material overlap is a necessary feature of biological reproduction. This debate is key because the methods used to generate synthetic genes seem to lack the type of material transfer between parents and offspring observed in most forms of biological reproduction. Godfrey-Smith considers certain types of virus replication (a form of biological reproduction) to be a paradigmatic example of a reproductive process that is formal rather than material. Griesemer disagrees, arguing there is material overlap in virus replication. The arguments in this debate bears on whether we should consider the scaffolding-like processes used to create synthetic genomes as having the right cohesive properties to situate them within a genetic lineage. I believe a point about ownership of material parts presented by Griesemer yields the result that synthetic genomes should be deemed parts of existing genetic lineages. But, to see why there are doubts as to whether a synthetic genome has the right cohesion to be a part of a genetic lineage, consider the following argument:

1. A fully synthetic eukaryotic genome would be a novel genome, because it lacks material overlap with any parts from an existing genetic lineage*.*
2. To be part of a genetic lineage, a eukaryotic genome must have a causal link that binds it to a spatiotemporally restricted lineage.
3. Material overlap is a necessary feature of biological reproductive, and the reproductive process (replication) of genes creates cohesion in a genetic lineage.
4. Therefore, a fully synthetic genome, even one that is perfectly identical to genomes that are parts of some genetic lineage, could not be part of that lineage*.*

To determine whether this argument (henceforth, argument for lineage-less genomes) is sound, I believe, requires closely examining the first and third premises. These are the conditions under which Piotrowska believes woolly mammoths could not be brought back from extinction. This is because completely synthetic genomes would lack any material connection to the original species (Piotrowska [2018], p.12). To be clear, Piotrowska makes this claim regarding the conservative methods of synthetically generating genomes. She argues that a mammoth genome generated through semi-conservative methods, like PCR, could generate a member of the species *M. primigenius* because these methods would have the right material connection to the original species*.* But what if a genome is created with no material connection, like in the argument above? Before moving on to consider these premises, let’s pause to note the possible implications that would come with accepting the conclusion of the argument for lineage-less genomes.

Piotrowska ([2018]) offers an interesting and insightful treatment of the implications that come with accepting an argument like the one above. She shows that if this argument were true, it would be a mistake to think scientists could synthetically bring back or recreate a species using wholly synthetic genomes. This would mean that deextinction efforts that just use genetic ‘information’ to build synthetic genomes would be a fool’s errand. Scientists might be able to create organisms that appear nearly identical to an extinct species, but they wouldn’t be members of the same species.

To see why, consider a case similar to one offered by Piotrowska: synthetically copying the information in a genome from the extinct species *Mammuthus primigenius* (wooly mammoths) would lack necessary cohesive connection to the species, *M. primigenius*, even if the rest of the necessary biological components (like the cell membrane, cytoplasm) had a sufficient connection to some existing species, say *Loxodonta africana pharaohensis* (Northern African Elephants). Even though an organism created in this way would participate in the species of *L. a. pharaohensis* through a cellular lineage, the genome itself wouldn’t have an obvious lineage. Therefore, the organism would have a hybrid-like tie to *L. a. pharaohensis*, but it couldn’t be part of the species *M. primigenius.* So even though a synthetic genotype might perfectly mimic a genome from the genetic lineage of *M. primigenius*, the organism wouldn’t have a material connection to the genes from the species, *M. primigenius* (we might wonder what kind of a hybrid, if it were a hybrid, such an organism would be). In essence, accepting the argument for lineage-less genomes would entail that synthetic genomes are the beginning points of entirelynew, historically unconnected genetic lineages in spite of their perfect resemblance to already existing genomes. And it would mean we could never bring back an extinct species using a synthetically derived genome.

**3. Synthetic eukaryotic chromosomes and material overlap**

This section provides details about how synthetic genomes are created, particularly the techniques used to assemble parts of synthetic eukaryotic genomes modeled on the species of yeast, *Saccharomyces cerevisiae*. Detailing these processes provides reasons for why the first premise of the argument for lineage-less genomes could be a reasonable premise to hold. This is because it appears we will soon possess the ability to create a genome that does not have share any obvious material parts from parts of the lineage from which it was copied.

We should begin by noting that prokaryotic synthetic genomes have been created and they have been shown to successfully function in unicellular organisms (Gibson, *et al.* [2008]). Thus, it seems as though the outcome of the argument for lineage-less genomes may already have implications which means the question of this paper is not just hypothetical. Yet, because horizontal gene transfer is common among prokaryotic organisms, understanding how to trace the lineages of these organisms is far from clear (consider O’Malley [2014]). Because certain species of microorganisms share genetic material horizontally, rather than vertically, their cellular lineages depart from their genetic lineages making their species designations particularly challenging. For this reason, there are more questions than answers as to how the tree of life should be constructed for prokaryotic organisms (Doolittle and Bapteste [2007]). Some argue that because the traditional, hierarchical tree of life cannot be applied to organisms like bacteria, it’s a mistake to apply species concepts to such organisms (Lawrence and Retchless [2010]). So while the advent of synthetic, prokaryotic genomes provides clear evidence that synthetic biologists can create genomes that are capable of functioning in an organism, there are still serious questions about how the lineages of such organisms are best understood. For these reasons, questions about genetic lineages, species and gene replication in eukaryotic genomes avoid some of the trickier questions that horizontal gene transfers raise. Therefore, the possibility of a synthetic eukaryotic genome makes the question of this paper more acute because the cohesive property that binds together most eukaryotic genomes into a single lineage tends to be found in vertical gene transfers. As it happens synthetic biologists are making rapid advances in creating a synthetic eukaryotic genome.

The *Saccharomyces cerevisiae* 2.0 project (Sc2.0) is aiming to create a fully synthetic, multi-chromosomal, eukaryotic genome. In 2014, researchers completed the synthesis of the third chromosome of *S. cerevisiae* that ‘establishes *S. cerevisiae* as the basis for designer eukaryotic genome biology’ (Annaluru, *et al.* [2014] p.55). Since that time Sc2.0 has made further progress by synthesizing six more complete chromosomes of *S. cerevisiae*. Some of the most recently designed chromosomes, such as ‘synX’ and ‘synV’, were designed using sophisticated techniques like *de novo* gene synthesis (Xie, *et al.* [2017]) and mapping procedures to identify and sort synthetic DNA from ‘wild-type’ DNA (Wu, *et al.* [2017]). Sc2.0 is successfully deploying some of the most advanced techniques used in synthetic biology.

Why are synthetic genomes of particular importance to tracing genetic lineages? Synthesis complicates the causal connection synthetic genomes have to the genetic lineage from which the genomes were copied.This is because the assembly process that researches are using to create functional chromosomes alters the normal process through which genes are created.[[2]](#footnote-2) Researchers begin with 60- to 70 mer overlapping oligonucleotides to create 750 base-pair building blocks. The process is completed using PCR. The use of PCR and other synthetic techniques, like CRISPR, are crucially important because these techniques weaken the material connection between parent-offspring strands of DNA. This is because by remotely coding each piece or ‘minichunk’ of original DNA from the native yeast, each replica minichunk is composed of nucleotides that did not come from some part that can be traced back to the species *S. cerevisiae*. These nucleotides are often supplied to labs by manufacturers. The chemical elements PCR use to build copies of genetic material aren’t, themselves, synthetic. But, the dNTPs (or deoxynucleotide triphosphates) that the machines use do not come from the native genetic material being copied. To verify this, researchers can track the synthetically built bits of genetic material using ‘watermark’ sequences (Gibson, *et al.* [2008]) to differentiate between ‘native’ and ‘synthetic’ genetic material. Thus, the dNTPs used to create synthetic genomes don’t necessarily have any material overlap with any parts of the species *S. cerevisiae*, which makes it look as though the process cannot be considered reproduction.

To be fair, there isn’t a complete loss of ‘native’ material in these synthetic chromosomes. The PCR methods use a parental genome, *in vitro*, as a template and then target small amounts of DNA to be synthetically altered, as Piotrowka ([2018]) notes. This is why she argues that there is material overlap in PCR. It’s also the case that backcrossing with native or wild-type yeast is often used to assure the viability of synthetic genes. But even with backcrossing, recent work performed on synX (Wu, *et al.* [2017]) targets and replaces bits of native genes with synthetically built strands that could bypass the need for backcrossing with native material. And recent, *de novo* work using CRISPR has perfectly matched the physical DNA sequence of the fifth chromosome of *S. cereisiae* to create synV (Xie, *et al.* [2017]). Thus, the researchers seem to be bypassing the semi-conservative, *in vitro* methods Piotrowska considers ([2018] pp.12-3). In essence, it appears as though researchers are getting far closer to breaking nearly all the material connection some chromosomes have to their native counterparts using CRISPR, PCR and other synthetic technologies in unique, ground breaking ways.

The application of such technology means researchers are getting closer to being able to exchange the ‘native’ parts of the original chromosome and replace them with material that has been assembled through these synthetic processes. And, the material in the template genome used in these processes are completely replaced with material from outside a genetic lineage. To date, six chromosomes have been synthesized using such techniques. Quite recently, researchers have even made a functional, synthetic eukaryotic genome, though it is all contained on a single chromosome (Luo, *et al.* [2018]). The result is that the material of eukaryotic synthetic chromosomes has a very different causal link to the genetic lineage of *S. cerevisiae* than native or “natural” chromosomes*.*

But does the creation of synthetic genomes support the idea that there will soon be eukaryotic genomes that will not be parts of *S. cerevisiae*’s genetic lineage because they have no material overlap? After all scientists have not yet created a complete multi-chromosomal synthetic eukaryotic genome that matches a ‘native’ genome. It’s also the case that the synthetic chromosomes that have been created still have genes that can be traced back to *S. cerevisiae.*

With regard to such worries, one might retort that it appears as though scientists are on the cusp of creating synthetically assembled eukaryotic genomes that are functional in organisms. And, they might add, that it has already been shown that synthetic genomes can function in single chromosomal organisms. Moreover, while eukaryotic genomes are far more complex, the remarkably rapid progress made by Sc2.0 over the last several years makes it look like more and more types of ‘natural’ or ‘native’ parts of complex genomes can be assembled through synthetic processes.

From this vantage point, it’s easy to imagine a future where there is an emergence of technologies that allow the creation of genomes that lack a direct material connection to a genetic lineage. After all, Sc2.0 is already using techniques that break such material connections. We need only imagine that the precise targeting methods, similar to those deployed to create synX or synV, are used to meticulously replace each bit of the native material without backcrossing with the native material to yield a functional genome. To push this point even further, in principle, it’s possible that the very chemicals used in the process could come to be synthetic given the advent of chemicals like XNA. Thus, the work being done in Sc2.0 makes of the thought of truly ‘synthetic genomes’, without any material connection to native genomes, not appear beyond the pale.

I think that projects like Sc2.0 make it reasonable to consider the first premise of the argument as being true, even if it hasn’t yet been completed empirically. And if the lack of empirical verification leaves room to doubt the first premise, I don’t actually think that turns out to be a serious problem for the question this paper is asking. Skepticism about the lack of material connection synthetic genomes have with their native counterparts only effects the practical implications of lineage-less genomes. If eukaryotic synthetic genomes that truly lack a material connection to some genetic lineage are never actually created, then certain questions about deextinction will not be as serious. But I don’t think a rejection of the first premise on empirical grounds means there’s nothing to learn about the nature of lineages by considering what a truly lineage-less genomes might look like. Thinking about the possibility of a synthetic genome is an ideal way to test that limits of how thin biological lineages can get before they break.

**4. Biological reproduction, material and information**

What reasons are there to think that material overlap between parent(s) and offspring is necessary in biological reproduction (and replication)? To see why material overlap is important to reproduction requires looking at the debate between formal or informational accounts of reproduction and material ones. According to Dawkins ([1976], [2004]), genes are the reproductive units and biological organisms are merely ‘survival machines’. He conceives of genes as ‘replicators’.[[3]](#footnote-3) Evolutionary progress allows certain genes to replicate or reproduce successfully, while others fail to do so. Organisms are just the imperfect survival machines of genes because they don’t replicate with the high fidelity we observe in genes. On this view, reproduction is formal rather than material. The formal nature of Dawkin’s account is clear in the way he analogizes genes to a written language. The arrangement of letters that makes words are like strings of nucleotides, marked by cistrons that code protein chains, etc. This means the particular sequence of nucleotides is the gene, not the particular molecular material that make up particular nucleotides. Genes are ‘immortal’ in their formal structure, not their material and it is this formal capacity for perfect replication that makes them the reproductive units.[[4]](#footnote-4)

In contrast to Dawkins, Godfrey-Smith ([2009]) presents different categories of reproduction. Categorizing reproduction allows Godfrey-Smith to allow some instances of purely ‘formal’ reproduction when the reproduction is done via scaffolding, like when external machinery is used to create a new entity that might not have any material connection to its parent. An example of a scaffolded reproducer is a retrovirus because it relies on structures other than itself to create new entities. Retroviruses infect cells, causing the DNA of the cell to form copies of the virus which are then transcribed back into the RNA of the virus. The production of the retroviruses use material found in the cell. Thus, while a ‘parent’ retrovirus is causally responsible for its offspring, it provides no material parts to the offspring (Godfrey-Smith [2009], pp. 79-80). However, Godfrey-Smith acknowledges that most mechanisms of reproduction typically have material overlap. And even with a broader set of reproductive categories, he concedes that ‘many actual cases fall outside and between these categories’ (Godfrey-Smith [2009], p.89). He goes so far as to say that any attempt to try to capture all instances of reproduction within Darwinian processes is akin to ‘herding cats’, which is to say the biological world is too diverse to categorize. If biological reproduction can be a purely formal affair, then synthetic genomes might well be parts of the lineage from which they’re copied.

However, there’s another account of reproduction that rivals Godfrey-Smith’s and Dawkins’ as a result of what seems to be a flaw in their views. Griesemer ([2000a], [2000b], [2005], [2014]) and Piotrowska ([2017], [2018]) offer accounts of reproduction that distinguish between reproduction and mere production by pointing out the difference between material and information. Griesemer argues that a necessary condition of biological reproduction is that there be material overlap between parent and offspring, ‘the parent-offspring relationship is not merely one of resemblance but also one of material overlap’ (Griesemer [2014], p.26). [[5]](#footnote-5) He obverses that informational overlap isn’t sufficient because replicators do ‘not provide resources to identify empirically the physical avatars of his (Dawkins) functional entities’ (Griesemer [2005], p.79). Because Dawkins and Godfrey-Smith’s accounts do not focus on the material transfer from parent to offspring, they include too many forms of copying as instances of reproduction. Consider an example used by Griesemer and Piotrowska against Dawkins: a photocopier makes copies of an original document, yet we do not think the document has reproduced because no material is shared between the original and the copy. It’s hard to see why this process isn’t reproduction on Dawkins’ account. To avoid including photocopiers as reproducers, Griesemer and Piotrowska include material overlap as a necessary condition of biological reproduction. Thus, Griesemer introduces the material overlap condition: ‘material overlap is a necessary condition for a system to count as biological, and thus that all cases of biological lineages of replicators include material overlap as a property’ (Griesemer [2005], p.79). For Griesemer and Piotrowska, transfer of material parts between parent and offspring, along with the conditions of development, and for Piotrowska ([2017]) persistence*[[6]](#footnote-6)*, reliably avoid the cases of mere production while including cases we intuitively take to be instances of biological reproduction. This is the Overlap, Development and Persistence (ODP) account of reproduction (Piotrowska [2017]).

Adjudicating between these accounts turns on the difference between material and information in biology, which has a history that goes back at least as far as Aristotle.[[7]](#footnote-7) And the literature on genetic ‘information’ is plentiful.[[8]](#footnote-8) Godfrey-Smith notes, ‘some have hailed the employment of informational concepts here as a crucial advance. Others have seen almost every biological application of informational concepts as a serious error, one that distorts our understanding and contributes to lingering genetic determinism. Most of the possible options between these extreme views have also been defended’ (Godfrey-Smith, [2007], p.2 – citations omitted). In a similar vein, Griesemer sees the dichotomy between information/material as a helpful heuristic, though it has the tendency to oversimplify many complex forms of scaffolded reproduction (Griesemer [2014]), as we shall see. Yet, he thinks that at one side of the spectrum the use of language is a clear case of informational overlap. Insofar as the dichotomy is useful, what would be the consequence if information can be sufficient for maintaining a cohesive biological lineage?

As Griesemer and Piotrowska saw, if informational overlap were sufficient for biological reproduction, reproduction becomes extremely permissive and makes it hard to see why things like photocopying would not count as biological reproduction. Is there a way to defend Godfrey-Smith and Dawkins against this charge? To rein in the over permissibility of informational overlap further constraints would have to be added for an account of reproduction that tracks our intuitions about what counts as biological reproduction. Perhaps adding further necessary conditions could yield such an account; for example, that reproduction can only take place between biological entities. But in my view, the most obvious attempts at new constraints look to fail for two reasons.

First, if information is sufficient for reproduction between biological organisms certain instances of language (clear instances of pure information[[9]](#footnote-9)) might fall within the constraints. Second, it’s unclear how much transfer of information would be sufficient for there to be an instance of reproduction. Consider our case of synthetic genomes. If scientists copy ninety-nine percent of the material of *S. cerevisiae*, but just happened to randomly generate the last one percent, would that be enough to meet the informational overlap condition, and hence count as an instance of reproduction? What about the inverse? What if scientists copy just one percent of the genetic notation of *S. cerevisiae* and guess the remaining ninety-nine? Constraining informational overlap looks to be extremely challenging, if not impossible. And while we could throw up our hands and say there’s no way to herd all of the cats, as Godfrey-Smith does, let’s look at Griesemer’s argument for why some of the problem cases, like retroviruses, can appear formal.

Griesemer disagrees with Godfrey-Smith’s claim that retroviruses replicate without any material overlap. Griesemer ([2014]) argues that whether we understand DNA replication in viruses (like HIV) as formal or material turns on how we ascribe ‘ownership’ to the material parts. Such viruses replicate using the material components of the host cell to assemble a new virus DNA. It seems the ‘offspring’ virus is composed of host cell material without sharing any of its ‘parents’ material. But Griesemer urges us to reframe how this process is viewed. Although it seems there is no direct transfer of material, one might see the process as one where a parent entity takes ownership of material as it disassembles some host or other. As we have seen, this is precisely what happens during scaffolded reproduction and it is where Griesemer and Godfrey-Smith differ in their accounts. Let’s examine why Griesemer ([2014]) thinks this process should be understood in this way.

On the one hand, we might trace the material parts of the offspring virus to the ‘source of components’, which would be the material of the cell – this is Godfrey-Smith’s view. On the other hand, we might trace the ownership of the material to the ‘assembler’ of the material parts, in which case the material belongs to the transcripting virus because the virus takes ownership of the material when it does its transcripting. Tracing the material parts to the transcripting virus turns the process into a form of biological reproduction in accordance with Griesemer’s view because there is material overlap. Why does tracing the material in this way make sense? It makes sense insofar as there must be some point at which external material is internalized and incorporated into an entity. For example, it would be odd to say that the material components that make up an organism should be thought of as the food eaten by the organism, rather than just parts of the organism. Ascribing ownership in this way would entail that the organism’s material never really belongs to it, but that the material continues to be property of the food. Using this notion of ownership leads to an infinite regress where material becomes metaphysically basic and it can never be considered part of some larger individual thing. It is rather the case that when an organism processes food into its own structure it takes ownership of the incoming material. I believe Griesemer thinks that we should understand the scaffolding process of retroviruses in a similar way. When a retrovirus dissembles genetic material to make a copy of itself using scaffolding, it takes ownership of that genetic material before passing onto its offspring.

Armed with Griesemer’s notion of ownership, we are now in a position to ask whether the processes used to create synthetic genomes generate the right kind of ownership of material parts to count as a reproductive process. I believe these synthetic processes take ownership in the right way – which means we can reject the first and the third premises of the argument for lineage-less genomes. The next section provides an account of how, during synthetic reproduction*,* assemblers take ownership of genetic material in such a way so as to locate synthetic genomes within an existing lineage.

**5. Synthetic reproductive processes and their implications**

To see why the synthetic processes used to create synthetic genomes are a form of biological reproduction requires carefully tracking the material that belongs to a biological entity and seeing that synthetic reproduction is a scaffolding-like process. The framework of Griesemer’s concept of ownership of material parts provides a way to do this. Tracing material in this way also reveals some interesting features of lineages and reproduction.

Synthetic reproduction is a scaffolding-like process because this form of reproduction uses external machinery. While Godfrey-Smith would likely view synthetic reproduction as it was described in section two as a process of formal reproduction, Griesemer’s account has established that there is material overlap during scaffolded reproduction, like during virus replication. Using this notion of ownership, it seems in assembling a synthetic genome the researchers using the actual machines and synthetic processes come to be the ‘owners’ of the material parts that contribute to the development of the genome.[[10]](#footnote-10) In this case, the researchers, machines, and other parts used in the process take ownership of the dNTPs and transmit them to the chromosome thereby serving as the material that forms overlap. So, I agree with Piotrowska’s ([2018]) general point that these processes are forms of reproduction. And, because these synthetic processes are a form of biological reproduction it seems as though a synthetically derived genome is a part of the genetic lineage from which it was copied, though I believe this extends to even the most conservative methods used to create genomes. The third premise of the argument against lineage-less genomes turns out to be false and the argument can be denied.

Griesemer’s notion of ownership allows for a seemingly tidy resolution to the question posed by synthetic genomes: the synthetic processes used to create synthetic genomes are reproduction because of the overlapping material, and this is what creates causal link to the original genetic lineage. Yet, this isn’t the end of the story. There are further implications for understanding synthetic processes in this way. This is where I diverge from Piotrowska’s argument ([2018]) that reproduction is sufficient for species membership. The most important of these implications is that biological reproduction, by itself, looks to be an insufficient criterion for something like species membership for the following reason.

In normal biological reproductive processes material and information appear to be necessarily entwined with each other. However, in synthetic reproduction they come apart. To see how they come apart, let’s return to the similarity synthetic reproduction has with scaffolded reproduction. Both processes have ‘external machinery’ that is responsible for the creation of the new entity. Normally, the external machinery responsible for creating a biological entity are parts of the same, single lineage. A reproducer yields something somewhat similar to itself, like when a member of the same species or that has a similar genome creates something similar to itself. But in synthetic reproduction the outside machinery responsible for the actual ‘reproducing’ of synthetically derived genomes doesn’t participate in the lineage in which the genome will be a part. The external machinery in synthetic reproduction are literally the methods and machines used by synthetic biologists, like PCR or CRISPR. Unlike retroviruses that create more retroviruses, the machines and humans involved in synthetic reproduction aren’t creating more machines and humans – they’re creating genomes. While we could say the genomes they create are part of the genetic lineage of *Homo sapiens* because *Homo sapiens* are parts of all that external machinery, that would be an extremely odd twist of evolution (to say the least). The more plausible alternative is to conclude that it’s metaphysically possible for vertical genetic lineages and reproduction to come apart. Looking at the process, this seems right. Synthetically creating a genome isa reproductive process in a true sense for all the reasons outlined in the ODP account of reproduction. A ‘parent’ of sorts overlaps material with the offspring, there is development, and persistence. Now we need an account of how a genetic lineage is maintained.

Lineages, it seems, can be maintained through information. A synthetically derived genome that is modeled on the molecular sequence of the genomes that are part of some genetic lineage can only be created by researchers and machines if there is a transfer of information (short of monkeys at typewrites just happening upon the right string of information). And this informational transfer that generates the causal continuity of the biological lineage, not the material. Without modeling the synthetic genome on the parts of the original or native genes there would be no way to construct a synthetic genome short of sheer luck. Even if by sheer luck researchers happened to land on the right genetic sequence, such luck would fail to be a causal link.

This means biological reproduction isn’t permissive, but biological lineages can be quite permissive when they are freed from the material constrains that are ubiquitous throughout most of the ‘natural’, biological world. Parental entities don’t have to be parts of the lineage of their offspring so long as they adhere to the constrains of the ODP. The possibility of synthetic reproduction shows that lineages can be maintained even though the entities involved in scaffolding reproductive process can be parts of other lineages (or even non-biological entities). The upshot is that lineages do the work of maintaining things like species, not reproductive processes (one might think that reproductive processes are responsible for the parts of the species). It seems that material and informational overlap are generally treated as a single process because, as an empirical matter in ‘natural’, biological processes, lineage information always comes in tandem with material. Insofar as reproduction usually takes place within the confines of a single genetic lineage, the differences between lineages and reproduction are opaque. I believe synthetic biology is beginning to let us see through this opacity. Freeing up lineages from their connection to biological material shows that the cohesive mechanism of a lineage turns out to be quite thin. So thin it can actually pass through other entities as information. In this way, synthetic genomes shed some light on the metaphysics of lineages (and perhaps species) as they show that genetic lineages can be realized through vastly different mediums. No wonder lineages and species can be so hard to locate and delimit.

If this account is right, while a synthetically derived genome would be a part of the lineage upon which it was modeled, its reproductive etiology would be vastly different than all other parts of the genetic lineage. Biological reproduction is not necessarily a guarantor of a cohesive lineage as the entities responsible for the reproduction can have an entirely different, even non-biological origin.

This also tells us that certain assumptions that have been made about the species generating cohesive properties of reproduction turn out to be wrong, for example (Hull [1992] or Ghiselin [1997]). An obvious case seems to be the Biological Species Concept. But also consider Hull’s claim that if a scientist made an organism that was identical to some already existing organisms, this organism would have to go on to interbreed with members (or parts) of that species before it would become a part of that species (Hull [1978], pp. 349-50). Its interbreeding might make it part of the species… or it might not. Consider instances where lineages fuse together in sort of inoculation, like when members of two different species interbred to create a fertile hybrid organism. No one claims that one of the parent organisms becomes a member of the other parent’s species after they’ve reproduced. I’m actually inclined to think that a scientifically created organism wouldn’t be lineage-less in the way that Hull implicitly thinks. Such an organism could only come into existence having a tie to the species and to the genetic lineage on which the scientists modeled it because the only truly lineage-less organisms would be ones generated through complete randomness. The scientists would actually be engaged in a truly reproductive, not merely productive, process in creating it. Analogously, synthetic genomes are not randomly generated, so they must have a historical cohesion to genes that participate in a lineage, even if they’re not reproductively linked to those genes. Yet they’re still the result of reproductive processes. Therefore, reproduction, by itself, can’t have the cohesive property Hull, Piotrowska and others attribute to it. And, this means that Jef Boeke’s (one of the key researchers involved in synthetic yeast projects) thought that ‘the two-chromosome yeast might qualify as a distinct species because it can’t breed with normal yeast, despite having near-identical DNA’ (Callaway [2018]) might be wrong. If all the genetic information comes from a unified, genetic lineage, even though it can’t interbreed with native yeast, I believe it is still a part of the same species.

We can now return to the question posed by Piotrowska ([2018]): is deextinction possible? Piotrowska believes that conservative methods of synthetic reproduction that lack material overlap can’t be reproduction, and hence those methods couldn’t create a link sufficient for species membership. Therefore, she argues that deextinction isn’t possible using informationally created synthetic genomes (though she thinks deextinction might be possible if semi-conservative methods are used). But, as we have seen, a genetic lineage doesn’t require reproduction. If an organism was created that had a synthetic genome modeled on mammoth DNA, it would at least be a hybrid mammoth assuming that genes play an important role in what determines species membership, even though it wouldn’t have any mammoth ancestors. Some of its ancestors might be humans and machines.

The process I’ve called ‘synthetic reproduction’ is particularly complex and very unlike other forms of reproduction seen in the biological world. It would be a form of reproduction where the genes (and-or machines) of one species (or no species) are the progenitors of genes of another species without obvious parent-offspring relations, though the synthetic genes participate in a genetic lineage. It shows how thin genetic lineages can be. I think this merely amounts to acknowledging that synthetic biology has a lot to teach us about the boundaries of the biological world. When it is placed alongside other biomedical and technological advancements, it looks to be another instance where science is shifting traditional boundaries in biology. This seems like it would be true with the advent of a synthetically created eukaryotic genome.

**Appendix**

**Acknowledgements**

I am deeply grateful to Monika Piotrowska and P.D. Magnus for providing extensive feedback and discussions on early versions of this paper. I would also like to thank Ron McClamrock, Matthew Mosdell, Henry Curtis and the participants at the University of Calgary Philosophy Dept. Graduate Conference on Philosophy of Biology for providing their thoughts on this paper. I am also grateful to the anonymous referees who provided valuable suggestions on how to improve this paper.

*Gunnar Babcock*

*Philosophy Dept., University at Albany*

*Albany, NY*

*gbabcock@albany.edu*

**References**

Annaluru, N., Muller, H., Mitchell, L. A., Ramalingam, S., Stracquadanio, G., Richardson, S. M., … Chandrasegaran, S., [2014]: ‘Total Synthesis of a Functional Designer Eukaryotic Chromosome’*, Science*, **344(6179)**, pp. 55–58

Callaway, E. [2018]: ‘Entire yeast genome squeezed into one lone chromosome’, *Nature,* available at <https://www.nature.com/articles/d41586-018-05857-9>

Dawkins, R. [1976]: The Selfish Gene, Oxford: Oxford University Press.

Dawkins, R. [2005]: *The Ancestor’s Tale: A Pilgrimage to the Dawn of Evolution,* Boston:Houghton Mifflin Harcourt Press.

Doolittle, W.F. and Bapteste, E. [2007]: ‘Pattern pluralism and the Tree of Life hypothesis’, *Proceeding of the National Academy of Sciences*, ***104*(7)**, pp. 2043-2049

Ereshefsky, M. [1992]: ‘Eliminative Pluralism’, *Philosophy of Science,* ***59*(4)**, pp. 671-690

Ghiselin, M. [1974]: ‘A Radical Solution to the Species Problem’, Systematic Zoology, **23**, pp. 536–544

Ghiselin, M. [1997]: *Metaphysics and the Origin of Species,* Albany, NY: State University of New York Press.

Gibson, D., Benders G., Andrews-Pfannkoch C., …Smith, H., [2008]: ‘Complete Chemical Synthesis, Assembly, and Cloning of a *Mycoplasma genitalium* Genome’, *Science*, **319(5867)**, pp. 1215-1220

Godfrey-Smith, P. [2000]: ‘Information, Arbitrariness, and Selection: Comments on Maynard Smith’, Philosophy of Science, **67**, pp.202–207.

Godfrey-Smith, P. [2007]: ‘Information in Biology’, in D. Hull and M. Ruse (*eds.*) The Cambridge Companion to the Philosophy of Biology, Cambridge: Cambridge University Press.

Godfrey-Smith, P. [2009]: *Darwinian populations and natural selection,* Oxford: Oxford University Press.

Griesemer, J. R. [2000a]: ‘Development, Culture, and the Units of Inheritance’, Philosophy of Science, **67** (Supplement. Proceedings of the 1998 Biennial Meetings of the Philosophy of Science Association. Part II: Symposia Papers): S348–S368.

Griesemer, J.R. [2000b]: ‘The units of evolutionary transition’, Selection, **1(1–3)**, pp. 67–80

Griesemer, J.R. [2005]: ‘The informational gene and the substantial body: on the generalization of evolutionary theory by abstraction’, in Idealization XII: Correcting the Model.

Griesemer, J.R. [2014]: ‘Reproduction and the Scaffolded Development of Hybrids’ in Caporael *et al.* (eds), *Developing Scaffolds in Evolution, Culture, and Cognition*, Cambridge: MIT Press.

Hull, D. [1976]: ‘Are species really individuals?’, *Systematic Zoology.* **25(2)**, pp. 174–191

Hull, D. [1978]: ‘A Matter of Individuality’, Philosophy of Science, **45**, pp. 335–360

Hull, D. [1992]: ‘An evolutionary account of science: A response to Rosenburg’s critical notice’, *Biology and Philosophy,* **7(2),** pp. 229-236

Lawrence, J. and Retchless A. [2010]: ‘The myth of bacterial species and speciation’, *Biology and Philosophy,* **25 (4)**, p.569-588

Lennox, J. [2006]: ‘SEP Entry in Aristotle’s Biology’, in E. N. Zalta (*eds.*), *Stanford Encyclopedia of Philosophy,* available at <<https://plato.stanford.edu/entries/aristotle-biology/>>.

Luo, J., Sun, X., Cormack, B.P., Boeke, J.D. [2018]: ‘Karyotype engineering by chromosome fusion leads to reproductive isolation in yeast’, *Nature,* **560**, pp. 392-396

O’Malley, M. [2014]: *Philosophy of Microbiology*, Cambridge: Cambridge University Press.

Preston, C. [2008]: ‘Synthetic Biology: Drawing a Line in Darwin’s Sand’, *Environmental Ethics*, **17**, pp. 23-39

Piotrowska, M. [2018]: ‘Is ‘Assisted Reproduction’ Reproduction?’ *The Philosophical Quarterly,* **68(270)**, pp.138-157

Piotrowska, M. [2018]: ‘Meet the New Mammoth, Same as the Old? Resurrecting the *Mammuthus primigenius*’, *Biology and Philosophy*, **33:5**, available at <https://doi.org/10.1007/s10539-018-9616-4>

Wu, Y., Li, B., Zhao, M., Mitchell, L., Xie, Z… Yuan, Y., [2017]: ‘Bug mapping and fitness testing of chemically synthesized chromosome X’, *Science*, **355(6329)**

Xie, Z., Li, B., Mitchell, L., Wu, Y., Qi, X… Yuan, Y., [2017]: ‘‘Perfect’ designer chromosome V and behavior of a ring derivative’, *Science*, **355(6329)**

Zhang, W., Zhao, G., Luo, Z., Lin, Y… Dai, J., [2017]: ‘Engineering the ribosomal DNA in a megabase synthetic chromosome’, *Science*, **355(6329)**

1. ‘PCR’ is a precise means of amplifying, or copying, a specific DNA target from a mixture of DNA molecules. [↑](#footnote-ref-1)
2. ‘Functioning’ in this context means that the organism with synIII has a level of fitness that is nearly the same as a native strain of *S. ceresisae*. The authors explain: ‘to check for negative effects of modifications of fitness of synIII-containing strains from the WT (BY4742), we examined colony size, growth curves, and morphology under various conditions. A growth curve analysis established that synIII and the isogenic native strain had no detectable fitness differences. The strains were also indistinguishable from each other on colony-size tests’ (Annaluru, *et al.* [2014], p.57). [↑](#footnote-ref-2)
3. David Hull shares a similar position; that reproduction takes place via replication and he argues that being a replicator is one of the marks of biological individuality (Hull [1978], p.340-1). [↑](#footnote-ref-3)
4. This does *not* mean that particular sequences that make up a gene are historically unconnected on Dawkins’ model: for example, the sequence ‘A-C-A’ might happen to remerge through mutation millennia apart, but it is not the same gene. Genes, though perfectly replicated, are linked through historical connection of one replicating the next. [↑](#footnote-ref-4)
5. In response to Godfrey-Smith ([2009]), Griesemer ([2014]) notes that scaffolded, hybrid reproducers do not necessarily require strict adherence to material overlap. His account can accommodate this because the condition of material overlap should be treated as a heuristic for reproduction in biology: ‘..my project is heuristic and empirical. I do think material overlap is basic to all cases of reproduction, so if there is less of it in the story of some central subjects than others, then my heuristic proposal is to look elsewhere for the material relations leading to adequate explanation of the development organization we observe [in reproduction]’ (Griesemer [2014], pp. 36-7). But even if material overlap is a heuristic rather than a necessary condition, it is still the case that Griesemer’s paradigmatic case of material overlap is the reproduction of eukaryotic chromosomes (Griesemer [2014], pp. 26-7) like those considered here. Furthermore, because we’re considering a hypothetical case that would seemingly have no material overlap at any point, for our purposes I will treat Griesemer’s account as holding to material overlap as a necessary condition. [↑](#footnote-ref-5)
6. Piotrowska ([2017]) convincingly argues that persistenceis a necessary addition to Griesemer’s account of reproduction, because without including persistence things like uterus transplants would be instances of biological reproduction. So, I will refer to this account of reproduction as the ‘ODP’ account. [↑](#footnote-ref-6)
7. Here we might understand information as form and material as matter. Consider (Lennox [2006]) and (Godfrey-Smith [2009], p.79). Insofar as Plato advocated that form was primary cause and Aristotle held a hylomorphic view, where biological reproduction requires a combination of form and matter (or material), Griesemer and Godfrey-Smith’s positions might represent a modern version of this debate. [↑](#footnote-ref-7)
8. While there is a difference between information and formal processes (for example, there can be formal notation that fails to convey any information), I will treat them as being equivalent. I do this because the types of informational and formal processes we’re considering share the important feature of being non-material*.*  [↑](#footnote-ref-8)
9. Consider Griesemer ([2014], p.29). Language is non-material information of sorts being transferred, which can be notated, between biological organisms. In some cases, like that between a human parent and child, linguistic information contributes to development, phenotypical characteristics of the child. There is no easy or obvious way of constraining non-material overlap to match up with our typical notion of biological reproduction as it pertains to genetic lineage. [↑](#footnote-ref-9)
10. This entails that even if the actual material used in the process is ‘natural’ (as opposed to synthetic chemical like XNA), the scaffolders ‘naturalize’ the material thereby taking ownership of it. Thus, it doesn’t matter if the chemical material is originally natural or synthetic. [↑](#footnote-ref-10)