Bonobo evolution: from the perspective of genomes

Sojung Han (한소정)

TESI DOCTORAL UPF / YEAR 2019

Thesis supervisor

Dr. Tomàs Marquès-Bonet

Dra. Aida M Andrés (UCL Genetics Institute, UK)

DEPARTAMENT DE CIÈNCIES EXPERIMENTALS I DE LA SALUT

사랑하는 마르틴과 루디에게
그리고 사랑하는 엄마에게

"Ignorance more frequently begets confidence than does knowledge: it is those who know little, not those who know much, who so positively assert that this or that problem will never be solved by science."

Charles Darwin, The Descent of Man

Acknowledgement

Starting a PhD program, studying evolution in bonobos using their genomes, has not been an easy decision to me, and completing it has been a great effort. Interestingly, I enjoyed a lot that long journey, which served as a compass in the dark corner of my life. Beautiful hands were supporting me through this. I am very grateful that they were there and that I could take them.

First, there were my two supervisors.
Tomas, you gave me the chance to open up this great journey and the support I needed.
Aida, you have been so understanding, patient and supportive throughout the way.

I also owe a lot to my dear group mates, all of whom enriched my days. Particularly, I would like to mention,
Jessica H, you are probably the most kind person I met. Sitting next to you was always warm.
Irene, you know how I liked our conversations. I will miss our time in the sun talking over all possible things.
Claudia, you were so considerate and helpful. Thank you for translating my abstract.
Marc, you always helped out with nice comments.

Jessica N, it is always fun to hang out with you. It is inspiring to talk to you.

I do not have enough words to describe my thanks to my family, for the great and strong support I have received.

Martin, I am a lucky person to meet you. It is amazing that you are there. What I am today cannot be explained without you.

Ludi, you made me ever bigger and stronger, and I learned that it was all possible. You are so beautiful, and so is my life with you.

태어나던 날부터 내 마지막 날까지 모든 하루 속에 존재하고 있을 엄마, 돌아보면 늘 같은 곳에 있는 변치않음에 감사해요.

때로는 언니같은 동생, 소윤이도 늘 고마워.

Lastly, the sun in Barcelona, you are so powerful.

Barcelona, May 20, 2019

Sojung Han (한소정)

Abstract

Bonobos are an interesting species to study due to their unique evolutionary features such as a peaceful social nature and hypersexual behavior. The prolonged exaggerated sexual swelling in females is a notable difference to chimpanzees, from which they diverged < 2 Million years ago. In my studies, I first show using whole-exomes, that there is genetic population substructure in bonobos. I estimate that the split between the western and the central bonobo population might be >100,000 years ago, comparable to modern human population split times. Second, I show that lineage-specific non-synonymous derived alleles in bonobos are enriched in genes associated with 'age at menarche', which suggests that the differences between bonobo and chimpanzee females in their exaggerated sexual swellings might be due to unique genetic changes in bonobos since their divergence. Finally, I show, across bonobos and four chimpanzee subspecies, that the efficacy of purifying selection correlates with effective population size, with bonobos and western chimpanzees showing similar levels.

Resum

El bonobo és una espècie d'especial interès per les seves característiques evolutives úniques: la seva naturalesa pacífica i social, i el seu comportament hipersexual. La principal diferència amb els ximpanzés, que van divergir dels bonobos fa 2 milions d'anys, és el llarg període d'inflor sexual exagerada en les femelles. En la recerca presentada en aquesta tesi, en primer lloc he demostrat que hi ha estructura poblacional genètica en bonobos, utilitzant dades d'exomes sencers. Alhora, he estimat que la separació entre els bonobos occidentals i centrals va ocórrer fa >100.000 anys, una data comparable amb la divergència amb els humans. En segon lloc, he examinat com els al·lels derivats no sinònims i específics de llinatge en bonobos estan enriquits en gens associats a "l'edat de menarquia". Això suggereix que les clares diferències en la inflor sexual exagerada entre bonobos i ximpanzés poden ser degudes a la l'evolució privada dels bonobos des de la seva divergència amb els ximpanzés. Finalment, he mostrat com l'eficàcia de la selecció purificadora correlaciona amb la mida poblacional efectiva tant amb bonobos com en els quatre llinatges de ximpanzés, sent els bonobos i els ximpanzés occidentals molt similars entre ells.

Preface

What makes bonobos bonobo?

In order to answer this philosophical and abstract question, one may revise what we know about bonobos (*Pan paniscus*). Bonobos are the only sister species to chimpanzees (*Pan troglodytes*), and they are both equally closely related to humans. Outside academia, bonobos are often referred to by their nicknames 'Hippie chimpanzee' or 'Pygmy chimpanzee'. As the names show, bonobos came into our attention through comparison to chimpanzees. Since the 1950s, when the renowned anthropologist Louis Leakey started a long-term expedition to study great ape behaviors in the wild, as a way to understand the origins of humans and of human behavior, we have rapidly become familiar with chimpanzees, the first great ape species he studied together with Jane Goodall. Discoveries of chimpanzee behavior in the wild made us realize how much similar they are to us: They use tools. They have personalities. They have friendship and politics. They have wars. These findings challenged us and made us ask if the definition of human as a 'thinking being' and 'tool-using being' could be enough, when chimpanzees could also think and use tools.

Not long after, studies of bonobo behaviors in the wild, which started in the 1970s, surprised us with completely new findings. We got to know that those chimpanzee-looking beings were completely different from chimpanzees. They were peaceful and egalitarian. They did not have wars. Adult females were not subordinate to adult males, unlike chimpanzees. They had much more frequent

sexual interactions, which was in fact at the core of their social behavior, including homosexual behavior, *e.g.* female-female genital rubbing behaviors, which indeed gave them the name 'Hippie chimpanzees'. Considering that the divergence between the two species is estimated to be 1 to 2 million years ago (Prado-Martinez et al. 2013; De Manuel et al. 2016), which is rather a short time window in the evolutionary time scale, such interspecific differences in their behavior appear striking.

How come they have evolved into such different ways? This apparent question, however, is still quite difficult to answer. The main reason which has been pointed out is that the available data on bonobos is limited (Stanford 1998; Boesch 2002). Bonobos are understudied, compared to chimpanzees, for a few reasons. The first reason is that they came into attention for research much later than chimpanzees. When they were introduced to Europe, they were first mistaken as chimpanzees, which were described with the species name '*Simia troglodytes*' very early (Blumenbach 1776). Not until 1929 were they recognized as a different subspecies of chimpanzees (Schwarz 1929), and only in 1933 were they considered a distinct species (Coolidge 1933). Another reason would be that their habitat is limited to the current Democratic Republic of Congo (DRC), unlike chimpanzees, whose habitat ranges across different climate zones across equatorial Africa (McGrew 1983; Nishida 1983; Boesch et al. 2008; Wrangham 1996; Whiten et al. 1999). The long-term studies of bonobos in the DRC, which began only in the 1970s, have often been complicated by political instability in the DRC, due

to civil wars and coups. However, the researchers from different fields of biology continue the endeavor.

The work I present here is a part of the effort to understand this special species and its evolution, by making use of their genomes. It heads towards the opening question: What makes bonobos bonobo? What was their past? What is their present?

Table of contents

Abbreviations

DNA deoxyribonucleic acid

DRC Democratic Republic of Congo

mtDNA mitochondrial DNA

N_e effective population size

RNA ribonucleic acid

SFS Site-Frequencies-Spectrum

SNC Single Nucleotide Change

SNP Single Nucleotide Polymorphism

1. INTRODUCTION

When Louis Leakey, who was a dedicated paleoanthropologist and archaeologist in discoveries of human evolution in Africa, set up his team to study three different great ape species in the wild, chimpanzees (*Pan troglodytes*), gorillas (*Gorilla*) and orangutans (*Pongo*) (Trimate: Morell 1993), his ambition was to reconstruct the human behavior in the past from understanding the present behavior of the species which is our living relative. As he said, behavior does not fossilize (Montgomery 1991), and only through comparing behaviors of the related species could we reconstruct the behavior of our common ancestors.

Such a comparative approach is also relevant in understanding the evolution and behavior of bonobos (*Pan paniscus*). We know that bonobos are peaceful and egalitarian, and that female bonobos have high social status and are at the center of the social dynamics in their community (Furuichi 1987; Idani 1991; Vervaecke, de Vries and van Elsacker 2000). Such traits can be recognized by comparing bonobos to their most closely related species, chimpanzees (Goodall 1986; de Waal 1982; Nishida 1983; Lehmann and Boesch 2008), since these are the only two species of the genus *Pan*. A better understanding should derive not only from cross-comparing them, but also from taking humans into account. This is true not only for behavioral biology but also for comparative genomics.

Until recently, it was still not clear whether gorillas or chimpanzees are closer to humans (Sarich and Wilson 1967). Only in the 1980s,

it became apparent, from direct sequencing of nuclear DNA and mitochondrial DNA (mtDNA) segments, that humans and chimpanzees share a more recent common ancestor (Hasegawa and Kishino 1984). Since then, with the technological development in molecular biology and the theoretical advancement in population genetics, we were able not only to explore the demographic history of the great apes (Langergraber and Prüfer 2012; Prado-Martinez et al. 2013; De Manuel et al. 2016), but also to detect potentially selected regions of their genomes on each branch (Cagan et al. 2016).

In the following sections, I will describe each topic I approached in the light of bonobo evolution, based on the advancement of the field I described. In the first section, I present evidence that there seems to be substantial divergence between bonobo populations, with an estimated split of at least 100,000 years ago, based on genetic population structure in wild bonobo genomes with unknown geographic origins. In the second section, I investigate female reproductive traits, by analyzing potential genetic changes which might explain the most noticeable phenotypic differences between bonobos and chimpanzees. Finally, I study mutational load, by exploring the evolutionary signature of all the known bonobo and chimpanzee subspecies, in terms of the efficacy of purifying selection in relation to their demographic history.

1.1. Phenotypic traits in female reproduction

At the behavioral level, the more egalitarian social interactions and much higher social status of females in bonobo, compared to

chimpanzees, are recognized as the most significant differences between bonobos and chimpanzees (Furuichi 1987; Kanō 1992; Wilson and Wrangham 2003; Mitani, Watts and Amsler 2010; Boesch et al. 2008). These are interesting, as they otherwise generally share a very similar social structure.

They both have a patriarchal society composed of multiple adult males and adult females, where females migrate out from their natal group on sexual maturity (Furuichi and Hashimoto 2004). Males in the same population are in theory genetically related, although paternity is usually unclear, as estrous females mate with multiple males which is described as 'promiscuous' (Goodall 1986; Kanō 1992). Seasonal fruits and other types of vegetation are their primary diet, with occasional animal consumption, such as termites and monkeys (Badrian, Badrian and Susman 1981; Boesch and Boesch-Achermann 2000; Mitani and Watts 1999; Hohmann and Fruth 1993). They both reside in equatorial Africa, even though their habitats do not overlap: Bonobos inhabit the current Democratic Republic of Congo (DRC), south of the Congo river, and chimpanzees live across central Africa, from Uganda to Guinea (Figure 1, De Manuel et al. 2016), encompassing different climates from topical to savanna (van Lawick-Goodall 1968; Nishida 1968; Boesch and Boesch-Achermann 2000; Pruetz and Bertolani 2007; Kanō 1992; Thompson 1997; Badrian, Badrian and Susman 1981).

There are several hypotheses to explain why they would be so different in their social behaviors nevertheless. Such hypotheses often start from assumptions on their ecology, particularly that

Figure 1. The habitat ranges of chimpanzees and bonobos: dark blue: western, red: Nigerian-Cameroon, green: central, orange: eastern chimpanzee population, and purple: bonobo (De Manuel et al. 2016).

bonobos have evolved in dense and rich forests (Kanō 1992; Hashimoto et al. 1998; Serckx et al. 2014), and that chimpanzees have evolved in environments where they needed to compete over food and were encouraged to use diverse tool using skills (Whiten et al. 1999; Gruber, Clay and Zuberbühler 2010). They both have a 'fission-fusion' society, where each community, a unit as a society, is composed of unique members who forage as a small group fluctuating in group size, and such a group is referred to as 'party' (Chapman, Chapman and Wrangham 1995). As bonobos tend to have bigger party sizes than chimpanzees, it used to be associated to their differential ecological environments: the average number of individuals per party is estimated to be 11 to 22 for bonobos and 4 to 10 for chimpanzees (Kuroda 1979; Idani 1991; Mulavwa et al. 2008; Sakura 1994; Nishida 1968; Boesch and Boesch-Achermann

2000). As better quality of food patches would confer decreased levels of competition and staying in a big group would be better for protection from predators, the bigger party size on average in bonobo populations has often been explained in relation to their richer environment, relatively to chimpanzees (Hohmann and Fruth 2002; Itoh and Nishida 2007; Janson and Goldsmith 1995; Lehmann, Korstjens, and Dunbar 2007). In order to test that hypothesis, ideally, on top of the survey of their current ecological environment, one would need to evaluate the actual environment they have evolved in. However, considering the variation in party sizes across chimpanzee communities with varying levels of resource quality and distribution (van Lawick-Goodall 1968; Nishida 1968; Boesch 1991), such an explanation appears probable. Unfortunately, those hypotheses cannot be directly tested using genetics or genomics approaches.

The high social status of female bonobos, however, is often related to the prolonged attractiveness of female bonobos, the longer and more frequent expression of maximal sexual swellings in female bonobos than female chimpanzees (White 1988; Furuichi 1989, 1997; Kanō 1992; Wrangham 1993, 2002; Furuichi and Hashimoto 2002), which is the most striking difference between bonobos and chimpanzees on the phenotypic level. Females of both species, when estrous, present maximal or exaggerated sexual swellings in the perineal part (Furuichi 1987; Wallis 1992) (Figure 2), which is a distinguishing feature in bonobos and chimpanzees, among the great apes, although it is a common trait in many Old World monkeys (Nunn 1999).

Figure 2. Maximal sexual swellings in a female bonobo on the left (Douglas et al. 2016), and in a female chimpanzee (Kappeler 2012).

Such sexual swellings are understood in line with pheromonal cues as a sexually selected trait (Darwin 1871; Deschner et al. 2004), which advertises a window of fertilizability (Nunn 1999; Emery and Whitten 2003). Female sexual advertisements are often observed in non-mammalian species, such as chameleons (Chamaeleo chamaeleon: Cuadrado, 1998) or some birds (Prunella collaris: Nakamura, 1990), whose females change coloring of a part of their body as a signal of sexual receptivity during the breeding season. But those female traits, which could be used by males to evaluate the reproductive quality of females, are observed rather rarely in mammalian females (Andersson 1994). The exaggerated swelling of the sexual skin in primates, including bonobos and chimpanzees, is understood as such a sexually selected trait (Pagel 1994; Nunn 1999).

On the physiological level, sexual swellings are equivalent to an estrogen dependent edema of perineal regions (Krohn and Zuckerman 1937; Zuckerman and Parkes 1939). Not only the tissue

of the swellings contains estrogen and progesterone receptors, like endometrial tissue (Kato, Onouchi, and Oshima 1980; Ozasa and Gould 1982; Tsuneko and Junzo 1983), but also progesterone appears related to detumescence of the swelling by inducing localized estrogen withdrawal and down-regulation of estrogen receptors (Gillman 1940; Gillman and Stein 1941; Carlisle, Brenner, and Montagna 1981; West and Brenner 1990).

What would be the adaptive function of this most likely costly trait? These exaggerated sexual swellings are found mainly in species where females actually or potentially have multiple mates (Clutton-Brock and Harvey 1976; Dixson 1983; Hrdy and Whitten 1987; Nunn 1999). Different theories have been proposed to explain it, as a reproductive strategy for females in this environment, which can be summarized as confusion or reassurance of paternity (Clutton-Brock and Harvey 1976; von Noordwijk 1985). It is described as a way to confuse paternity by successfully attracting multiple males at the same time which otherwise may lead to infanticide, or to promote sexual competition among males to get a chance to copulate by which females could select the best male candidate.

Indeed, infanticides, the killing of a newborn individual by a conspecific, are quite frequently observed across different mammalian species (Lukas and Huchard 2014), including primates, like Hanuman Langurs (Sugiyama 1984) and chimpanzees (Nishida et al. 2003; Williams et al. 2008; Wilson et al. 2014). One established hypothesis explains it as a reproductive strategy in males as it could be a way to remove the unrelated infants sired by

other males, which is understood as sexual competition among males, and this way they could also have an immediate access to fertile females who otherwise would not conceive during months or years due to lactation (Hrdy 1974, 1979). Infanticides are, however, very costly for females who need to invest time and resources for conception, lactation and care-taking in general. Therefore one would assume that females need counter-strategies, like promiscuous copulation, to confuse the paternity, in order to avoid such risks.

However, would promiscuous mating preclude females' choices on the males to sire? In mammalian species, reproductive success in males is understood as the number of their access to fertilizing females which could result in a large number of offspring, and in females as the quality of the male the gametes come from, as the potential maximum number of their offspring is limited throughout their lifetime and the expected investment to each offspring is large compared to males, due to pregnancy and lactation (Bateman 1948; Trivers 1972). Therefore, females are expected to be choosy in selecting the sire (Andersson 1994), as they could increase the chances for male protection, better access to food and other resources (Kirkpatrick and Ryan 1991), or receive 'good genes' for their offspring, which could be advantageous for survival and contest competition (Darwin 1871; Emlen and Oring 1977; Clutton-Brock et al. 1988; Smith 1991). It should be noted that female choice does not simply mean the strongest male. There is evidence that female primates do choose the males who have dissimilar genes at the major histocompatibility complex (MHC), which may help

their offspring to be better equipped against pathogens (Schwensow, Eberle and Sommer 2007; Setchell et al. 2009; Huchard et al. 2010). Various theories suggested potential strategies females could employ to choose the right male in a promiscuous mating system, particularly where they have sexual swellings. Postcopulatory sexual selection would be a good example for it, which is composed of sperm competition and cryptic female choice (Parker 1970; Evans et al. 2003) (Figure 3). Sperm competition, which is primarily a male-male competition, refers to the male sexual evolution to compete with the gametes of other males to first reach the egg of the female (Parker 1970). This could involve penile elongation, as the maximal sexual swellings of females around the ovulation period elongate the vaginal tract due to the inflated perineal part (Dixson and Mundy 1994), which could be a way for females to select competitive traits in males. Cryptic female choice refers to the theory that females have morphological and physiological mechanisms inside to choose certain gametes of their interest (Thornhill 1983; Eberhard 1985, 1996, 2009).

On the other hand, there are at large two different models in predicting the function of the sexual swellings: as an accurate indicator for the ovulation timing (Hamilton 1984) or rather as a confusing sign as an inaccurate predictor for ovulation (Nunn 1999). If the sexual swellings function as a precise sign for ovulation, high-ranked males would have a direct benefit from monopolizing the females with the maximal swellings, which could be beneficial also for females as the strongest male becomes the sire and potentially provides protection for the female and the offspring.

Otherwise, females could make use of the impreciseness of the sexual swellings as an ovulation advertisement, as then it would not be rewarding for males to monopolize females and it would give them a better chance if the females choose them from their own interest.

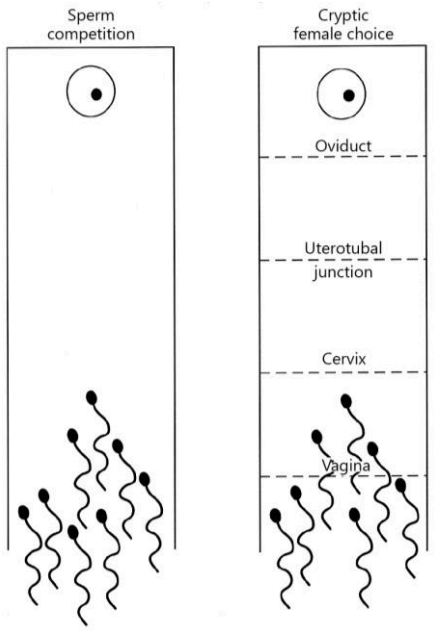


Figure 3. A diagram to represent the concept of postcopulatory sexual selection: sperm competition (left) and cryptic female choice (right) (Dixson 2009).

Indeed, in chimpanzees, the size of sexual swellings has been related to cycle quality and to the proximity of ovulation, both of which have to do with the chances of conception (Deschner et al. 2004; Nunn 1999; Emery and Whitten 2003). However, in bonobos,

it has been reported that their sexual swellings are not precise indicators for the chances of conception, in comparison to chimpanzees (Douglas et al. 2016). Moreover, in bonobos, both the duration of the swelling cycle (Dahl 1986; Furuichi 1987) and the duration of the maximum swelling phase (Blount 1990; Dahl 1986; Furuichi 1987; Kanō 1992; Thompson-Handler, Malenky and Badrian 1984) seems to be longer than in chimpanzees. Another striking difference is that female bonobos appear estrous even when they are not in the phase to ovulate (e.g. during lactating period, Furuichi, 1987; Thompson-Handler et al. 1984), which is referred to as 'pseudo estrous'. Such traits in female bonobos make them more often sexually receptive than female chimpanzees, which (Kanō 1992) describes as 'semi-continuous receptivity'. As the presence of such prolonged maximal sexual swellings is related to elevated female attractiveness to males (Furuichi and Hashimoto 2004; Stanford 1998) and also to females in their genital-genital rubbing behavior (Ryu, Hill and Furuichi 2015), which is known as homosexual behavior, it could be understood as that the female sexual attractiveness is inflated in bonobo society, compared to chimpanzee's. It has been described as 'hypersexuality' (de Waal 1987; Wrangham 1993) and 'higher estrous sex ratio' (Furuichi 2011), and since the lifetime number of offspring per females does not differ between bonobos and chimpanzees, considering their similar interbirth intervals (Kanō 1992; Furuichi and Hashimoto 2002), the more frequent sexual behaviors among bonobos have been suggested to function as a way for females not only to avoid male monopolies and have a choice on the sire, but also to build

alliances among females and earn a higher social status through that, a strategy female chimpanzees do not have. This nature of female bonobo sexuality was also proposed as an explanation of potential sexual selection on non-aggressive males (Wrangham 2018; Self-domestication theory: Hare, Wobber and Wrangham, 2012) and the egalitarian social dynamics described above (Furuichi 2011).

Bonobos and chimpanzees, which diverged 1 to 2 Million years ago (Mya), serve as a good model to compare and understand how the colliding interests on reproductive success between males and females lead to diverse reproductive strategies, which shape phenotypic traits of both sexes. However, what would have been their ancestral state? Which lineage has derived from that? Or can it be that they both have diverged into different states? These are not easy questions to answer. If we include the human state, which is a pair-bonding system with 'continuously receptive females', it gets even more difficult. Some suggest that bonobos are derived (Shea 1983; Wrangham and Pilbeam 2001), while others suggest that chimpanzees are (Kanō 1992), but from a comparative behavioral approach, we can still not answer these questions.

It might be useful to go down to the molecular level for approaching these questions. These phenotypic differences between female bonobos and chimpanzees in their reproductive traits are likely to have physiological mechanisms, which differ from each other. Indeed, at the hormonal level, it has been shown, using urinary testosterone, that female bonobos are on average three years

younger than female chimpanzees when they experience the onset of puberty (Behringer et al. 2014), which is an important clue for their differential sexual development. Such differences in physiology are most likely based on lineage-specific genetic changes. However, previous studies in their genomes, with the goal to find lineage-specifically selected signatures, did not find evidence that female reproduction-related traits were under selection (Prüfer et al. 2012; Cagan et al. 2016). This could be due to the particular sensitivity of each method used regarding the time frame traits were selected at, or the lack of knowledge on the genes involved in the traits. Therefore, in my study, I approach the questions by identifying the lineage-specific genetic changes in each lineage, and investigate if there are selective differences in these lineages possibly associated to these phenotypes.

1.2. Mutational load within the context of demographic history

Natural selection works on heritable phenotypic traits that modify fitness (Darwin 1859; Zimmer and Emlen 2013). Evolutionary theory explains that random mutations occur in organisms, and advantageous mutations for survival and reproduction would be selected and fixed, whereas disadvantageous mutations would be removed from a population, which in the long term results in adaptation and might be related to speciation (Darwin 1859). Naturally, in ecology and behavioral biology, questions mostly

focus on species-specific phenotypic traits and their adaptive functions.

Regarding the molecular level, functional changes in a protein have an immediate influence on phenotypic traits. Early studies in molecular evolution focused on comparing amino acid changes between species, and discovered unique properties of molecular evolution. Most amino acid substitutions in proteins do not change the function of the proteins substantially. The number of amino acid substitutions, *i.e.* amino acid changes that became fixed within a population, compared between two species was approximately proportional to the time since their divergence (Zuckerkandl and Pauling 1965; Margoliash 1963; Doolittle and Blomback 1964). Finally, such changes occurred less frequently in proteins which are functionally important, such as hemoglobins and cytochrome c, than in proteins which are less important, like fibrinopeptides (Margoliash and Smith 1965; Zuckerkandl and Pauling 1965). The results were suggesting that, on the molecular level, natural selection works on certain regions which have greater functional consequences and most likely an appreciable amount of substitutions is neutral, meaning that they do not have an influence at the phenotypic level.

This idea has been proposed under the name 'Neutral theory', stating that molecular evolution is mainly governed by random genetic drift, which means that most of the substitutions, or alleles that were replaced with different alleles in a population are neither selected for nor selected against (neutral mutations: Jukes and

Kimura 1984; Kimura 1968; King and Jukes 1969). This has created a niche for such neutral mutations (Figure 4) in evolutionary theory, beyond the concepts of advantageous and deleterious mutations. However, it was questioned by the selectionists or "neo-Darwinians" who viewed substitutions mostly as a consequence of positive selection (Simpson 1964; Mayr 1965), or of balancing selection (Mayr 1963; Ford 1964). By now, however, the Neutral theory is generally accepted as a fundamental concept in evolutionary genetics (Yoder et al. 2018; Jensen et al. 2019), and generally deemed to contribute to our understanding of natural selection, instead of opposing it, as it does not dismiss the principle

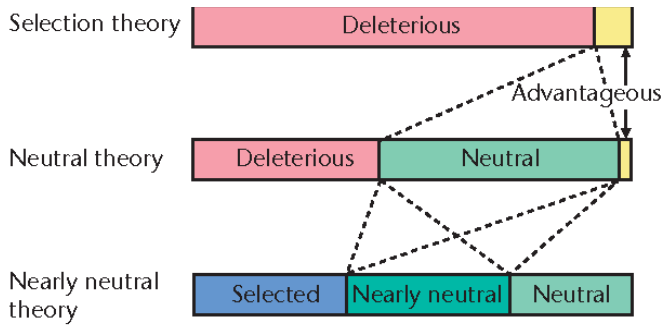


Figure 4. A representation of mutation classes in proportion across the theories: Selection theory, Neutral theory and Nearly neutral theory from top to bottom (Ohta 2013).

that the evolutionary change of phenotypic characters is primarily caused by new mutations on the molecular level (Nei 2005).

In the meantime, Tomoko Ohta has proposed another concept of mutation, 'nearly neutral' or 'slightly deleterious' mutations[1] (Nearly neutral theory: Ohta 1972, 1973) (Figure 4), which are more likely to be selected against than selected for. She pointed out the observations from previous molecular studies, namely that double fixation of mutations, *i.e.* concomitant amino acid substitutions in the same protein region, is observed more often than expected (Fitch and Margoliash 1967), and that evolution is more rapid in the paired region than in the non-paired region of tRNA (Dayhoff and McLaughlin 1972). She speculated that if the first mutation to occur is only slightly deleterious, rather than strictly neutral, the second could compensate it. This could be interpreted as another force in evolutionary change. On the other hand, Ohta has also argued that per-year rates of amino acid substitution are not equal across species and negatively correlate with the effective population size (N_e) of the species, which is the size of an idealized population that would give rise to the same variance of gene frequency or inbreeding rates as the actual population (Wright 1931; Crow and Kimura 1970; Caballero 1994). This was seemingly in contrast to expectations from the neutral theory (Ohta and Gillespie 1996). She explains it again by using the concept of slightly deleterious mutations, that is, if most amino acid changes are slightly deleterious, as their selection coefficient is assumed to be near the reciprocal of the N_e ($s \approx 1/N_e$), when N_e decreases, the

[1] I note that the concept of 'Nearly neutral' mutations refer to both 'slightly deleterious' and 'slightly advantageous' mutations, both of which have little fitness effect and been described with the expected selection coefficient, $|N_e s| < 1$ (Ohta 1972). In my study, I focus only on the 'slightly deleterious' mutations, and I use the term 'nearly neutral' in an interchangeable manner.

previously slightly deleterious change would behave neutrally, and when N_e increases, the same mutations would be under strong selection. In other words, it means that slightly deleterious mutations appear effectively neutral when N_e decreases (Charlesworth and Eyre-Walker 2007; Eyre-Walker and Keightley 2007; Nielsen and Yang 2003).

This suggested differential effects of purifying selection depending on the N_e of populations. If it was true, as slightly deleterious changes are expected to be relatively well tolerated (Ohta 1972, 1973; Ohta and Gillespie 1996), a larger proportion of such single nucleotide changes (SNCs) would be expected to reach high frequencies in a population under a relaxed purifying selection pressure given a small N_e than in one with the more efficient purifying selection given a large N_e. Moreover, it would be expected that such slightly deleterious changes should be observed more often in homozygous state in a population with small N_e, as such changes are expected to behave like neutral changes which are not effectively selected against (Ohta 1972, 1973). However, it was not straightforward to test this, as it was not trivial to identify the category of mutations which are 'slightly deleterious'. Although it has been known that the level of the purifying selection coefficient was expected to be different depending on the function and significance of a protein or the regions of the same protein (Ohta and Gillespie 1996; Subramanian and Lambert 2012), and for example, that almost half of coding SNCs in a single genome are inferred to be deleterious (Subramanian and Lambert 2012), it is still not easy to assess the deleteriousness of each possible mutation.

Several different approaches have devised ways to address this problem. For instance, the Grantham Score (Grantham 1974; Li, Wu and Luo 1984), is a rather straightforward measure to assess the deleteriousness of a single amino acid change. Based on the known physical and chemical properties of amino acid changes, it predicts how radical the amino acid change from A to B should be. SIFT (Sorting Tolerant From Intolerant; Kumar, Henikoff and Ng 2009) predicts the deleteriousness of amino acid changes, based on the degree of homology or conservation inferred from BLAST-based sequence alignments. According to the observation on how conserved or polymorphic a certain amino acid appears to be, it estimates the deleteriousness of the change at the residue of a protein. PolyPhen-2 (Polymorphism Phenotyping v2; Adzhubei et al. 2010), on the other hand, predicts the possible impact of amino acid changes on the structure and function using both physical property and multiple sequence alignments. There are also approaches to assess the deleteriousness of the SNCs in non-coding regions. The C-score (Kircher et al. 2014) is based on machine learning and prioritizes functional, deleterious, and pathogenic variants on a genome-wide scale including both coding and non-coding variants, using data such as allelic diversity, annotations of functionality, pathogenicity, severity of associated diseases, known regulatory effects from experiments, and complex trait associations, as training datasets. GWAVA (Ritchie et al. 2014) predicts the functional impact of genetic variants using a similar approach to C-score, but by using a number of different inputs. Finally, the GERP score (Genomic Evolutionary Rate Profiling scores; Davydov et al.

2010) compares, based on multiple alignments, the number of observed substitutions to the number of hypothetical substitutions under the assumption of neutral changes. It assumes that a deficit of observed substitutions as "Rejected Substitutions" serves as a measure of constraint on the element.

These methods have been employed in the endeavor to understand whether or not, or to which degree different demographic histories influence the efficacy of purifying selection in populations, which is still an ongoing debate in the field of population genetics. It has been tested in different species of animals and plants, including archaic and modern humans (Lohmueller et al. 2008; Castellano et al. 2014), dogs (Marsden et al. 2016), and rice populations (Liu et al. 2017). For example, Lohmueller et al. 2008, among other studies in human populations, clearly showed that, in populations with small N_e or population bottlenecks, genetic diversity is lower, which is reflected in the number of heterozygous genotypes per individual, compared to the populations that have maintained a larger N_e. Also, the rate of random fixation of deleterious derived alleles is higher, which is the proportion of fixed substitutions and number of homozygous derived alleles per individual.

In their study, they compare African American (AA) and European American (EA) populations, representing the populations with a large and a small N_e, respectively. Using the PolyPhen method, they show that the deleterious derived alleles are significantly more accumulated at homozygous positions in the EA population (Figure 5), which is in agreement with the findings in other human studies

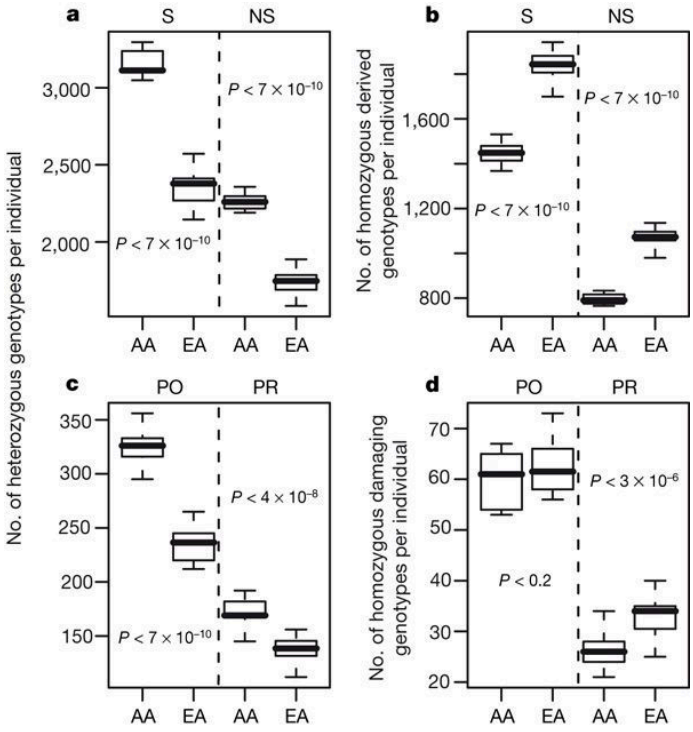


Figure 5. Distribution of the number of heterozygous and homozygous genotypes per individual in African American (AA) and European American (EA) populations. S: synonymous changes, NS: Non-synonymous changes, PO: Possibly damaging alleles and PR: Probably damaging alleles. Note that 'Probably damaging' is the category in PolyPhen which predict an allele change to be 'deleterious' (Lohmueller et al. 2008).

(Kidd et al. 2012; Torkamani et al. 2012; Hodgkinson et al. 2013). This is also highlighted in Henn et al. 2016, where they show a correlation of the numbers of such deleterious changes in homozygous state and the distance from Africa. Since the Out of Africa event, as humans continued to spread across continents, they are inferred to have experienced further and further population size

reductions and drift, so decreases in the efficiency of purifying selection would be expected to broadly form a continuum across populations from Africa to South America.

On the other hand, as mentioned earlier, more and more alleles become effectively neutral with lower N_e, and so they can accumulate in the population, reaching high frequencies. This has been tested in a comparative study of cultivated and wild types of rice species (Liu et al. 2017). They show, using the Site-Frequency-Spectrum (SFS) of synonymous and deleterious changes in different rice species, that the small N_e in the cultivated rice species is

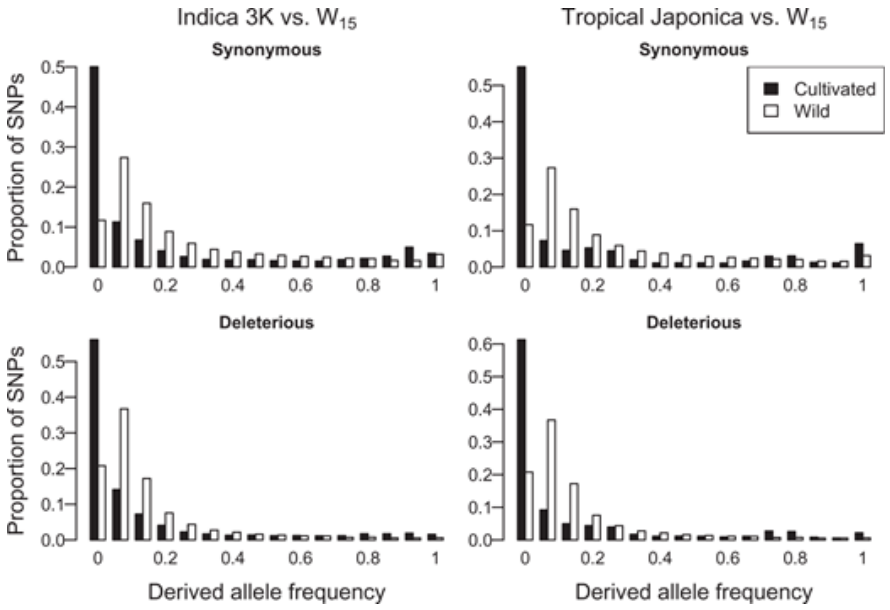


Figure 6. The Site-Frequencies-Spectrum for cultivated Asian rice species, Indica 3K (left) and Tropical Japonica (right), in comparison to their wild type counterpart W15, *Oryza rufipogon*, in synonymous (top row) and deleterious (bottom row) changes (Liu et al. 2017).

correlated with a much higher proportion of deleterious changes reaching high frequencies, compared to the wild species (Figure 6). This is expected for domesticated plants, as they experienced severe population bottlenecks due to strong selection by humans (Meyer and Purugganan 2013).

Bonobos and chimpanzees are interesting species to ask the same question, as their demographic histories since their divergence have been shown to be very different, as inferred from their genomes (Fischer et al. 2011; Prado-Martinez et al. 2013; De Manuel et al. 2016). Generally, chimpanzees are expected to have maintained a large N_e, reflected by the average number of heterozygous sites per genome or the number of Single Nucleotide Polymorphisms (SNPs), whereas bonobos had a smaller N_e with population bottlenecks (Prado-Martinez et al. 2013). However, it has been well demonstrated, among the four chimpanzee subspecies, that their demographic histories are markedly different. The central chimpanzees have experienced the largest N_e (24,400–48,700) and the western the smallest N_e (9,800–19,500), which is comparable to that of bonobos (11,900–23,800) (Prado-Martinez et al. 2013; De Manuel et al. 2016). For those reasons, bonobos and chimpanzees serve as interesting populations to test if their demographic histories would correlate with the differential level of purifying selection pressures.

Indeed, several studies have analyzed the efficacy of purifying selection in the *Pan* clade, however leaving an inconclusive picture.

By comparing the exomes of central, eastern and western chimpanzee populations, it has been shown that the efficacy of purifying selection correlates with N_e (Bataillon et al. 2015). Four different neutrality tests have been applied to detect signals of purifying, positive and balancing selection with different sensitivity to the evolutionary time windows, in the genomes of all the great ape lineages: the McDonald–Kreitman test, the Hudson–Kreitman–Aguadé test, the Extended Lineage Sorting test and Fay and Wu's H statistic, where the results pointed to the same conclusion that the N_e correlates with the efficacy of purifying selection (Cagan et al. 2016). On the other hand, (de Valles-Ibáñez et al. 2016) explored the same question in the great ape species, including only the eastern and Nigeria-Cameroon populations for chimpanzees, by comparing the load of loss-of-function (LoF) mutations, which probably have severe consequences. Their results suggested that, regarding the most severe type of variants, the efficacy of purifying selection does not vary depending on the N_e, which seemingly contradicts the previous studies. However, we should note that they do not address exactly the same question, as nearly deleterious mutation could be tolerated and appear neutral, whenever the threshold allows (Ohta 1972, 1973), and the strongly deleterious and damaging mutations with immediate functional consequences like LoF mutations are outside of this consideration.

Although the efficacy of purifying selection seems to correlate with N_e across all the great ape lineages, including the *Pan* species (Cagan et al. 2016), the accumulation of slightly deleterious mutations in all the chimpanzee and bonobo populations has not

been tested yet. It is important to include all the lineages, considering that the four chimpanzee subspecies have experienced variable N_e, with central and western chimpanzees being the largest and the smallest, respectively. Particularly, western chimpanzees are estimated to have a similar N_e as bonobos. The N_e of bonobos has been most likely small for a long time since they split from chimpanzees, possibly 1 to 2 Mya (Prado-Martinez et al. 2013), whereas the N_e of the western chimpanzees has been small as a consequence of a bottleneck after the split from the Nigeria-Cameroon chimpanzees about 250,000 years ago (De Manuel et al. 2016). Comparing the patterns of slightly deleterious mutations in these two populations is of particular interest, as it could inform not only about the efficacy of purifying selection but also about the potential purging effect in their populations.

In this study, I compare the accumulation of slightly deleterious allele changes across the four chimpanzee subspecies and bonobos, in order to infer the efficacy of purifying selection with regard to their N_e, with particular interest in the western chimpanzee and bonobo comparison. To do so, I exploit the six different deleteriousness measures for allele change diagnosis: the Grantham score, SIFT, PolyPhen-2, C-score, GWAVA and GERP scores. It has been reported that each method has different sensitivity in recognizing the deleterious variants depending on the dataset used (Mahmood et al. 2017). We decided to use the six measures, which take different and unique approaches in variant assessment, as explained above, as a way to confirm the robustness of the results.

1.3. Evidence of population substructure and divergence

Bonobos, chimpanzees and humans all belong to the order of Primates, which is divided into Haplorhini and Strepsirrhini (Groves 2001). Their divergence has been estimated to be around 74-76 Mya (Pozzi et al. 2014; Hirai, Imai and Go 2012). Afterwards, Strepsirrhini is divided into lemurs, pottos and lorises, which reside in Africa and Asia. Around 55 to 90 Mya, they most likely have radiated even in Europe and the Americas, where they went extinct later (Rose 2006). Haplorhini, on the other hand, are composed of the New World monkeys, the Old World monkeys, apes and tarsiers. By now, the New World monkeys inhabit South and Central America, after having diverged from the Old World monkeys and apes 33-40 Mya (Goodman et al. 1998; Nei and Glazko 2002). The Old World monkeys and apes are found in Asia and Africa, and are estimated to have appeared 25-30 Mya (Stevens et al. 2013). The African apes, chimpanzees, bonobos and gorillas and the Asian apes, gibbons and orangutans are the only extant non-human ape species, even though they are all recognized by now as endangered, if not critically endangered, by the IUCN red list (Romero Zarco 2018).

Such estimates in lineage branching were heavily based on paleontology and taxonomy. In the early 1960s, it was believed, based on such data, that humans branched out from a common ancestor with chimpanzees and gorillas roughly 20 Mya (Simpson 1963; Schultz 1966) (Figure 7), which was corrected later by

biochemical and molecular studies. One of the first approaches was estimating the divergence among primate species using protein albumin (Sarich and Wilson 1967). Even though it was still unclear

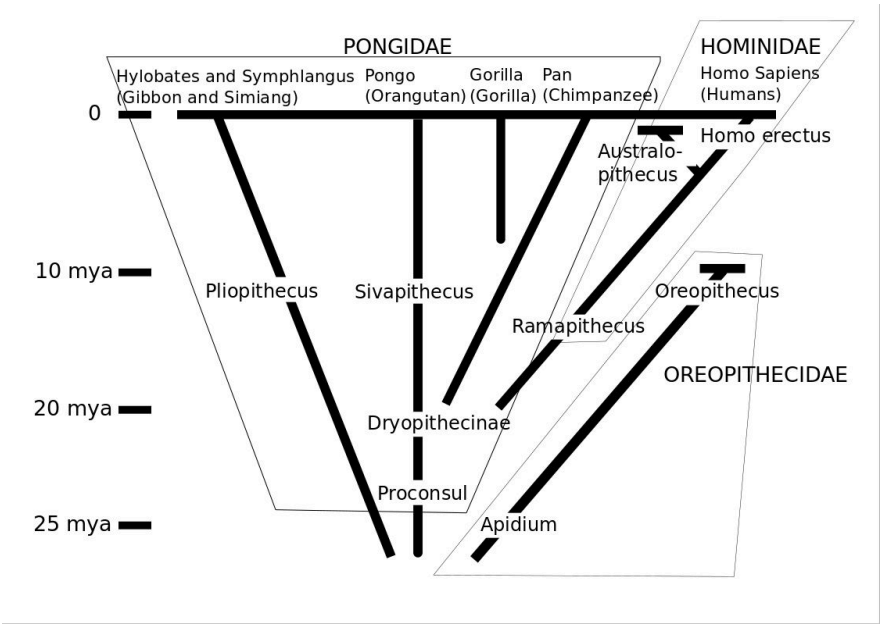


Figure 7. Classical hominoid family tree presented in (Kanō 1992), after the reports of Simpson 1963 and Schultz 1966, which is rejected nowadays.

whether chimpanzees or gorillas are more closely related to humans, their results implied that humans have diverged from those African apes around 5 Mya. With the technical development in molecular biology, it was revealed that chimpanzees are the most closely related sister species to humans (much closer than gorillas) (Hoyer et al. 1972), and based on comparative studies using nuclear and mitochondrial DNA, it has been estimated that the human-

chimpanzee divergence may have been as low as between 2 and 4 Mya, and their divergence from gorillas 3 to 6 Mya (Sarich and Wilson 1967; Hasegawa and Kishino 1984).

Thereafter, the astonishing speed of technological development in genetics was greatly improving the picture. New theories and tools in population genetics, together with a growing body of genome sequences, such as the first human genome in 2003 (Collins, Morgan and Patrinos 2003), the chimpanzee genome in 2005 (Sequencing and Consortium 2005), the bonobo genome in 2012 (Prüfer et al. 2012), and subsequent population-scale datasets such as the 1000 Human Genome Project (1000 Genomes Project Consortium et al. 2010), the Simon Human Diversity Panel

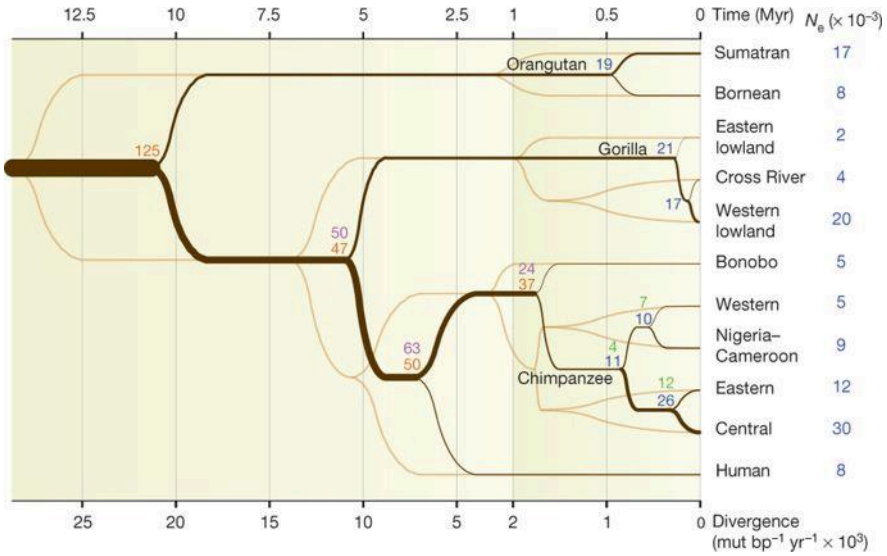


Figure 8. Inferred population history in the great apes: Population split time and N_e during the great ape evolution. Dark brown: split times and light brown: divergence times (Prado-Martinez et al. 2013).

(Mallick et al. 2016), and the Great Ape Genome Diversity Project in 2013 (Prado-Martinez et al. 2013), allowed us to better resolve their demographic histories (Figure 8), based on their molecular diversity.

For example, we know by now that humans and *Pan* lineages have diverged at least 6 Mya, and chimpanzees and bonobos 1 to 2 Mya. Chimpanzees have been previously recognized as four distinct subspecies, which are eastern (*Pan troglodytes schweinfurthii*), central (*Pan troglodytes troglodytes*), Nigeria-Cameroon (*Pan troglodytes ellioti*) and western chimpanzees (*Pan troglodytes verus*) (Groves 2001). Different repertoires or levels of cultural and social behaviors have been described for these populations, which has been interpreted within their environment as adaptation and diversity (Nishida 1968; van Lawick-Goodall 1968; Boesch and Boesch-Achermann 2000; Sugiyama and Koman 1987). Chimpanzee habitats are also known to range from tropical to woodland-savanna regions with varying climatic and environmental conditions, which was considered as an important factor driving the diversity across chimpanzee populations (Pruetz and Bertolani 2007; Poulsen and Clark 2004).

It is now clear from their genomes that the four populations are genetically distinct, having distinguished ancestries and differential demographic histories (Prado-Martinez et al. 2013). For example, it has been estimated that the first split into two separate branches in chimpanzees happened about 0.5 Mya, which was followed by further splits into two other lineages in each branch: Western and

Nigeria-Cameroon chimpanzees about 250,000 years ago, and central and eastern chimpanzees less than 200,000 years ago, respectively (Prado-Martinez et al. 2013; De Manuel et al. 2016). Among the four chimpanzee populations, central chimpanzees were inferred to have maintained the largest N_e, which is reflected in their higher genetic diversity, like heterozygosity, whereas western chimpanzees are supposed to have experienced severe genetic drift (Won and Hey 2005; Prado-Martinez et al. 2013; De Manuel et al. 2016), marked by the lowest genetic diversity comparable to Eastern lowland gorillas (Prado-Martinez et al. 2013).

Introgression between different lineages within the *Pan* clade, however, has not been very clearly comprehended. Won and Hey 2005, for example, reported a potential gene flow from the western to the central chimpanzees, based on their observations of 54 genomic loci, with no evidence of introgression between chimpanzees and bonobos. On the other hand, Becquet et al. 2007 reported later using 470 microsatellites an excess of derived allele sharing between bonobos and the central chimpanzees in comparison to between bonobos and the western chimpanzees, with no evidence of gene flow among chimpanzee lineages. However, they speculated that the observation might be due to the greater genetic drift in the western chimpanzees, rather than differential levels of introgression among them.

It is in fact not very surprising that the two studies appear opposing, considering the limited amount of data used in the analyses. A low level of genetic exchange between populations after a split, which

might be the case between bonobos and chimpanzees, and also between chimpanzee lineages, would need a sufficient amount of genome-wide segregating genotypes of high coverage, with which heterozygosity could be efficiently recognized. Prüfer et al. 2012 has made use of the first bonobo whole-genome in addressing the same issue, where they found no evidence of differential gene flow between bonobos and any of chimpanzee lineages, which was most likely due to the low-coverage genomes used in the analysis, as 19 additional bonobo and chimpanzee genomes of 1-fold coverage were included.

Indeed, SFS-based modelling, allele sharing and a haplotype-based local tree estimate approach, using 59 chimpanzee and 10 bonobo high-coverage (on average of 25-fold) whole-genomes, suggested at least one event of gene flow from bonobos into the ancestors of the central and the eastern chimpanzees, potentially followed by an additional gene flow event from bonobos into the central chimpanzees, along with bidirectional gene flow events between the central and the eastern chimpanzees and also between their ancestors and the Nigeria-Cameroon chimpanzees between 200 and 550 thousands of years ago (De Manuel et al. 2016) (Figure 9). The estimated migration rates from bonobos to chimpanzees were around 0.1 scaled by N_e, which clearly represents that multiple high-coverage whole-genomes had a greater power in detecting such a low level of gene flow between populations.

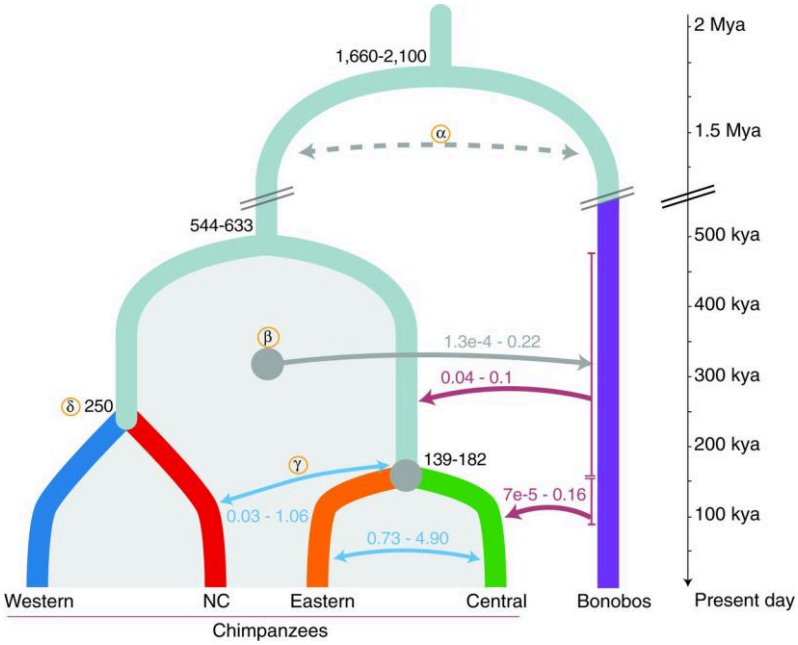


Figure 9. Conceptual model of a complex population history in *Pan* lineages. Split times (Kya) and migration rates correspond to 95% confidence intervals. Several introgression events between chimpanzees and bonobos after their divergence were inferred (De Manuel et al. 2016).

More recently, gene flow from an unknown extinct population, which seems to have diverged from the ancestral *Pan* population at least 3 Mya, into bonobos has been suggested (Kuhlwilm et al. 2019), using the same published data from De Manuel et al. 2016. The study made use of two different methods, *S** (Vernot and Akey 2014) and Skov HMM (Skov et al. 2018), which detects private alleles falling outside the internal variation or an unexpected density of private alleles in a given genetic segment of a given individual, respectively. The two methods estimated this migration to have occurred around 500 thousand years ago to an extent of 0.9 to 4.2%.

This study highlighted that the demographic history of the *Pan* clade is as much complicated as the one of *Homo*, involving introgression with extinct lineages (Prüfer et al. 2014), and also that high-coverage whole-genomes could shed a light on understanding the past history when fossil records are absent (Vernot and Akey 2014).

After these significant discoveries of divergence and introgression in the *Pan* clade, however, our understanding of divergence within bonobo populations is still limited. Bonobo groups in the wild have not been claimed so far to constitute uniquely distinct populations with differential cultures or behaviors, apart from differences in hunting and drumming behaviors (Hohmann and Fruth 2003). Bonobos were generally considered to reside in the rather homogeneous tropical region inside the current Democratic Republic of Congo (DRC) (Kanō 1992; Hashimoto et al. 1998; Serckx et al. 2014) and bonobo populations were often described as homogeneous in genetic studies (Fischer et al. 2011), marked by low levels of heterozygosity and high coefficients of inbreeding (Prado-Martinez et al. 2013). However, mitochondrial DNA (mtDNA) studies from seven wild bonobo populations suggested that there are at least three distinct clusters among them, which they described as eastern, central and western populations, according to their geographical origin (Kawamoto et al. 2013). Furthermore, it has not been pointed out in other publications, that substantial divergence between bonobo groups may exist, possibly comparable to that in chimpanzees, as implied in their phylogeny estimated from their mitochondrial DNA and some nuclear loci (Fischer et al.

2011) and several different ancestry components in 10 bonobo genomes (Prado-Martinez et al. 2013). This was probably not discussed further in those studies, due to the limitations of data, as only mtDNA and a few nuclear loci have been used in (Fischer et al. 2011) and only 10 individuals with unknown geographic origins were used in (Prado-Martinez et al. 2013). However, as the mtDNA haplotype tree suggests (Kawamoto et al. 2013), it is likely that distinct bonobo populations exist. The fact that some bonobos have adapted to the forest-savanna mosaic area (Serckx et al. 2014) also implies that bonobo populations might have differentially adapted to their unique environments. Understanding the extent of divergence in bonobos and their adaptations has a significance in recognizing features of the "prototype bonobo" and diverse characteristics in present-day bonobos (Boesch, Hohmann and Marchant 2002). One straightforward way to test it, on the genetic level, would be to sequence whole genomes of samples from wild populations, which is difficult. This is because samples like blood, the easiest material for sequencing, are often not available from the wild and other materials such as hair or fecal samples, which are easier to access, are not yet sufficient for genome-sequencing. Alternatively, it is promising to exploit published data to determine the extent of divergence in bonobos.

In this study, I make use of 20 published bonobo whole-exomes (Teixeira et al. 2015), which is the largest bonobo exome dataset to date, and explore genomic features to test whether wild-born bonobos have a clear genomic structure, allowing to group them as different populations. Based on the grouping, I estimate the extent

of divergence among them, together with published bonobo genomes representing each group (Prado-Martinez et al. 2013). For this analysis, the software G-PhoCS (Generalized Phylogenetic Coalescent Sampler; Gronau et al. 2011) was used.

2. OBJECTIVES

1. Can the remarkable differences in the physiology of exaggerated sexual swellings between bonobos and chimpanzees be associated to lineage-specific genetic changes with functional implications? If so, would it suggest on which lineage it has been a derived feature in comparison to their ancestral state?

2. How do the variations in demographic history across the four known chimpanzee populations, which are western, Nigeria-Cameroon, central and eastern chimpanzees, correlate with the behavior of slightly deleterious mutations, or the efficacy of purifying selection? How would bonobos fit into that, particularly in comparison to western chimpanzees?

3. Based on the genomic inferences, how much can we learn about divergence in the wild bonobo populations?

3. RESULTS

3.1. Genetic variation in *Pan* species is shaped by demographic history and harbors lineage-specific functions

Sojung Han, Aida M Andrés, Tomas Marques-Bonet, Martin Kuhlwilm

Genome Biology and Evolution. 2019

DOI: https://doi.org/10.1093/gbe/evz047

Han S, Andrés AM, Marques-Bonet T, Kuhlwilm M. Genetic Variation in Pan Species Is Shaped by Demographic History and Harbors Lineage-Specific Functions. Genome Biol Evol. 2019 Apr 1;11(4):1178–91. DOI: 10.1093/gbe/evz047

3.2. Divergence in bonobos: genomic evidence

Sojung Han*, Cesare de Filippo*, Genís Parra, Juan Ramon Meneu, Romain Laurent, Martin Kuhlwilm, Ilan Gronau, Svante Pääbo, Tomas Marques-Bonet, Aida M Andrés

in preparation

Divergence in bonobos: genomic evidence

Sojung Han*, Cesare de Filippo*, Genís Parra, Juan Ramon Meneu, Romain Laurent, Martin Kuhlwilm, Ilan Gronau, Svante Pääbo, Tomas Marques-Bonet, Aida M Andrés

UCL Genetics Institute, Department of Genetics, Evolution and Environment, University College London, London, UK.
Department of Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.
Institut de Biologia Evolutiva (Consejo Superior de Investigaciones Científicas–Universitat Pompeu Fabra), Barcelona, Spain.
Institucio Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.
CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain
Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Edifici ICTA-ICP, Barcelona, Spain.

Correspondence e-mail addresses:
Aida M Andrés (a.andres@ucl.ac.uk)

Highlights

- We identified three genetically distinct populations in 20 published bonobo exomes from individuals in African sanctuaries, two of which we infer to come from the west, and one of which from the central region within the bonobo habitat range.
- The first split time among the three populations is estimated at > 100,000 years ago, followed by the second split which is possibly tens of thousands of years ago.
- The three populations seem to have differential levels of genetic diversity, and one among them appears to have experienced strong genetic drift.

Summary

Bonobos are, together with chimpanzees, the closest species to humans. They are generally described to be egalitarian and peaceful marked by high female social status within their social network, in comparison to humans and chimpanzees, which has been intensively studied. However, other aspects of their biology and demography are less understood, probably due to their small census size (15,000–20,000 individuals by IUCN) and limited geographic range within the Democratic Republic of Congo (DRC). Unlike chimpanzees, of which there are four distinct recognized subspecies with various behavioral repertoire, there are no known subspecies of bonobo. Population level diversities in bonobos are not well understood, but it is probable that differential adaptation to unique environments might exist across bonobo populations. Early theories assumed bonobo habitat as rich tropical forests. However, it is known now that some bonobo populations live in forest-savanna mosaic regions. Bonobos were classically considered homogeneous at the genetic level, even though the same studies contained observations suggesting potential genetic diversity across populations. Nuclear genomic data is scarce and comes mostly from individuals of unknown geographic origin, which hinders the study of population substructure. Using 20 bonobo published exomes we show the presence of genetically distinct populations within bonobos, with substantially different levels of genetic diversity. Using mitochondrial DNA of individuals of known origin, we infer that these might be two western and one central populations within the DRC, which split at least 100,000 years ago, making their genetic differentiation comparable to that of modern human populations.

Keywords

Pan paniscus, Bonobo, divergence, population structure, diversity, exome, genome

Introduction

Bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*) are the closest extant species to humans. Chimpanzees are recognized to have at least four genetically and taxonomically distinct populations (Groves, 2001): eastern (*P. t. schweinfurthii*), central (*P. t. troglodytes*), Nigerian-Cameroon (*P. t. ellioti*) and western chimpanzees (*P. t. verus*). Their habitats range across a large area of central Africa with varying climatic and environmental conditions, from tropical to woodland-savanna areas (Boesch & Boesch-Achermann, 2000; Nishida, 1968; Poulsen & Clark, 2004; Pruetz & Bertolani, 2007; Sugiyama & Koman, 1987; Van Lawick-Goodall, 1968). Bonobos, in contrast, inhabit a smaller geographic area that spans only the current Democratic Republic of Congo (DRC) (Hashimoto et al., 1998; Kanō, 1992; Serckx et al., 2014).

Bonobos are generally described as social animals with an egalitarian and peaceful nature, in comparison to other primates (Idani, 1991; Kanō, 1992; Kuroda, 1980), where adult females have a high social status (Furuichi, 1987). Their hypersexuality, including homosexual interactions, has been described as the unique evolutionary feature of this species (Furuichi, 1987; Wrangham, 1993). Such features in bonobos were often understood as prototype features of the species, within their rather homogeneous tropical forest environment (Hashimoto et al., 1998; Kanō, 1992; Serckx et al., 2014), although it has been argued that diversity in bonobo is little understood, likely due to a paucity of data on bonobos (Stanford, 1998; Wrangham, 2002).

By now, accumulating observations suggest that there might exist a substantial diversity within bonobo populations in the wild. For example, it was known that there are differences in vegetation across bonobo habitats (Kanō, 1992). Later, it has been reported that some bonobo populations inhabit a forest-savanna mosaic area (Serckx et al., 2014; Thompson, 1997), suggesting a potentially unique adaptation of these populations to the environment. It was also reported that some differences in social and foraging behaviors exist across five different bonobo populations (Hohmann & Fruth, 2003).

Genetically, bonobos have been described to be homogenous as well (Fischer et al., 2011), marked by low heterozygosity and a high inbreeding coefficient (Prado-Martinez et al., 2013), although geographic differentiation has been observed. The mitochondrial DNA (mtDNA) phylogeny from seven wild bonobo populations across field sites in DRC suggested a structure of three genetic clusters, which were designated as eastern, central and western populations (Kawamoto et al., 2013). However, even though it has not been emphasized in those studies, the published analyses contained suggestive observations of substructure within bonobos: the phylogeny estimated for mtDNA and some nuclear loci of bonobos appeared to have a few distinct lineages, potentially

comparable to chimpanzee lineages (Fischer et al., 2011), and population structure analysis of ten bonobo genome showed several different ancestry component among them (Prado-Martinez et al., 2013). Such results imply a possible genomic substructure of distinct bonobo populations, which was not pointed out, probably due to limitations given the type and the size of the data. Fischer et al., 2011 have used 20 bonobo mtDNA sequences, which is not only small but also not sufficient for phylogeny inference (Hurst & Jiggins, 2005), and Prado-Martinez et al., 2013 have used only ten bonobos for whole-genome analyses.

Understanding bonobo divergence is relevant in evaluating differential natural selection across populations, which is possible, considering potential diversity in some bonobo populations regarding their habitat climate (Serckx et al., 2014; Thompson, 1997), vegetation (Kanō, 1992) and behaviors (Hohmann & Fruth, 2003). It would also help us better understand the genetic background in divergence and speciation (Mayr, 1942).

In this study, we make use of 20 bonobo whole exomes published in Teixeira et al., 2015, the largest bonobo exome dataset to date, to test whether wild-born bonobos show a clear genomic structure to group them as different populations. Combining this information with full genomes, we also estimate the split times among groups.

Results

Three genetically distinct populations of wild-born bonobos

A Principal Component Analysis (PCA) separates the 20 exomes of wild-born bonobos into three distinct groups (Figure 1A) that we will refer to as B1, B2 and B3. The same three groups are evident in a phylogenetic analysis (Neighbor-Joining tree, Figure 1B) and an ADMIXTURE analysis (Figure 1C) of the same data, confirming the presence of three genetically distinct units. In order to quantify the level of differentiation between the three groups, we measured pairwise F_{ST}, which estimates the differences in the allele frequency of single nuclear changes between pre-defined groups. We observed the highest differentiation between B1 and B3 (average F_{ST} = 0.145), while B2 shows lower differentiation with both B1 and B3 (F_{ST} = 0.093 and F_{ST} = 0.088, respectively, Table 1A), in agreement with the PCA results where B2 is placed in between of B1 and B3 on the PC1 (Figure 1A). In comparison, the F_{ST} between B1 and B3 falls within the range of the F_{ST} between chimpanzee subspecies (although it is much lower than the F_{ST} between *P.t. verus*

and the other subspecies, due to their extremely low *Ne*, Table 1B). On the other hand, the F_{ST} values for B1-B2 and B2-B3 fall within the range of the F_{ST} between African and non-African human populations (Table 1C). These results suggest that the genetic differentiation among the three bonobo populations might be similar to the level of central and eastern chimpanzees, which split around 200 kya (De Manuel et al., 2016), or to the level of African and non-African human populations.

Two populations seem to come from the western and one from the central part of the DRC

The 20 bonobo samples in our study come from an African sanctuary without known geographic origin. Previously, Kawamoto et al., 2013 and Takemoto et al., 2017 analyzed mtDNA data from seven bonobo wild populations and estimated their haplotype divergence. They defined three mtDNA clusters, western, central and eastern, based on the UPGMA tree built from the pairwise F_{ST}. As the genetic structure in our three bonobo populations could reflect geographic isolation, we combined the mtDNA sequences of our bonobo individuals, which were published in (Fischer et al., 2011), to place our bonobo samples within the context of the known mtDNA groups. B3 populations clustered with central mtDNA groups, which live in the center of the bonobo habitat range and have the highest mtDNA diversity among the three clusters (Figure 2). B1 and B2 populations, on the other hand, both clustered with the western mtDNA groups. Particularly, B1 seems to group tightly with the Malebo population, which is a western population in (Kawamoto et al., 2013; Takemoto et al., 2017), living at the periphery of the bonobo range, which has been reported to extend into the forest-savanna mosaic region (Pruetz & Bertolani, 2007; Thompson, 1997). Although we interpret that B1 and B2 populations are genetically distinguishable based on mtDNA, it is still not clear on which extent the two populations are genetically differentiated.

The split between the B1/B2 and B3 populations is estimated to be over 100kya ago

In order to estimate the potential split times between the three populations, we used the program G-PhoCS (Gronau, Hubisz, Gulko, Danko, & Siepel, 2011), which summarizes the information over local genealogies at neutral loci in approximate linkage equilibrium. Since exomes are far from ideal for demographic inferences, and neutral regions of the genome are required for this method (Gronau et al 2011), we added the exonic regions of the ten published high coverage bonobo genomes (Prado-Martinez et al., 2013) to the 20 exomes, and identified which individuals belong to each of the three clusters, based on a PCA (Figure S3). We then selected one individual representing each of the three bonobo populations, and determined neutral loci across the

non-coding fraction of their genome for the G-PhoCS analysis (Methods). We find that that the first split between the B1/B2 and the B3 populations was estimated around 100,000 years ago, and the two western populations split around 48,000 years ago (Figure 3). The split between bonobos and chimpanzees was inferred to be 1.5 Mya, while within chimpanzees we infer a first split ~560,000 years ago, the split between western and Nigeria-Cameroon ~200,000 years ago and that between the central and eastern chimpanzees ~220,000 years ago.

The putative western population B1 appears to have the lowest genetic diversity

The three distinct bonobo populations also show differences in their levels of genetic diversity. The B1 population has the lowest, whereas B3 has the highest genetic diversity (Figure S5). This agrees well with the limited information we have based on mtDNA (Kawamoto et al., 2013; Takemoto et al., 2017). The genetic diversity at the nuclear level in B3, expressed as the normalized Theta value (which is based on the number of polymorphic sites), is almost twice as high for B3 as for B1.

The low genetic diversity on both mtDNA and nuclear DNA levels in the B1 population indicates that B1 may have the smallest long-term effective population size (*Ne*), higher levels of recent inbreeding (the mating of closely related individuals), or both. To investigate this, we examined the distribution of runs of homozygosity (ROH), which are continuous segments depleted of heterozygous positions likely because the two chromosomes are derived from the same recent ancestor. As exome data are composed mostly of the genic part of the genome and cannot be used to infer with precision the ends of ROHs, we restricted the analysis to the comparison among individuals. We then compared the average ROH value in each bonobo group with the value of central chimpanzees and the value of Yoruba humans sequenced in the identical way (Teixeira et al., 2015). As a group, bonobos have on average longer ROHs than humans and chimpanzees (Figure S6), but among bonobos, B1 individuals have the longest ROHs, which are on average 29% longer than in B2 and B3 bonobos, 61% than in humans and 139% than in chimpanzees, which suggests that B1 have likely endured bottlenecks that are more severe than those experienced by other bonobo, humans, or central chimpanzee groups.

We further investigated the presence of inferred identity by state segments in our sample, which are identical chromosome fragments between two individuals that (when long enough) are best explained by the two chromosomes being inherited from the same recent common ancestor, that is, the segments are identical by descent. Thus, identity by descent (IBD) segments are indicative of

the presence of close genetic relationships between sampled individuals. While all bonobos in our sample have inferred IBD segments (shared predominantly with other members of their group) in comparison to chimpanzee and human (Figure S7), bonobos in B1 show more and longer IBD segments than bonobos from the other groups (Figure 3), which is another observation that B1 might have been isolated from other populations, possibly after a strong population bottleneck.

Recently, archaic admixture from an unknown ape lineage into bonobos has been described (Kuhlwilm et al. 2019). We find that coalescence times cannot be meaningfully calculated due to very sparse data, where vast parts of the genome would be considered continuous "internal" fractions between exon boundaries. This also leads to low inferred amounts of gene flow compared to whole genomes, here estimated at 0.14%. We find no significant differences in estimated admixture times or admixture fractions between the three groups of bonobos ($p > 0.05$, Wilcoxon rank test). An overall younger estimated admixture time of 294 kya, compared to ~500 kya from whole-genome data, is most likely due to sparse and discontinuous data. We also find no significant differences between groups for the amount of confidently called "archaic" sequence and for inferred introgression times from tract length.

Discussion

To which extent have bonobos diverged into genetically differentiated populations? As some studies report, bonobo populations might inhabit diverse geographic regions with different climates from tropical to forest-savanna mosaic area (Poulsen & Clark, 2004; Pruetz & Bertolani, 2007; Thompson, 1997), and with different types of vegetation (Kanō, 1992). It is probable that genetic diversity in bonobos might be as much diverse, which needs to be understood. In this study, we show that there are three genetically distinct populations (Figure 1). The average pairwise F_{ST} values point out that the genetic differentiation among the three bonobo populations might be comparable to that between central and eastern chimpanzee subspecies or between African and non-African human populations (Table 1).

In the phylogenetic analysis on the mtDNA haplotypes, the B1 population clusters with the Malebo group, which lives in the savanna-forest mosaic area in the western corner of the DRC (Poulsen & Clark, 2004; Pruetz & Bertolani, 2007). This is interesting, as it may imply that the genomic differentiation of the B1 population, reflected by the pairwise F_{ST} values (Table 1), suggests that this

population might have experienced stronger genetic drift. Based on the mtDNA haplotype analysis, it has been hypothesized that the bonobo ancestors have come to the Congo basin from the North, after which they have migrated across the DRC (Takemoto et al., 2017). In this scenario, the Malebo population would be the population expected to have experienced the most severe genetic drift. It has been shown in humans that the severity of population bottleneck signatures correlates with the distance from Africa to where the population comes from (Henn et al., 2016). The mutational load comparison across the three populations supports this, as the number of derived mutations at homozygous loci in the B1 population appears much higher than in B3 (Figure S8 and S9), although it is not significantly different compared to B2.

Our results from the G-PhoCS analyses suggest that the first split between the two western and the central populations happened at least 100,000 years ago, with the subsequent split between the two western populations, B1 and B2, more than 40,000 years ago. Considering the social structure of bonobos, where females migrate to a new group on sexual maturity, and that in primates admixture seems to be abundant (De Manuel et al., 2016; Kuhlwilm et al., 2016; Prüfer et al., 2012; Veeramah et al., 2015, Kuhlwilm et al., 2019), it is reasonable to assume that the three bonobo populations had some gene flow in between them, which in that case would shift the split time further back. Unfortunately, our analysis using migration scenarios failed to reach stable values, likely due to G-PhoCS failing to support one of two similarly likely specific models: a high migration rate with deeper split, or a small migration rate with more recent split (personal communication). The populations in the eastern corner of the DRC, which based on mtDNA seem the most diverged among the wild bonobo populations (Kawamoto et al., 2013; Takemoto et al., 2017) appear absent from our exome data. This suggests that there might be even deeper genetic substructure in bonobo populations, the extent of which needs to be understood.

Regarding genetic diversity, the B1 population appears to have the lowest levels. This is consistent across analyses: the normalised theta was lowest in B1 (Figure S5), the distribution of runs of homozygosity (ROH), showed that the average length of ROH fragments is highest in B1, (Figure S6) and the IBD analysis demonstrated that B1 have larger proportion of the genomic chunks, which are supposed to come from recent related shared ancestors, suggesting that they might have experienced severe bottlenecks (Figure 3). These are in agreement with the discussion earlier, that B1 might have gone through population bottlenecks resulting from the bonobo migration history, which might prioritize them in the conservation efforts, although all the bonobo populations in the

wild are recognized to be endangered (Romero Zarco, 2018).

In this study, we have identified genetically distinct bonobo populations, with the estimate that their first split might be at least 100,000 years ago. These populations seem genetically differentiated comparable to African and non-African humans. We propose that sampling a wider range of bonobo populations in the wild, in particular including the eastern populations, would allow us to better comprehend their divergences and interaction in the past, and that, although we do not have enough data yet from the wild, we might get to know more in the future about behavioral and adaptive diversity across bonobo populations in the wild.

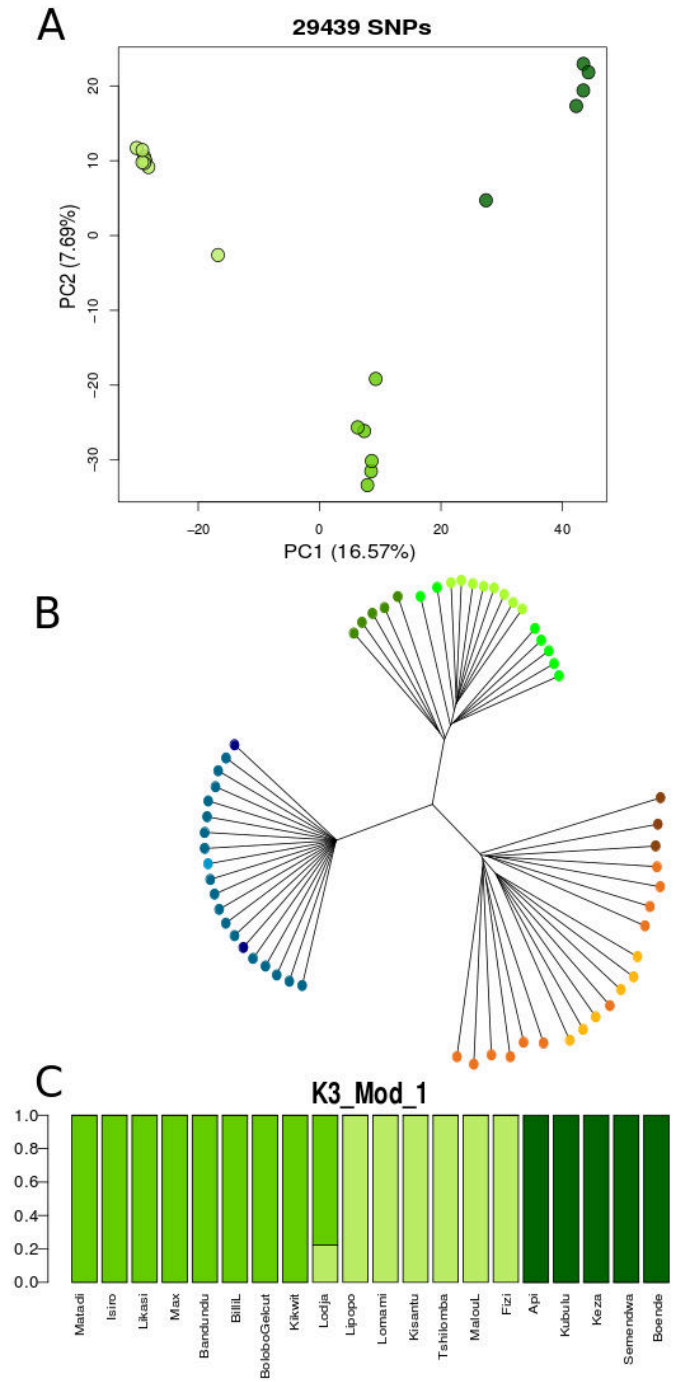


Figure 1. Separation of bonobos into three distinct groups. A) PCA, B) Neighbor-joining tree, with colors of groups according to PCA, and C) ADMIXTURE of three species: bonobo (k=3).

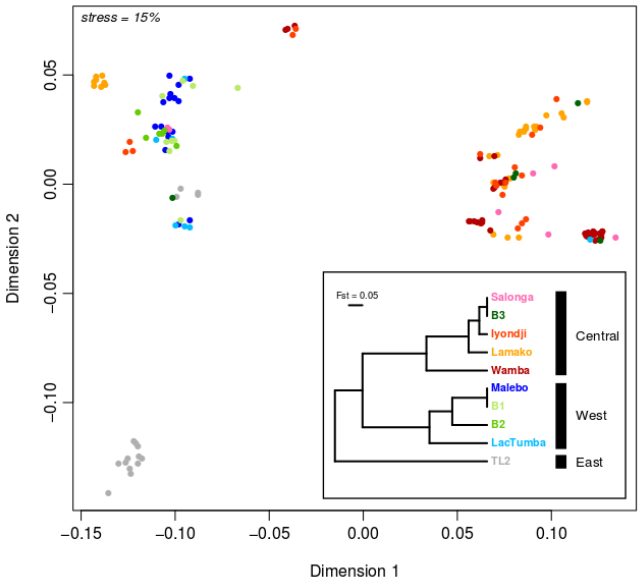


Figure 2. Multidimensional scaling (MDS) and UPGMA tree of bonobo individuals combined from (Kawamoto et al., 2013) and (Fischer et al., 2011).

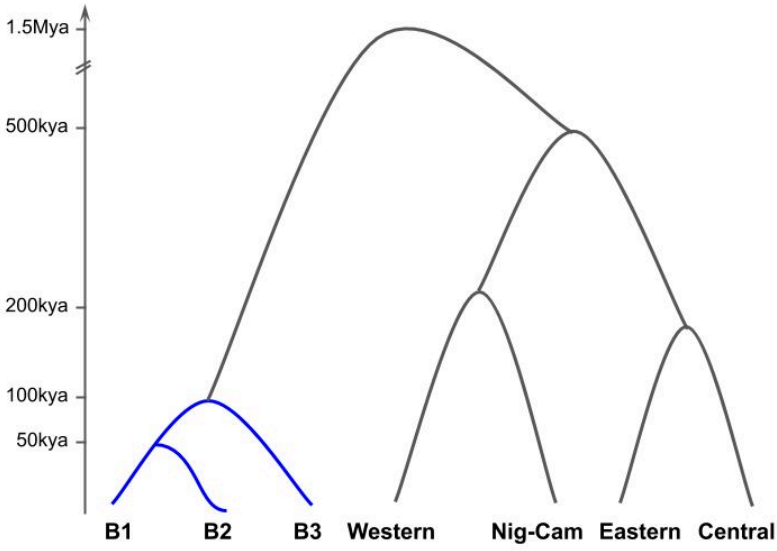


Figure 3. Conceptual model of split times in the *Pan* clade, estimated by G-PhoCS.

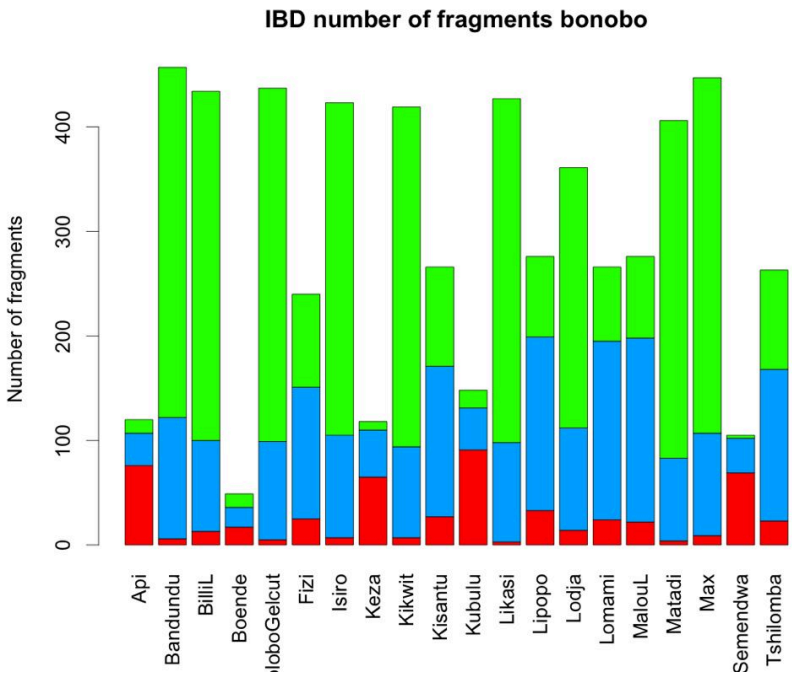


Figure 4. IBD analysis based on the grouping of the bonobos. Total length of IBD fragments per individual, normalized by the maximum length observed in the sample.

Table 1. Fst values across groups. A) Bonobo and chimpanzee groups in our exome data, B) chimpanzee subspecies (Prado-Martinez et al., 2013) and C) three continental comparisons in humans using the 1000 Genomes dataset. *We consider two populations per continent, so we present one Fst value within continent, and the average Fst of four pairwise population comparisons between continents.

A.

Bonobo	B1	B2	B3
B1	-		
B2	0.093	-	
B3	0.145	0.088	-

B.

Chimpanzee	Elioti	Schweinfurthii	Troglodytes	Verus
Elioti	-			
Schweinfurthii	0.163	-		
Troglodytes	0.166	0.122	(0.059)	
Verus	0.18	0.234	0.227	-

C.

Human	Africa	Europe	Asia
Africa	0.01		
Europe	0.1	0.015	
Asia	0.116	0.079	0.1

Methods

Data Preparation

We analyzed whole-exome high-coverage (~20X) Illumina sequencing data of 20 humans, 20 central chimpanzees (*Pan troglodytes troglodytes*), and 20 bonobos (*Pan paniscus*) published in (Teixeira et al., 2015). Human samples belong to the Yoruba population from HapMap; bonobo and chimpanzee blood samples were collected in African sanctuaries (Lola ya bonobo sanctuary in Kinshasa, Democratic Republic Congo; and Tchimpounga sanctuary, Jane Goodall Institute, Republic of Congo, respectively).

Base calling was performed with Ibis (Teixeira et al., 2015), and reads with more than 5 bases with a base quality score lower than 15 were discarded. Reads were aligned to the human reference genome hg19 using BWA with default parameters. Mapping all individuals to the same reference genome prevented complications from mapping to genomes of different quality. Only reads with a mapping quality (MQ) $\geq$ 25 and mapping outside of known segmental duplications in the three species were considered for further analysis. Specifically, the average coverage for each individual is 18.9× in human, 17.9× in chimp, and 17.9× in bonobo.

Data Analysis

Principal Components Analyses (PCA) were performed using the R-function 'glPca' from the package adegenet (Jombart, 2008) and run for all individuals and separately per species. Neighbor-Joining (NJ) trees were also constructed for all individuals together, and separately per species. The distance matrix is counts of pairwise nucleotide differences among individuals. The distance between an heterozygote and any homozygote is 1, and the distance between the two different homozygotes is 2. The number of differences were summed across all autosomal sites (SNPs and/or fixed differences) per species: 228,488 for chimp, 86,250 for bonobo, and 106,832 for Yoruba.

To further test the substructure observed within each species on the PCA and the phylogenetic trees, we performed structure analyses. We performed the analyses on a subset of SNPs that minimizes linkage disequilibrium (LD). In order to do this, SNPs in LD were removed using the software plink (Purcell et al., 2007) with the following steps:

1. create a window of 200 SNPs

2. calculate LD (as r^2) between each pair of SNPs in the window

3. if $r^2 > 0.5$ remove one of a pair of SNPs

4. shift the window 20 SNPs and repeat the procedure.

Afterwards, we ran ADMIXTURE (Alexander, Novembre, & Lange, 2009) from K (number of clusters)=1 to K=8, each with 10 replicates using the following command line: "admixture -s time --cv INPUT.ped k". We run the cross-validation procedure ("--cv" flag) as described in (Alexander et al., 2009) to determine the number of K that best fits the data. The software CLUMPP (Jakobsson & Rosenberg, 2007) was used to condense the 10 admixture runs per K in order to identify modes where different runs have similar outcomes (>90%). This was done by selecting pairs of replicates having a symmetric similarity coefficient G' > 0.9.

As a measurement of population differentiation, we calculated the average F_{ST} across all sites that are polymorphic in at least one population of each population pair. F_{ST} for each site was calculated according to the formula of (Weir & Cockerham, 1984). The pairwise comparisons are: 1) among bonobo groups using our exome data; 2) among all chimpanzee subspecies using the GAGDP dataset, and among two central chimpanzee subgroups identified with our exome data; 3) among six human populations from Africa (Yoruba and Luhya), Europe (Toscani and Finns) and Asia (Han Chinese and Japanese) using the 1000 Genomes data (1000 Genomes Project Consortium et al., 2010).

Kawamoto et al., 2013 analyzed the D-loop of the mitochondrial DNA (mt) of 136 bonobo's individuals and found that 83% of the 54 unique haplotypes were specific to the seven sampled locations. This allows us to geographically place our 20 bonobos with reasonable confidence. To do so, we used the previously published complete mitochondrial genomes (Fischer et al., 2011) of our 20 bonobos. The 156 mtDNA sequences were aligned with the software mafft v7 (Katoh & Standley, 2013), and then we removed positions with indels and missing data retaining a total of 1,101bp. We calculated pairwise difference between sequences using the Kimura 2-parameters model (Kimura, 1980) and performed multidimensional scaling (MDS) on these distances using the R-function 'cmdscale'. In order to visualize all samples, we randomly added noise as half of the variance for each dimension of the MDS. We then calculated Φ_{ST} among 10 bonobo groups according to the formula of (Michalakis & Excoffier, 1996), which then was used for building UPGMA tree. For plotting, we used the R-function 'upgma' of package phangorn with default parameters.

In order to infer the divergence time among the three bonobo populations, B1, B2 and B3, we used the Generalized Phylogenetic Coalescent Sampler (G-PhoCS; Gronau et al., 2011), which is a Bayesian sampling strategy summarizing the information over local genealogies at short, putatively neutral loci in approximate linkage equilibrium. Our analysis followed the guidelines described in previous studies (Gronau et al., 2011; Kuhlwilm et al., 2016; Schlebusch et al., 2017). Eight filters were downloaded from the UCSC genome annotation database for hg19 (http://hgdownload.cse.ucsc.edu/goldenpath/hg19/database/, the last date of access 29/04/2019), which targets known genic regions (refGene, knownGene), simple and complex repeat regions (simpleRepeat, genomicSuperDups), CpG islands (cpgIslandExt), repeat masker (rmsk), conserved regions across 46 placental species (phastConsElements46wayPlacental), and synteny net between the assemblies hg19 and PanTro4 (netPanTro4). The length of the fragments was set to 1kb, which is a length of the optimal trade-off between minimizing the impact of recombination and maximizing information for coalescence analysis in human genomes (Gronau et al., 2011).

Each fragment that, after applying the filters, contained less than 20% missing data within each individual was considered for the analysis. After applying the filters, the loci were used 1) as a whole and 2) by randomly choosing every third fragment in two batches, following the recommendation from the author of the software to try out different filters. The exact number of fragments used in the G-PhoCS run was 32,569 fragments. In each, the MCMC was run for 1,000,000 iterations. To calculate mean and median split times (Tau), the results from the last 200,000 runs were used, after removing the first 800,000 runs as burnins. Those values are presented as a table of mean split times with standard deviations (Table S1). To convert Tau to calendar years, the human mutation rate of 0.43 x 10-9 per site per generation was used (Besenbacher, Hvilsom, Marques-Bonet, Mailund, & Schierup, 2019).

G-PhoCS requires the expected divergence nodes, based on which it could calculate the rates of split times and effective population sizes. We have used the UPGMA tree from their mtDNA haplotypes (Figure 2) for the divergence among the three bonobo populations, and added four chimpanzee individuals representing the four known chimpanzee subspecies, for validation (Figure S4).

Identity-by-state (IBS) can be observed at a given locus, for any given pair of individuals with genotype information, with three possible outcomes: the individuals have two different alleles (IBS0) or they share one (IBS1) or two (IBS2) alleles in common. Two individuals who share 1 or 2

alleles IBS at a given locus may have inherited the shared alleles from a recent common ancestor, in which case these alleles are identical-by-descent (IBD). IBD regions tend to be short between pairs of individuals derived from a given population that are not closely related, primarily because their last common ancestor was many generations ago; they tend to be long among closely related individuals. Using the set of informative nucleotide positions for the three species, we computed the number and length of the IBD regions (defined here as regions between IBS0 alleles) among all the individuals.

In order to analyze patterns of such an archaic admixture in our bonobos, we analyzed the exome data using a method for detection of gene flow without source genomes (Skov et al. 2018). We intersected the genotypes of 20 bonobos with the genotypes at the same genomic coordinates of 30 central and western chimpanzees from whole genomes (de Manuel et al. 2016). We generated exome-specific files of private alleles, local mutation rates and genomic coverage, analogous to the procedures described in detail in (Kuhlwilm et al. 2019) and (Skov et al. 2018). We then calculated parameters as described in these previous studies.

References

1000 Genomes Project Consortium, Abecasis, Altshuler, Auton, Brooks, Durbin, … McVean. (2010). A map of human genome variation from population-scale sequencing. *Nature*, *467*(7319), 1061–1073.

Alexander, Novembre, & Lange. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664.

Besenbacher, Hvilsom, Marques-Bonet, Mailund, & Schierup. (2019). Direct estimation of mutations in great apes reconciles phylogenetic dating. *Nature Ecology & Evolution*, *3*(2), 286–292.

Boesch, & Boesch-Achermann. (2000). *The chimpanzees of the Taï Forest : Behavioural ecology and evolution.* Oxford University Press.

De Manuel, Kuhlwilm, Frandsen, Sousa, Desai, Prado-Martinez, … Marques-Bonet. (2016). Chimpanzee genomic diversity reveals ancient admixture with bonobos. *Science*, *354*(6311).

Fischer, Prüfer, Good, Halbwax, Wiebe, André, … Pääbo. (2011). Bonobos fall within the genomic variation of chimpanzees. *PLoS ONE*, *6*(6).

Furuichi. (1987). Sexual swelling, receptivity, and grouping of wild pygmy chimpanzee females at Wamba, Zaïre. *Primates*, *28*(3), 309–318.

Gronau, Hubisz, Gulko, Danko, & Siepel. (2011). Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, *43*(10), 1031–1034.

Groves. (2001). *Primate taxonomy.* Smithsonian Institution Press.

Hashimoto, Tashiro, Kimura, Enomoto, Ingmanson, Idani, & Furuichi. (1998). Habitat use and ranging of wild Bonobos (*Pan paniscus*) at Wamba. *International Journal of Primatology*, *19*(6), 1045–1060.

Henn, Botigué, Peischl, Dupanloup, Lipatov, Maples, … Bustamante. (2016). Distance from sub-Saharan Africa predicts mutational load in diverse human genomes. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(4).

Hohmann, & Fruth. (2003). Culture in Bonobos? Between-species and within-species variation in behavior. *Current Anthropology*, *44*(4), 563–571.

Hurst, & Jiggins. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1572), 1525–1534.

Idani. (1991). Social relationships between immigrant and resident Bonobo (*Pan paniscus*) females at Wamba. *Folia Primatologica*, *57*(2), 83–95.

Jakobsson, & Rosenberg. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801–1806.

Jombart. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405.

Kanō. (1992). *The last ape : pygmy chimpanzee behavior and ecology.* Stanford University Press.

Katoh, & Standley. (2013). MAFFT Multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780.

Kawamoto, Takemoto, Higuchi, Sakamaki, Hart, Hart, … Furuichi. (2013). Genetic structure of wild bonobo populations: Diversity of mitochondrial DNA and geographical distribution. *PLoS ONE*, *8*(3).

Kimura. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, *16*(2), 111–120.

Kuhlwilm, Gronau, Hubisz, de Filippo, Prado-Martinez, Kircher, … Castellano. (2016). Ancient gene flow from early modern humans into Eastern Neanderthals. *Nature*, *530*(7591), 429–433.

Kuhlwilm, Han, Sousa, Excoffier and Marques-Bonet. (2019). Ancient Admixture from an Extinct Ape Lineage into Bonobos. *Nature Ecology & Evolution*, 3(6):957-965.

Kuroda. (1980). Social behavior of the pygmy chimpanzees. *Primates*, *21*(2), 181–197.

Mayr. (1942). *Systematics and the origin of species.* Harvard University Press.

Michalakis, & Excoffier. (1996). A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics*, *142*(3), 1061–1064.

Nishida. (1968). The social group of wild chimpanzees in the Mahali Mountains. *Primates*, *9*(3), 167–224.

Poulsen, & Clark. (2004). Densities, distributions, and seasonal movements of gorillas and chimpanzees in swamp forest in northern Congo. *International Journal of Primatology*, *25*(2), 285–306.

Prado-Martinez, Sudmant, Kidd, Li, Kelley, Lorente-Galdos, … Marques-Bonet. (2013). Great ape genetic diversity and population history. *Nature*, *499*(7459), 471–475.

Pruetz, & Bertolani. (2007). Savanna chimpanzees, *Pan troglodytes verus*, hunt with tools. *Current Biology*, *17*(5), 412–417.

Prüfer, Munch, Hellmann, Akagi, Miller, Walenz, … Pääbo. (2012). The bonobo genome compared with the

chimpanzee and human genomes. *Nature*, *486*(7404), 527–531.

Purcell, Neale, Todd-Brown, Thomas, Ferreira, Bender, … Sham. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575.

Romero Zarco. (2018). *The IUCN red list of threatened species 2018.* International Union for Conservation of Nature - IUCN.

Schlebusch, Malmström, Günther, Sjödin, Coutinho, Edlund, … Jakobsson. (2017). Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago. *Science*.

Serckx, Huynen, Bastin, Hambuckers, Beudels-Jamar, Vimond, … Kühl. (2014). Nest grouping patterns of bonobos (*Pan paniscus*) in relation to fruit availability in a forest-savannah mosaic. *PLoS ONE*, *9*(4).

Skov, L. et al. 2018. "Detecting archaic introgression using an unadmixed outgroup." *PLoS Genet.* 14, e1007641.

Stanford. (1998). The social behavior of chimpanzees and bonobos: Empirical evidence and shifting assumptions. *Current Anthropology*, *39*(4), 399–420.

Sugiyama, & Koman. (1987). A preliminary list of chimpanzees' alimentation at Bossou, Guinea. *Primates*, *28*(1), 133–147.

Takemoto, Kawamoto, Higuchi, Makinose, Hart, Hart, … Furuichi. (2017). The mitochondrial ancestor of bonobos and the origin of their major haplogroups. *PloS One*, *12*(5).

Teixeira, de Filippo, Weihmann, Meneu, Racimo, Dannemann, … Andres. (2015). Long-term balancing selection in *LAD1* maintains a missense trans-species polymorphism in humans, chimpanzees, and bonobos. *Molecular Biology and Evolution*, *32*(5), 1186–1196.

Thompson. (1997). *The history, taxonomy and ecology of the bonobo (Pan Paniscus Schwarz, 1929) with a first description of a wild population living in a forest/savanna mosaic habitat.*

Van Lawick-Goodall. (1968). The behaviour of free-living chimpanzees in the Gombe stream reserve. *Animal Behaviour Monographs*, *1*, 161-IN12.

Veeramah, Woerner, Johnstone, Gut, Gut, Marques-Bonet, … Hammer. (2015). Examining phylogenetic relationships among gibbon genera using whole genome sequence data using an approximate bayesian computation approach. *Genetics*, *200*(1), 295–308.

Weir, & Cockerham. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*(6), 1358.

Wrangham. (1993). The evolution of sexuality in chimpanzees and bonobos. *Human Nature*, *4*(1), 47–79.

Wrangham. (2002). The cost of sexual attraction: is there a trade-off in female Pan between sex appeal and received coercion? In Boesch & Hohmann (Eds.), *Behavioural Diversity in Chimpanzees and Bonobos* (pp. 204–216).

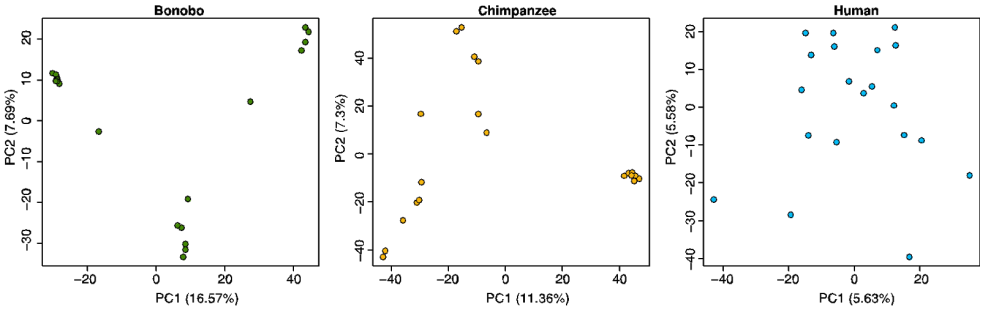
Supplementary

Sojung Han*, Cesare de Filippo*, Genís Parra, Juan Ramon Meneu, Romain Laurent, Martin Kuhlwilm, Ilan Gronau, Svante Pääbo, Tomas Marques-Bonet, Aida M Andrés

UCL Genetics Institute, Department of Genetics, Evolution and Environment, University College London, London, UK.
Department of Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.
Institut de Biologia Evolutiva (Consejo Superior de Investigaciones Científicas–Universitat Pompeu Fabra), Barcelona, Spain.
Institucio Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.
CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain
Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Edifici ICTA-ICP, Barcelona, Spain.

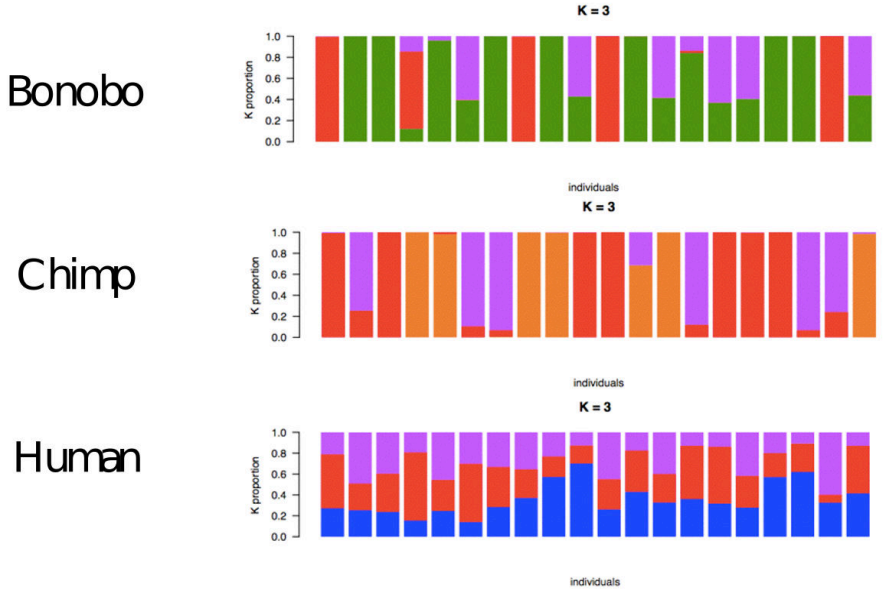
Correspondence e-mail addresses:
Aida M Andrés (a.andres@ucl.ac.uk)

Figure S1



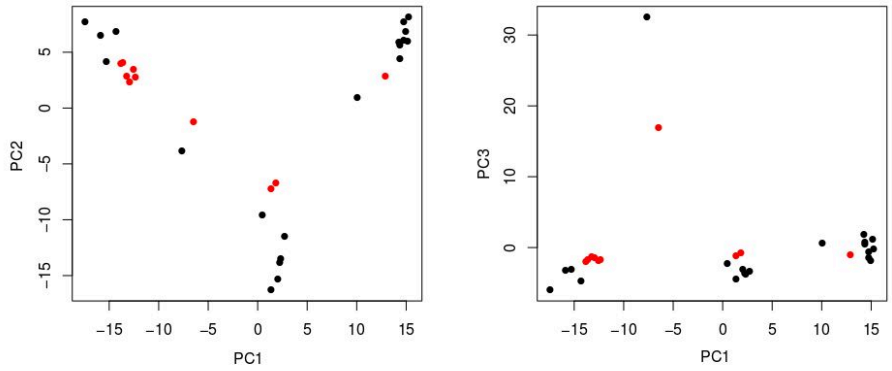
PCA results of the analysis of each species separately. We observe three groups in bonobo, two groups in chimpanzee, and no grouping in humans.

Figure S2



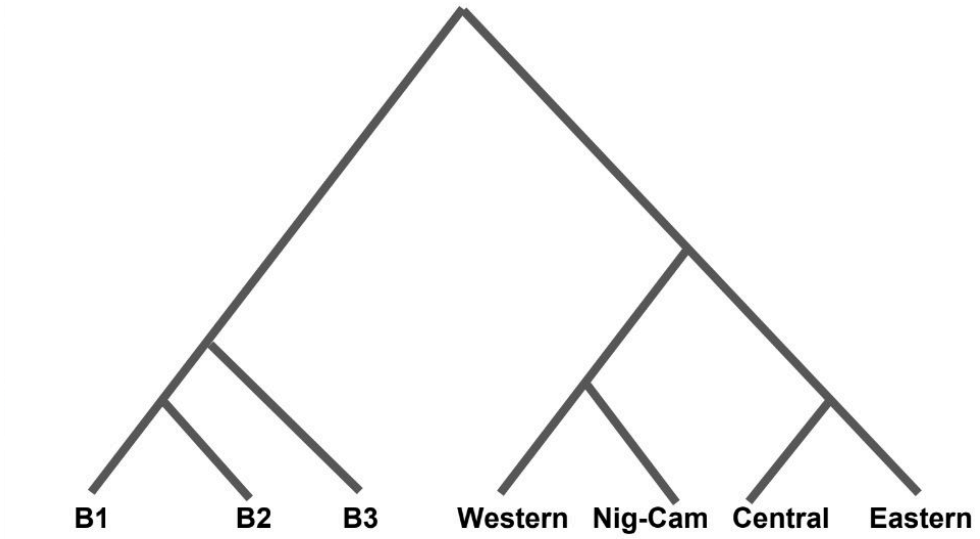
STRUCTURE results for bonobo, chimpanzee and human, K=3.

Figure S3



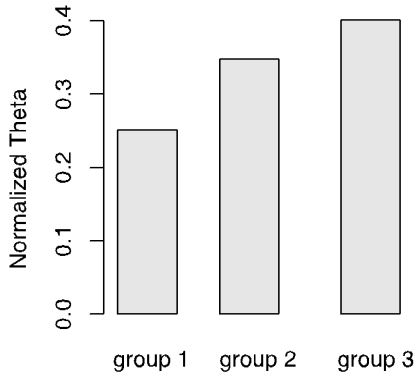
PCA, exomes and genomes together

Figure S4



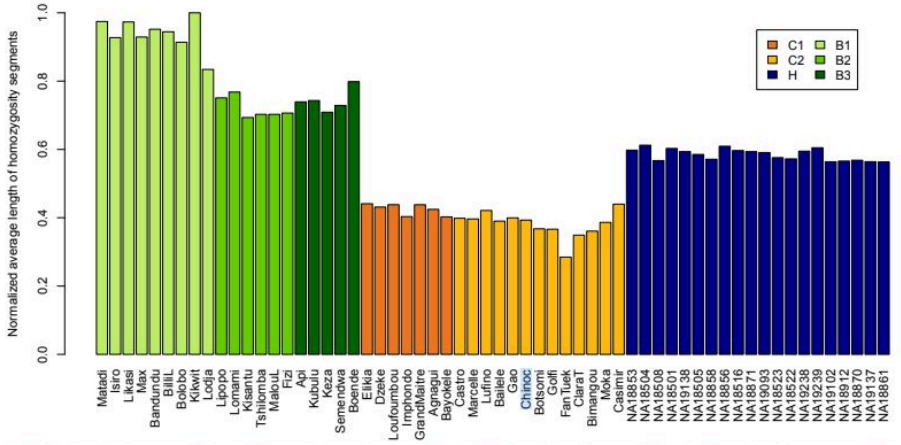
Schematic drawing of the tree used for G-PhoCS.

Figure S5



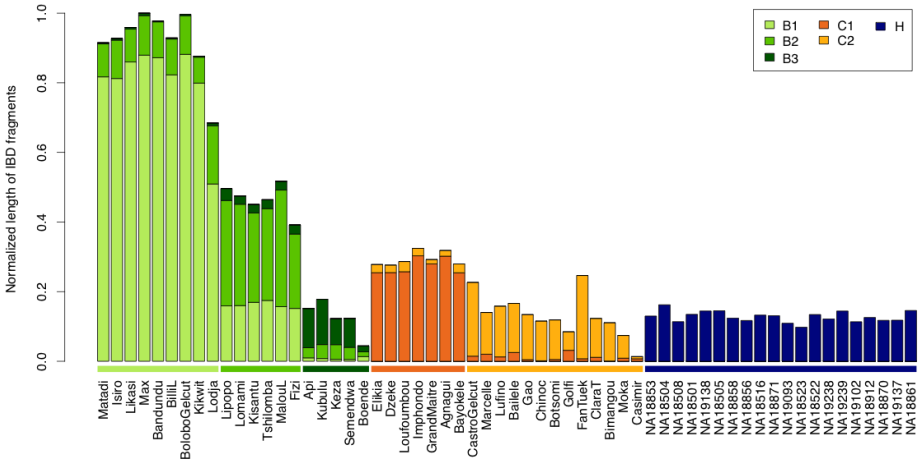
Normalized diversity levels of the different bonobo groups. Theta S {Watterson}, based on the number of segregating sites.

Figure S6



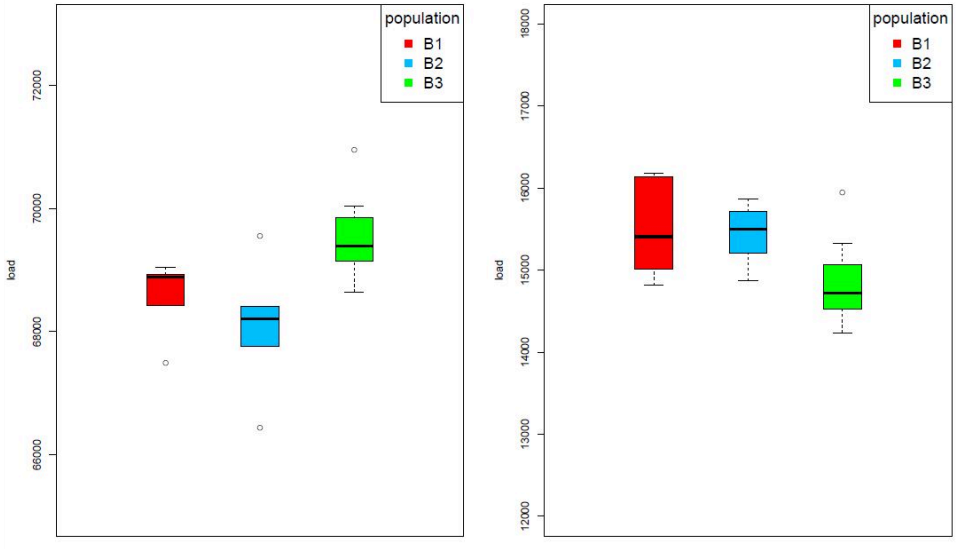
Normalized average length of homozygosity segment in bonobos, chimpanzees and humans.

Figure S7



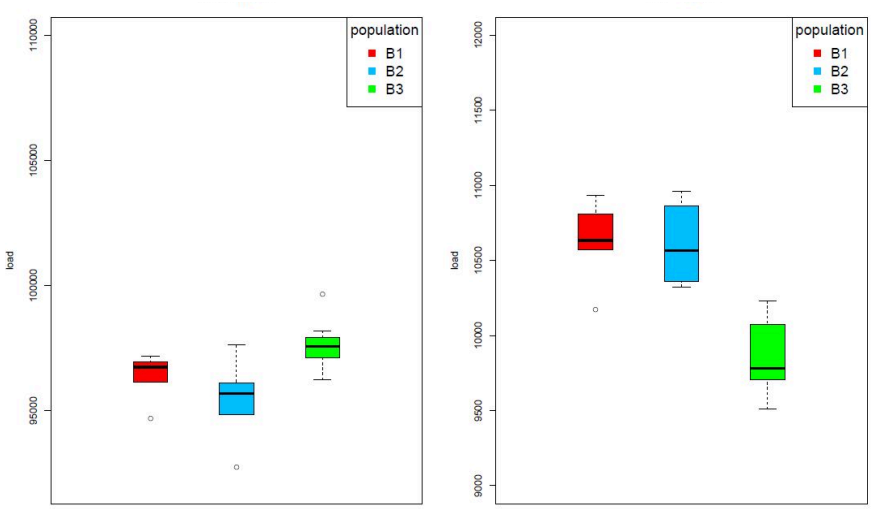
Total length of IBD fragments per individual, normalized by the maximum length observed in the sample. Each individual is represented in a vertical bar, sorted by species and by number of singletons. Each segment of the vertical bar represents the length of IBD fragments shared between this individual and any individual belonging to the group of the corresponding color. Horizontal bars represent the group each individual belongs to.

Figure S8



Mutational load, which is a count of loci with derived segregating alleles, as heterozygote (left) and as homozygote (right) among missense mutations.

Figure S9



Mutational load, which is a count of loci with derived segregating alleles, as heterozygote (left) and as homozygote (right) among synonymous mutations.

Table S1. Split time estimates (tau estimates scaled by mutation rate) from G-PhoCS.

	B1/B2	B12/B3	W/NC	C/E	chimpanzee	Pan
10k	93,023 (0)	93,023 (0)	209,253 (1,511)	223,968 (11,263)	557,799 (3,403)	1,555,541 (17,377)
10k_sub1	43,316 (8,124)	99,598 (10,473)	214,394 (9,689)	220,862 (12,432)	554,708 (9,261)	1,547,036 (17,257)
10k_sub2	51,872 (9,795)	94,493 (5,660)	209,285 (6,138)	222,785 (12,218)	564,119 (10,304)	1,548,102 (16,409)
5k	46,747 (2,329)	93,110 (1,417)	209,245 (1,206)	230,744 (6,829)	558,195 (1,157)	1,569,032 (17,798)
5k_sub1	160,150 (16,831)	160,188 (16,825)	209,235 (5,302)	230,231 (10,100)	565,788 (10,926)	1,417,662 (32,076)
5k_sub2	47,952 (5,605)	93,566 (3,510)	201,752 (10,916)	212,858 (10,713)	564,671 (10,461)	1,546,170 (20,039)

4. DISCUSSION

4.1. Phenotypic traits in female reproduction

Bonobos are an interesting species for evolutionary biology. Their frequent sexual relationships, including homosexual interactions (de Waal 1987; Wrangham 1993), the notably peaceful and egalitarian social dynamics (Kanō 1992; Furuichi 2011), even in intergroup encounters (Idani 1990), and the high social status of females and their alliances mark their unique niche in our understanding of primates (Thompson-Handler, Malenky, and Badrian 1984; Wrangham 1993), including ourselves. A valid understanding of bonobo evolution is facilitated by a comparison to other closely related species, such as chimpanzees and humans, not only considering their behavior and ecology, but also their genomes. Such a comparative approach provides us with inferences to comprehend the ancestral state, from which we could understand unique evolution in each species. That means, the seemingly different questions 'What makes humans human?' and 'What makes bonobos bonobo?', are actually interrelated.

Understanding bonobo sexuality, likewise, is directly related to understanding chimpanzee and human sexuality, which is of evolutionary significance, as it is immediately associated with fitness (Darwin 1871) and involves striking differences among the three closely related species. Humans are supposed to have diverged from *Pan* at least 6 Mya, and bonobos from chimpanzees about 2 Mya (Prado-Martinez et al. 2013; De Manuel et al. 2016). Humans

are distinct from the other two, as they do not have sexual swellings, although some reports suggest that the attractiveness of human female faces increases around their ovulation period (Roberts et al. 2004). Humans have a pair bonding system, where females are considered to have permanent sexual receptivity. Bonobos and chimpanzees, on the other hand, have sexual swellings (Nunn 1999), which is a sexual advertisement for females, while their sexual receptivity is very different: female bonobos are more often receptive, marked with their exaggerated sexual swelling, even when they are not ovulating, which is in stark contrast with female chimpanzees (Thompson-Handler, Malenky, and Badrian 1984; Furuichi 1987). Therefore, it is not easy to draw a conclusion on the ancestral status in terms of female sexuality (Kanō 1992; Wrangham 1993).

Exaggerated sexual swellings in bonobo females have been described to coin the bonobo society as 'hypersexual', supposedly not a good indicator for predicting ovulation timing. As some hypotheses suggest, this trait in female bonobos might have allowed them to avoid male monopoly during the fertile windows of females (Kanō 1992; Furuichi 1987; Wrangham 1993), to ally with other females (Ryu, Hill, and Furuichi 2015), and even to select males in the manner of reducing the level of aggressiveness/violence (contributing to their egalitarian society: Furuichi 2011; to self-domestication: Hare, Wobber, and Wrangham 2012; to extremely low levels of active aggression: Wrangham 2018). It has been reported, both in bonobos and chimpanzees, that high-rank males have a higher reproductive success (Wroblewski et al. 2009;

Surbeck et al. 2017). This is interesting, as it shows, even though male monopoly of females with a maximal sexual swelling is not common in bonobos, that high-rank males (or attractive males) end up having higher reproductive success, which implies active female choice on those males.

Our comparative analyses on bonobo and chimpanzee genomes, published in Han et al. 2019, where we assessed lineage-specific Single Nucleotide Changes (SNCs), suggest that such notable physiological differences in female sexual swellings might be explained by genetic changes in bonobos, which supports the hypothesis that it might be a derived feature in bonobos (Wrangham 1993). The main result in this line is the investigation of the putative effect of lineage-specific SNCs on phenotypes, by using the NHGRI-EBI GWAS Catalog (MacArthur et al. 2017), encompassing data on a total of 2,385 traits from genome-wide association studies on genetic loci in humans. We found that such lineage-specific nonsynonymous SNCs appear enriched in the genes associated to 5 unique traits on the bonobo lineage and 17 unique traits on the chimpanzee lineage (Supplementary Table S13 in (Han et al. 2019). “Menarche (age at onset)” is among the 5 unique traits enriched in bonobos, which is an obvious trait related to female reproduction. For validation, we investigated further 307 genes associated with “age at menarche” in the most recent and comprehensive analysis of this trait to date (Day et al. 2017), which was not included in the GWAS database. We found again that nonsynonymous bonobo-specific SNCs in menarche-associated genes were significantly more abundant than chimpanzee-specific

SNCs, regarding both the number of SNCs and the number of genes carrying these, and significantly more than by random expectation (Figure 4 in Han et al 2019, page 1186). These findings suggest that we could identify genetic candidate changes that underlie the physiological differences between bonobo and chimpanzee females, and understand their molecular mechanisms in future studies. Comprehending bonobo sexuality might eventually help us to understand human sexuality, described as monogamy and concealed ovulation (Benshoof and Thornhill 1979), and moreover to grasp the diverse adaptations of female reproductive strategies and their relation to features of sociality in general.

In chimpanzees, we did not find an enrichment for traits related to female reproduction. Instead, we found such an enrichment for traits which appear to be rather related to sociality more generally, like "Loneliness". This is interesting, as (Kanō 1992) once hypothesized that female chimpanzees evolved into social isolation. He explained that they stay alone with their dependent offspring and avoid taking part in group foraging, for a few years until the offspring is weaned and independent. This would need further tests across chimpanzee populations with variations in their resource distribution, but this finding of the enrichment of their genetic changes associated to this relevant trait is an interesting implication.

We acknowledge that these analyses have limitations coming first from the size of the data (69 genomes in total, which is still the largest dataset to date for bonobo and chimpanzee genomes), and second from our restricted understanding of genes and their

functions in humans and other species. We are still progressing in identifying the functions of genes, their interaction in networks, and environmental influences on them. The GWAS type of analyses and interpretation of Gene Ontology enrichments are still challenged for their implications for this reason (Altmüller 2001, Cheung 2010, Khatri 2005). Furthermore, since our knowledge on gene annotation and the genetic basis of phenotypic traits is massively biased towards humans, we should always be cautious in interpreting such an analysis of genes and genomes in other species.

However, despite these restrictions, we note that our observation of an enrichment of bonobo-specific genetic substitutions in genes associated to the female reproductive trait, age at menarche, yields a significant implication which could potentially explain one biologically relevant species-specific feature in bonobos, the prolonged maximal sexual swellings. This suggests potential future studies that could investigate the functional consequences of such candidate changes, or to determine whether or not those changes were rejected after inter-species admixture events, which would imply a more essential role in species-specific adaptation (Martin and Jiggins 2017).

4.2. Mutational load within the context of demographic history

The concept of nearly neutral mutations (Ohta 1972, 1973), *i.e.* not fully neutral mutations as they are supposed to confer minor fitness

effects, is a widely accepted and very influential theory in population genetics. The assumption is that such mutations are tolerated and behave as if they were neutral, not under strong enough selection pressure, and also that the threshold for the selection pressure depends on the effective population size (N_e) (Ohta 1972, 1973; Ohta and Gillespie 1996). It has been proposed that the efficacy of purifying selection positively correlates with the N_e of the populations, and multiple studies supported this based on observations of differences between taxa (archaic and modern humans, Lohmueller et al. 2008; Castellano et al. 2014; Kidd et al. 2012; Torkamani et al. 2012; Hodgkinson et al. 2013; dogs, Marsden et al. 2016; rice populations, Liu et al. 2017). In our study, we investigated several different genomic signatures to infer the efficacy of purifying selection, by using multiple genomes of bonobos and all the known chimpanzee subspecies. Our question was particularly whether western chimpanzees, which experienced a small N_e and population bottlenecks in the past (Won and Hey 2005; Prado-Martinez et al. 2013; De Manuel et al. 2016), would appear similar to bonobos in this aspect.

Our results (Han et al. 2019) suggest that the efficacy of purifying selection in western chimpanzees seems to be the lowest among chimpanzees, whereas it is quite similar to bonobos, which followed our expectation. For example, a population-wise version of the neutrality index (NI) (Rand and Kann 1996), is highest in western chimpanzees (1.51), compared to the other chimpanzee subspecies (1.19–1.28), while the NI in bonobos is lower than that in western, although higher than all other chimpanzee subspecies (1.36). The

direction of selection (DoS) (Stoletzki and Eyre-Walker 2011), also suggests that purifying selection might be most relaxed in western chimpanzees (0.1), compared to that in the other chimpanzee subspecies (0.04-0.06), and somewhat similar to that in bonobos (0.08). The same pattern was observed for non-coding sites in functional elements, such as 5' UTRs and the upstream and downstream regions of genes (Fig. 1B and Supplementary Table S2, in Han et al 2019, page 1182), which suggests that, also in non-coding loci, the efficacy of purifying selection seems to correlate with the N_e.

In another analysis, we measured the ratio of deleterious-to-neutral derived alleles across the site frequency spectrum (SFS) in each population (Fig. 2 and Supplementary Fig. S2 in Han et al 2019, page 1183). We found no consistent pattern across the 6 different deleteriousness measures, based on the expectation that deleterious alleles should reach high frequencies proportionally more than neutral alleles. However, proportionately more deleterious alleles were observed at high frequencies in bonobos, in comparison to all chimpanzees, using C-score and GERP score, both of which are genome-wide predictions, which consider more loci across genomes and hence may have more power than measures for protein-coding regions, especially for nearly neutral mutations. Bonobos likely have experienced a small N_e for a long period of time, since their divergence from chimpanzees, and the observations fit well into the expectation that they may also have accumulated more deleterious alleles than chimpanzees, which would segregate at high frequencies. C-score and GERP score are at least to some degree

phylogeny-based approaches, meaning that they include the western chimpanzee reference genome as part of their underlying data, which might make them biased towards derived alleles in western chimpanzees. Therefore, it is seems quite possible that those measures underestimate the proportion of deleterious alleles in western chimpanzees. However, across all the measures, there seems to be no clear correlation between N_e and proportion of deleterious alleles segregating at high frequencies. We speculate that it might correlate with other variables of demographic history, like severe bottlenecks or inbreeding, as observed in a study comparing domesticated rice to their wild type counterpart (Liu et al. 2017). Apart from that, we consistently observe across the 6 deleteriousness measures that proportionately more deleterious alleles appear at low frequencies in non-central chimpanzees, in comparison to central chimpanzees. This suggests that in central chimpanzees, which have the largest genetic diversity among all the chimpanzee lineages, deleterious alleles are more efficiently removed than in the other subspecies.

The analysis of individual mutational load, on the other hand, *i.e.* the number of sites with putatively deleterious derived alleles per individual, either in heterozygous state or in homozygous state, reveals a positive correlation with N_e for heterozygous sites and a negative correlation for homozygous sites (Fig. 3A, B in Han et al 2019, page 1184). This follows the expectation that a population with a large N_e has higher genetic diversity, which would be reflected by a higher heterozygosity due by singletons (Robertson 1965). In contrast, slightly deleterious mutations in a population

with a small N_e are effectively neutral, which should make them appear as homozygotes more often. Across 6 deleteriousness measures, we generally see that the mutational load in bonobos at homozygous sites is much higher than in all chimpanzees, as expected. Furthermore, the deleterious load in western chimpanzees is higher than that in other chimpanzee subspecies, except for C-score and GERP score. As mentioned earlier, this could be due to their phylogeny-based approach, and the use of a western chimpanzee as reference genome. This is also reflected in that the total numbers of deleterious derived alleles in western chimpanzees are lower than those in other chimpanzees only for these two methods (Supplementary Fig. S15 in Han et al 2019). However, the distributions of the homozygous mutational load of central chimpanzees are significantly different from the other three chimpanzee populations in C-score and GERP score ($P<0.01$, Wilcoxon rank test), as expected. In general, we interpret these results as that the mutational load at homozygous sites represents the efficacy of purifying selection, resulting in a higher load in bonobos than in chimpanzees, and also in western chimpanzees compared to the other chimpanzee subspecies.

These results lead to the conclusion that among chimpanzees, purifying selection was less efficient in western chimpanzees, which is in line with previous findings (Bataillon et al. 2015). They further suggest that the behavior of slightly deleterious alleles in western chimpanzees is comparable to that in bonobos, depending on the measure used. For example, selection indices, such as the NI and DoS, the proportion of slightly deleterious changes in non-

coding functional elements and the mutational load at homozygous sites show remarkably similar patterns, while the ratio of SFS of deleterious to neutral alleles separate only bonobos from chimpanzees, without differentiating chimpanzee lineages. This might be due to differential sensitivities of the analyses to specific demographic features, which need to be explored in detail in future studies. It has been reported in previous studies that variables in demographic history, such as timing and number of population bottlenecks a population experienced, recent expansion, or their interactions may result in differential genetic diversity (Masatoshi Nei, Maruyama, and Chakraborty 1975; Henn et al. 2015; Peischl et al. 2013), which might also invoke differential levels of purifying selection as well.

Slightly deleterious mutations seem to follow general expectations from the Nearly neutral mutation hypothesis (Ohta 1972, 1973), and reflect a differential efficacy of purifying selection in populations (Lohmueller et al. 2008; Bataillon et al. 2015; Cagan et al. 2016; Henn et al. 2016; Marsden et al. 2016; Liu et al. 2017; Han et al. 2019). However, it is still enigmatic how much this relates to the overall fitness of a population. This questions is in particular relevant for conservation biology, for instance, in relation to prioritizing populations for primary care, in order to avoid inbreeding depression (Franklin 1980; Lynch and Gabriel 1990), or to interpret such patterns as implications to extinction threat (Rogers and Slatkin 2017). Our study suggests that *Pan* lineages may be a good model to investigate this. Bonobos and western chimpanzees are of particular interest, because the slightly

deleterious mutations in those populations show quite similar patterns, although they have some differences (*e.g.* SFS), which might be related to unique variables in their demographic history, which we need to understand more deeply in the future.

4.3. Evidence of population substructure and divergence

Did bonobo groups substantially diverge from each other? Genetic divergence refers to the accumulation of new mutations in daughter populations, after the split from an ancestral population, followed by reproductive isolation (Palumbi 1994), although occasional gene flow may commonly occur between taxa, also in mammals (between humans and Neandertals, Kuhlwilm et al. 2016; Prüfer et al. 2014; from bonobos to chimpanzees, De Manuel et al. 2016; among horses, Gaunitz et al. 2018; Librado et al. 2017; among whales, Árnason et al. 2018; Foote et al. 2019; a ghost population to bonobos, Kuhlwilm et al. 2019).

In biological classification, the concept of speciation involves divergence. Darwin describes this as a purely quantitative problem, meaning that after accumulation of small differences, eventually qualitative differences would arise (Darwin 1859), although his arguments are mainly based on phenotypic divergence (Kozak et al. 2011). The most useful and influential definition of species probably comes from Mayr (Mayr 1942) which is the natural occurrence of free interbreeding, actual or potential, between

members of a group/population or between such populations. Different species, therefore, are populations having intrinsic factors that will act to prevent interbreeding between the populations of a degree as free as that within each population (Mayr 1942; Wilson and Brown 1953). Such a mechanism to prevent interbreeding between two populations, which is called reproductive isolation, can be achieved through multiple different mechanisms, including differential selection and hybrid sterility (Dobzhansky 1952). These considerations accommodate different concepts of divergence both at the phenotypic level and the genetic level.

However, regarding the divergence of distinct groups within a species, which is described as subspecies (Mayr 1942; Monroe 1982), the debate is more complicated. Subspecies is the only unit below the level 'species' recognized by the International Code of Zoological Nomenclature (Ferraris 2000), and refers to at least two populations within a species living in different geographic ranges, with varying morphological characteristics (Mayr 1942; Monroe 1982). In mammals, for example, leopards, lions, gorillas and chimpanzees are recognized to have different subspecies across their geographic habitat ranges (Miththapala, Seidensticker and O'Brian 1996; Uphyrkina et al. 2001; Kitchener et al. 2017; Wilson and Reeder 2005; Groves 2001). Such a classification is typically made in consideration of phenotypic differences and the geographic distribution of populations, as long as they could normally interbreed in captivity (Mayr 1942, 1982; Monroe 1982). Even though the concept of subspecies is often employed to refer to such a distinction across populations, it has been frequently argued that

this is an arbitrary concept, and easily misunderstood (Mayr 1942, 1982; Wilson and Brown 1953; Patten 2015). One strong argument is that the phenotypic traits do not always correlate with genetic divergence (maize, Bar-Hen et al. 1995; cucumber, Bernet et al. 2003; pepper, Kwon et al. 2005). For example, domestic dog and gray wolf differ only to 0.2% in their mtDNA sequences (Wayne 1993), which may be surprising given their stark differences in phenotypic traits, like tameness. Moreover, striking phenotypic divergence might be driven by geographic isolation or a severe population bottleneck, rather than long-term genetic separation (Mayr 1942, 1982).

Despite the controversies over the concept of subspecies, there seems to be a consensus that subspecies is a description referring to the diversity within a species, often correlating to geographic distances (Mayr 1942; Monroe 1982; Patten 2015), even though phenotypic diversities were highlighted (Patten 2015). Since this study aims primarily at addressing the genetic diversity across bonobo populations, without knowing the full extent of phenotypic diversity of bonobos in the wild, we do not employ the concept of subspecies. However, I refer to the distinct chimpanzee populations from four geographic origins as subspecies, following previous studies and current knowledge, while I refer to genetic divergence within bonobos using the broader term population, with genetic inferences on potential natural selection.

Going back to the question whether bonobo populations have substantially diverged from each other, our analyses do suggest that

this is the case. Using 20 published bonobo exomes (Teixeira et al. 2015), a principle component analysis (PCA), a Neighbor-joining (NJ) tree and ADMIXTURE suggest that they represent three genetically distinct groups, or populations (Figure 1, in the manuscript "Divergence in bonobos: genomic evidence", page 10), which appear to be genetically divided, as seen in pairwise F_{ST} values (Table 1, in the manuscript "Divergence in bonobos: genomic evidence", page 13). The level of differentiation seems to be comparable to that between closely related chimpanzee subspecies, or between African and non-African human populations (Table 1, in the manuscript "Divergence in bonobos: genomic evidence", page 13). This is surprising, since it suggests that wild bonobo populations might carry a level of diversity similar to that between different chimpanzee or human populations, contrasting with the early understandings of bonobos as homogeneous and low in genetic diversity (Fischer et al. 2011; Prado-Martinez et al. 2013). The three bonobo populations, as inferred from their mtDNA haplotypes, seem to be part of the previously suggested broader western and central populations (Kawamoto et al. 2013), with B1 and B2 being related to western and B3 to central bonobos (Figure 2, in the manuscript "Divergence in bonobos: genomic evidence", page 11). In particular, the B1 population appears to be indistinguishable from the Malebo population, which is known to inhabit the forest-savanna mosaic area in the western corner of the DRC (Pruetz and Bertolani 2007; Poulsen and Clark 2004). This would predict that B1 has experienced a population bottleneck as a consequence of their migration history (Takemoto et al. 2017),

which seems to be summarized in F_{ST} values. This is also in line with the analysis of runs of homozygosity (ROH), the Identity by descent (IBD), and mutational load analyses, which all demonstrated that the B1 population has the lowest genetic diversity, with proportionally more shared recent ancestors, and a higher mutational load. These results imply that high F_{ST} values in the three bonobo populations might be driven by the highest amount of genetic drift of the B1 population.

The population parameter estimates from G-PhoCS suggest that the split between the western and the central populations occurred at least 100,000 years ago, followed by the split between the two western populations roughly 40,000 years ago (Figure 3, in the manuscript "Divergence in bonobos: genomic evidence", page 11). This is indeed within the range of divergence between rather deeply divergent modern human populations, as suggested by the F_{ST} analysis. If we consider the social structure of bonobos, where females migrate out on sexual maturity, and introgression, which happens frequently in primates, complicating divergence estimates (Prüfer et al. 2014; Veeramah et al. 2015; Kuhlwilm et al. 2016; De Manuel et al. 2016), it is reasonable to assume some amount of gene flow across the three bonobo populations after their initial split. If that is the case, the actual split of the populations might be further back in time. However, our G-PhoCS analysis assuming a few different migration scenarios failed to reach stable values, which we speculate to be a consequence of the G-PhoCS program failing to support one conclusion over the other: Either a high migration rate with a deeper split or a small migration rate with a shorter split time

(Ilan Gronau, personal communication). Therefore, it should be cautioned that the estimates we present in our study might be the most conservative ones, suggesting their actual divergence to be deeper.

Regarding the time to the most recent common ancestor (TMRCA) of all wild bonobos, it would most likely be substantially deeper than our estimates. Our 20 exome samples seem to cover only the western and central populations of bonobos in the wild, and Takemoto et al 2017 estimated the time to the TMRCA of bonobos to be 0.64 or 0.95 million years ago, using the mtDNA of wild bonobos, and particularly including the eastern population carrying the deepest-splitting lineages. This suggests that the genome-wide divergence of bonobo populations in the wild might be deeper, and potentially involve unique and distinct adaptations, in interaction with their environment. Bonobos have long been thought of as a species which evolved in a rather homogenous dense tropical forest (Hashimoto et al. 1998; Furuichi and Thompson 2007) with varying types of vegetation (Hashimoto et al. 1998; Kanō 1992), but a growing body of studies indicates a broader diversity in their habitats (Thompson 2002) and behaviors (Hohmann and Fruth 2003).

Our observations on genomic features in bonobos imply that their within-species divergence might overlap those between closely related chimpanzee subspecies, which needs to be studied further in the future by sequencing more individuals to represent bonobo populations. As biosamples from the wild bonobo field sites, which

are sequencing materials, are limited to non-invasive collection, we need to make use of them, such as feces, hairs and saliva (Inoue et al. 2007, Andronic 2009, Hernandez-Rodriguez et al. 2017), for which techniques for whole-genome sequencing are under development. Ancient DNA sequences would help understanding genetic diversity in bonobos as well. However, great ape fossils are extremely rare, especially in Africa, due to the climate and soil conditions, which are not good for bone and DNA preservation (McBrearty and Jablonski 2005, Kuhlwilm et al. 2019). Alternatively, genetic diversity in the relatively recent past could be investigated by making use of samples in museums, which are valuable resources to study the diminishing genetic diversity during the recent decades (Van der Valk et al. 2019).

Contributions to other publications

1.

"The Genomic Footprints of the Fall and Recovery of the Crested Ibis."

Shaohong Feng, Qi Fang, Ross Barnett, Cai Li, **Sojung Han,** Martin Kuhlwilm, Long Zhou, Hailin Pan, Yuan Deng, Guangji Chen, Anita Gamauf, Friederike Woog, Rober Prys-Jones, Tomas Marques-Bonet, Thomas Gilbert, Guojie Zhang.

Current Biology. 2019.

DOI: https://doi.org/10.1016/j.cub.2018.12.008.

2.

"Ancient Admixture from an Extinct Ape Lineage into Bonobos."

Martin Kuhlwilm, **Sojung Han,** Vitor C. Sousa, Laurent Excoffier, and Tomas Marques-Bonet.

Nature Ecology & Evolution, 2019.

DOI: https://doi.org/10.1038/s41559-019-0881-7.

Bibliography

1000 Genomes Project Consortium, The 1000 Genomes Project, Gonçalo R Abecasis, David Altshuler, Adam Auton, Lisa D Brooks, Richard M Durbin, Richard A Gibbs, Matt E Hurles, and Gil A McVean. 2010. "A Map of Human Genome Variation from Population-Scale Sequencing." *Nature* 467 (7319): 1061–73.

Adolph H. Schultz. 1966. *Age Changes, Sex Differences, and Variability as Factors in the Classification of Primates.*

Adzhubei, Ivan A, Steffen Schmidt, Leonid Peshkin, Vasily E Ramensky, Anna Gerasimova, Peer Bork, Alexey S Kondrashov, and Shamil R Sunyaev. 2010. "A Method and Server for Predicting Damaging Missense Mutations." *Nature Methods* 7 (4): 248–49.

Andersson, M. B. 1994. *Sexual Selection.* Princeton University Press.

Arandjelovic, M. et al. 2009. "Two-step multiplex PCR improves the speed and accuracy of genotyping using DNA from noninvasive and museum samples." *Molecular Ecology Resources* 9: 28-36.

Árnason, Úlfur, Fritjof Lammers, Vikas Kumar, Maria A. Nilsson, and Axel Janke. 2018. "Whole-Genome Sequencing of the Blue Whale and Other Rorquals Finds Signatures for Introgressive Gene Flow." *Science Advances* 4 (4): eaap9873.

Badrian, Noel, Alison Badrian, and Randall L. Susman. 1981. "Preliminary Observations on the Feeding Behavior of *Pan Paniscus* in the Lomako Forest of Central Zaïre." *Primates* 22 (2): 173–81.

Bar-Hen, A., A. Charcosset, M. Bourgoin, and J. Guiard. 1995. "Relationship between Genetic Markers and Morphological Traits in a Maize Inbred Lines Collection." *Euphytica* 84 (2): 145–54.

Bataillon, Thomas, Jinjie Duan, Christina Hvilsom, Xin Jin, Yingrui Li, Laurits Skov, Sylvain Glemin, et al. 2015. "Inference of Purifying and Positive Selection in Three Subspecies of Chimpanzees (*Pan Troglodytes*) from Exome Sequencing." *Genome Biology and Evolution* 7 (4): 1122–32.

Bateman, A J. 1948. "Intra-Sexual Selection in Drosophila." *Heredity* 2 (3): 349–68.

Becquet, C.,Patterson, N., Stone, A. C., Przeworski, M., Reich, D. 2007. "Genetic Structure of Chimpanzee Populations." *PLoS Genetics* 3 (4).

Behringer, V., T. Deschner, C. Deimel, J.M.G. Stevens, and G. Hohmann. 2014. "Age-Related Changes in Urinary Testosterone Levels Suggest Differences in Puberty Onset and Divergent Life History Strategies in Bonobos and Chimpanzees." *Hormones and Behavior* 66 (3): 525–33.

Benshoof, Lee, and Randy Thornhill. 1979. "The Evolution of Monogamy and Concealed Ovulation in Humans." *Journal of Social and Biological Structures* 2 (2): 95–106.

Bernet, G. P., S. Bramardi, D. Calvache, E. A. Carbonell, and M. J. Asins. 2003. "Applicability of Molecular Markers in the Context of Protection of New Varieties of Cucumber." *Plant Breeding* 122 (2): 146–52.

Blount, Ben. 1990. "Issues in Bonobo (*Pan Paniscus*) Sexual Behavior." *American Anthropologist*.

Boesch, C. 2002. “Cooperative Hunting Roles among Taï Chimpanzees.” *Human Nature* 13 (1): 27–46.

Boesch, Christophe., and Hedwige Boesch-Achermann. 2000. *The Chimpanzees of the Taï Forest : Behavioural Ecology and Evolution.* Oxford University Press.

Boesch, Christophe. 1991. “The Effects of Leopard Predation on Grouping Patterns in Forest Chimpanzees.” *Behaviour*. Brill.

Boesch, Christophe, Catherine Crockford, Ilka Herbinger, Roman Wittig, Yasmin Moebius, and Emmanuelle Normand. 2008. “Intergroup Conflicts among Chimpanzees in Taï National Park: Lethal Violence and the Female Perspective.” *American Journal of Primatology* 70 (6): 519–32.

Boesch, Christophe, Gottfried Hohmann, and Linda F Marchant. 2002. “Behavioural Diversity in Chimpanzees and Bonobos.”

Caballero, Armando. 1994. “Developments in the Prediction of Effective Population Size.” *Heredity* 73 (6): 657–79.

Cagan, Alexander, Christoph Theunert, Hafid Laayouni, Gabriel Santpere, Marc Pybus, Ferran Casals, Kay Prüfer, et al. 2016. “Natural Selection in the Great Apes.” *Molecular Biology and Evolution* 33 (12): 3268–83.

Carl J. Ferraris, William N. Eschmeyer. 2000. *International Code of Zoological Nomenclature: Fourth Edition*.

Carlisle, Kay S, Robert M Brenner, and William Montagna. 1981. “Hormonal Regulation of Sex Skin in *Macaca Nemestrina*.” *BIOLOGY OF REPRODUCTION*. Vol. 25.

Castellano, S, G Parra, F A Sánchez-Quinto, F Racimo, M Kuhlwilm, M Kircher, S Sawyer, et al. 2014. “Patterns of Coding Variation in the Complete Exomes of Three Neandertals.”

Chapman, C, L Chapman, and R Wrangham. 1995. “Ecological Constraints on Group Size: An Analysis of Spider Monkey and Chimpanzee Subgroups.” *Behavioral Ecology and Sociobiology* 36 (1): 59–70.

Charlesworth, J., and A. Eyre-Walker. 2007. “The Other Side of the Nearly Neutral Theory, Evidence of Slightly Advantageous Back-Mutations.” *Proceedings of the National Academy of Sciences* 104 (43): 16992–97.

Clutton-Brock, T.H. and P. H. Harvey. 1976. “Growing Points in Ethology : Based on a Conference Sponsored by St. John’s College and King’s College, Cambridge.” In *Growing Points in Ethology, Bateson and Hinde*, 548. Cambridge University Press.

Clutton-Brock, T. H., D. Green, M. Hiraiwa-Hasegawa, and S. D. Albon. 1988. “Passing the Buck: Resource Defence, Lek Breeding and Mate Choice in Fallow Deer.” *Behavioral Ecology and Sociobiology* 23 (5): 281–96.

Collins, Francis S, Michael Morgan, and Aristides Patrinos. 2003. “The Human Genome Project: Lessons from Large-Scale Biology.” *Science (New York, N.Y.)* 300 (5617): 286–90.

Coolidge, Harold J. 1933. “*Pan Paniscus*. Pigmy Chimpanzee from South of the Congo River.” *American Journal of Physical Anthropology* 18 (1): 1–59.

Crow, JF, and M Kimura. 1970. “An Introduction to Population Genetics Theory.” *Cabdirect.Org*.

Cuadrado, Mariano. 1998. “Models Painted with Female-like Colors Elicited Courtship by Male Common Chameleons: Evidence for a Courtship Releaser.” *Journal of Ethology* 16 (2): 73–79.

Dahl, Jeremy F. 1986. “Cyclic Perineal Swelling during the Intermenstrual Intervals of Captive Female Pygmy Chimpanzees (*Pan Paniscus*).” *Journal of Human Evolution* 15 (5): 369–85.

Darwin, C. 1859. *On the Origin of Species.* London: John Murray.

———. 1871. *The Descent of Man, and Selection in Relation to Sex.*

Davydov, Eugene V., David L. Goode, Marina Sirota, Gregory M. Cooper, Arend Sidow, and Serafim Batzoglou. 2010. “Identifying a High Fraction of the Human Genome to Be under Selective Constraint Using GERP++.” Edited by Wyeth W. Wasserman. *PLoS Computational Biology* 6 (12): e1001025.

Day, Felix R, Deborah J Thompson, Hannes Helgason, Daniel I Chasman, Hilary Finucane, Patrick Sulem, Katherine S Ruth, et al. 2017. “Genomic Analyses Identify Hundreds of Variants Associated with Age at Menarche and Support a Role for Puberty Timing in Cancer Risk.” *Nature Genetics* 49 (6): 834–41.

Dayhoff, M, and P McLaughlin. 1972. *Atlas of Protein Sequence and Structure.* National Biomedical Research Foundation.

Deschner, Tobias, Michael Heistermann, Keith Hodges, and Christophe Boesch. 2004. “Female Sexual Swelling Size, Timing of Ovulation, and Male Behavior in Wild West African Chimpanzees.” *Hormones and Behavior* 46 (2): 204–15.

Dixson, AF. 1983. “Observations on the Evolution and Behavioral Significance of ‘Sexual Skin’ in Female Primates.” *Advances in the Study of Behavior* 13 (January): 63–106.

———. 2009. *Sexual Selection and the Origins of Human Mating Systems.* Oxford University Press.

Dixson, AF, and NI Mundy. 1994. “Sexual Behavior, Sexual Swelling, and Penile Evolution in Chimpanzees (Pan Troglodytes).” *Archives of Sexual Behavior* 23 (3): 267–80.

Dobzhansky, T. 1952. “Genetics and the Origin of Species.” In *Science*, 115:430–31. American Association for the Advancement of Science.

Doolittle, R, and B Blomback. 1964. “Amino-Acid Sequence Investigations of Fibrinopeptides from Various Mammals: Evolutionary Implications.” *Nature* 202 (4928): 147–52.

Douglas, Pamela Heidi, Gottfried Hohmann, Róisín Murtagh, Robyn Thiessen-Bock, and Tobias Deschner. 2016. “Mixed Messages: Wild Female Bonobos Show High Variability in the Timing of Ovulation in Relation to Sexual Swelling Patterns.” *BMC Evolutionary Biology* 16 (1): 140.

Eberhard, WG. 1985. *Sexual Selection and Animal Genitalia.* Harvard Univ. Press.

———. 1996. *Female Control : Sexual Selection by Cryptic Female Choice.* Princeton University Press.

———. 2009. “Static Allometry and Animal Genitalia.” *Evolution* 63 (1): 48–66.

Emery, Melissa A., and Patricia L. Whitten. 2003. “Size of Sexual Swellings Reflects Ovarian Function in Chimpanzees (*Pan Troglodytes*).” *Behavioral Ecology and Sociobiology*. Springer.

Emlen, Stephen T., and Lewis W. Oring. 1977. “Ecology, Sexual Selection, and the Evolution of Mating Systems.” *Science*. American Association for the Advancement of Science.

Ernst Schwarz. 1929. “Das Vorkommen Des Schimpansen Auf Den Linken Kongo-Ufer.” *Revue de Zoologie et de Botanique Africaines*.

Evans, Jonathan P., Lorenzo Zane, Samuela Francescato, and Andrea Pilastro. 2003. “Directional Postcopulatory Sexual Selection Revealed by Artificial Insemination.” *Nature* 421 (6921): 360–63.

Eyre-Walker, Adam, and Peter D. Keightley. 2007. “The Distribution of Fitness Effects of New Mutations.” *Nature Reviews Genetics* 8 (8): 610–18.

Fischer, Anne, Kay Prüfer, Jeffrey M. Good, Michel Halbwax, Victor Wiebe, Claudine André, Rebeca Atencia, Lawrence Mugisha, Susan E. Ptak, and Svante Pääbo. 2011. “Bonobos Fall within the Genomic Variation of Chimpanzees.” Edited by Etienne Joly. *PLoS ONE* 6 (6): e21605.

Fitch, WM, and E Margoliash. 1967. “Construction of Phylogenetic Trees.” *Science (New York, N.Y.)* 155 (3760): 279–84.

Foote, Andrew D., Michael D. Martin, Marie Louis, George Pacheco, Kelly M. Robertson, Mikkel-Holger S. Sinding, Ana R. Amaral, et al. 2019. “Killer Whale Genomes Reveal a Complex History of Recurrent Admixture and Vicariance.” *BioRxiv*, January, 520718.

Ford, E.B. 1964. *Ecological Genetics.*

Franklin, I.R. 1980. “Evolutionary Change in Small Populations.” In *Conservation Biology : An Evolutionary-Ecological Perspective*, 395. Sinauer Associates.

Furuichi, T. 1987. “Sexual Swelling, Receptivity, and Grouping of Wild Pygmy Chimpanzee Females at Wamba, Zaïre.” *Primates* 28 (3): 309–18.

———. 1989. “Social Interactions and the Life History of Female Pan Paniscus in Wamba, Zaire.” *International Journal of Primatology* 10 (3): 173–97.

———. 1997. “Agonistic Interactions and Matrifocal Dominance Rank of Wild Bonobos (Pan Paniscus) at Wamba.” *International Journal of Primatology* 18 (6): 855–75.

———. 2011. “Female Contributions to the Peaceful Nature of Bonobo Society.” *Evolutionary Anthropology: Issues, News, and Reviews* 20 (4): 131–42.

Furuichi, T, and C Hashimoto. 2002. “Why Female Bonobos Have a Lower Copulation Rate during Estrus than Chimpanzees.” In *Behavioral Diversity in Chimpanzees and Bonobos*.

———. 2004. “Sex Differences in Copulation Attempts in Wild Bonobos at Wamba.” *Primates* 45 (1): 59–62.

Furuichi, T, and M Thompson. 2007. *The Bonobos : Behavior, Ecology, and Conservation*. Springer.

Gaunitz, Charleen, Antoine Fages, Kristian Hanghøj, Anders Albrechtsen, Naveed Khan, Mikkel Schubert, Andaine Seguin-Orlando, et al. 2018. “Ancient Genomes Revisit the Ancestry of Domestic and Przewalski’s Horses.” *Science (New York, N.Y.)* 360 (6384): 111–14.

Gillman, Joseph. 1940. “Experimental Studies on the Menstrual Cycle of the Baboon (*Papio Porcarius*). VI. The Effect of Progesterone upon the First Part of the Cycle in Normal Female Baboons.” *Endocrinology* 26 (1): 80–87.

Gillman, Joseph, and H. B. Stein. 1941. “A Quantitative Study of the Inhibition of Estradiol Benzoate by Progesterone in the Baboon (*Papio Porcarius*).” *Endocrinology* 28 (2): 274–82.

Goodall, Jane. 1986. *The Chimpanzees of Gombe : Patterns of Behavior.* Belknap Press of Harvard University Press.

Goodman, Morris, Calvin A. Porter, John Czelusniak, Scott L. Page, Horacio Schneider, Jeheskel Shoshani, Gregg Gunnell, and Colin P. Groves. 1998. "Toward a Phylogenetic Classification of Primates Based on DNA Evidence Complemented by Fossil Evidence." *Molecular Phylogenetics and Evolution* 9 (3): 585–98.

Grantham, R. 1974. "Amino Acid Difference Formula to Help Explain Protein Evolution." *Science* 185 (4154).

Gronau, Ilan, Melissa J Hubisz, Brad Gulko, Charles G Danko, and Adam Siepel. 2011. "Bayesian Inference of Ancient Human Demography from Individual Genome Sequences." *Nature Genetics* 43 (10): 1031–34.

Groves, Colin P. 2001. *Primate Taxonomy.* Smithsonian Institution Press.

Gruber, Thibaud, Zanna Clay, and Klaus Zuberbühler. 2010. "A Comparison of Bonobo and Chimpanzee Tool Use: Evidence for a Female Bias in the *Pan* Lineage." *Animal Behaviour* 80 (6): 1023–33.

Hamilton, W.J. III. 1984. "Significance of Paternal Investment by Primates to the Evolution of Adult Male-Female Associations." van Nostrand Reinhold Co.

Han, Sojung, Aida M Andrés, Tomas Marques-Bonet, and Martin Kuhlwilm. 2019. "Genetic Variation in *Pan* Species Is Shaped by Demographic History and Harbors Lineage-Specific Functions." Edited by Naruya Saitou. *Genome Biology and Evolution* 11 (4): 1178–91.

Hare, B, V Wobber, and R Wrangham. 2012. "The Self-Domestication Hypothesis: Evolution of Bonobo Psychology Is Due to Selection against Aggression." *Animal Behaviour* 83 (3): 573–85.

Hasegawa, Masami, Taka-aki Yano, and Hirohisa Kishino. 1984. "A New Molecular Clock of Mitochondrial DNA and the Evolution of Hominoids." *Proceedings of the Japan Academy. Ser. B: Physical and Biological Sciences* 60 (4): 95–98.

Hashimoto, Chie, Yasuko Tashiro, Daiji Kimura, Tomoo Enomoto, Ellen J. Ingmanson, Gen'ichi Idani, and Takeshi Furuichi. 1998. "Habitat Use and Ranging of Wild Bonobos (*Pan Paniscus*) at Wamba." *International Journal of Primatology* 19 (6): 1045–60.

Henn, Brenna M., Laura R. Botigué, Carlos D. Bustamante, Andrew G. Clark, and Simon Gravel. 2015. "Estimating Mutation Load in Human Genomes." *Nature Reviews. Genetics* 16 (6): 333.

Henn, Brenna M. et al. 2016. "Distance from Sub-Saharan Africa Predicts Mutational Load in Diverse Human Genomes." *Proceedings of the National Academy of Sciences of the United States of America* 113 (4): E440-9.

Hernandez-Rodriguez, J. et al. 2017. "The impact of endogenous content, replicates and pooling on genome capture from faecal samples." *Mol. Ecol. Resour.* 18, 319–333.

Hirai, Hirohisa., Hiroo. Imai, and Yasuhiro. Go. 2012. *Post-Genome Biology of Primates.* Springer.

Hodgkinson, Alan, Ferran Casals, Youssef Idaghdour, Jean-Christophe Grenier, Ryan D Hernandez, and Philip Awadalla. 2013. "Selective Constraint, Background Selection, and Mutation Accumulation Variability within and between Human Populations." *BMC Genomics* 14 (July): 495.

Hohmann, G, and B Fruth. 1993. “Field Observations on Meat Sharing among Bonobos (*Pan Paniscus*).” *Folia Primatologica* 60 (4): 225–29.
———. 2002. “Dynamics in Social Organization of Bonobos (*Pan Paniscus*).” In *Behavioral Diversity in Chimpanzees and Bonobos*.
———. 2003. “Culture in Bonobos? Between- species and Within- species Variation in Behavior.” *Current Anthropology* 44 (4): 563–71.
Hoyer, Bill H., N.W. van de Velde, Morris Goodman, and R.B. Roberts. 1972. “Examination of Hominid Evolution by DNA Sequence Homology.” *Journal of Human Evolution* 1 (6): 645–49.
Hrdy, SB. 1974. “Male-Male Competition and Infanticide among the Langurs (Presbytis Entellus) of Abu, Rajasthan.” *Folia Primatologica* 22 (1): 19–58.
———. 1979. “Infanticide among Animals: A Review, Classification, and Examination of the Implications for the Reproductive Strategies of Females.” *Ethology and Sociobiology* 1 (1): 13–40.
Hrdy, SB, and PL Whitten. 1987. “Patterning of Sexual Activity.” In *Primate Societies*.
Huchard, E, LA Knapp, J Wang, M Raymond, and G Cowlishaw. 2010. “MHC, Mate Choice and Heterozygote Advantage in a Wild Social Primate.” *Molecular Ecology* 19 (12): no-no.
Idani, G. 1990. “Relations between Unit-Groups of Bonobos at Wamba, Zaire: Encounters and Temporary Fusions.” *African Study Monographs*.
———. 1991. “Social Relationships between Immigrant and Resident Bonobo (*Pan Paniscus*) Females at Wamba.” *Folia Primatologica* 57 (2): 83–95.
Inoue, E. M. Inoue-Murayama, O. Takenaka, T. Nishida. 2007. “Wild chimpanzee infant urine and saliva sampled noninvasively usable for DNA analyses.” 48 (2): 156–159.
Itoh, Noriko, and Toshisada Nishida. 2007. “Chimpanzee Grouping Patterns and Food Availability in Mahale Mountains National Park, Tanzania.” *Primates* 48 (2): 87–96.
Janson, Charles H., and Michele L. Goldsmith. 1995. “Predicting Group Size in Primates: Foraging Costs and Predation Risks.” *Behavioral Ecology* 6 (3): 326–36.
Jensen, Jeffrey D., Bret A. Payseur, Wolfgang Stephan, Charles F. Aquadro, Michael Lynch, Deborah Charlesworth, and Brian Charlesworth. 2019. “The Importance of the Neutral Theory in 1968 and 50 Years on: A Response to Kern and Hahn 2018.” *Evolution* 73 (1): 111–14.
Johann F. Blumenbach. 1776. “De Generis Humani Varietate Nativa Liber Cum Figuris Aeri Incisis.”
Jukes, TH, and M Kimura. 1984. “Evolutionary Constraints and the Neutral Theory.” *Journal of Molecular Evolution* 21 (1): 90–92.
Kanō, Takayoshi. 1992. *The Last Ape : Pygmy Chimpanzee Behavior and Ecology*. Stanford University Press.
Kato, Junzo, Tsuneko Onouchi, and Kiyoshi Oshima. 1980. “The Presence of Progesterone Receptors in the Sexual Skin of the Monkey.” *Steroids* 36 (6): 743–49.
Kawamoto, Yoshi, Hiroyuki Takemoto, Shoko Higuchi, Tetsuya Sakamaki, John A. Hart, Terese B. Hart, Nahoko Tokuyama, et al. 2013. “Genetic Structure of Wild Bonobo Populations: Diversity of Mitochondrial DNA and

Geographical Distribution." Edited by David Caramelli. *PLoS ONE* 8 (3): e59660.
Kidd, J. M. et al. 2012. "Population Genetic Inference from Personal Genome Data: Impact of Ancestry and Admixture on Human Genomic Variation." *American Journal of Human Genetics* 91 (4): 660–71.
Kimura, M. 1968. "Evolutionary Rate at the Molecular Level." *Nature* 217 (5129): 624–26.
King, JL, and TH Jukes. 1969. "Non-Darwinian Evolution." *Science (New York, N.Y.)* 164 (3881): 788–98.
Kircher, M, D M Witten, P Jain, B J O'Roak, G M Cooper, and J Shendure. 2014. "A General Framework for Estimating the Relative Pathogenicity of Human Genetic Variants." *Nature Genetics* 46 (3): 310–15.
Kirkpatrick, Mark, and Michael J. Ryan. 1991. "The Evolution of Mating Preferences and the Paradox of the Lek." *Nature* 350 (6313): 33–38.
Kitchener, A. C., C. Breitenmoser-Würsten, E. Eizirik, A. Gentry, Lars Werdelin, A. Wilting, N. Yamaguchi, and Warren E. Johnson. 2017. *A Revised Taxonomy of the* Felidae *: The Final Report of the Cat Classification Task Force of the IUCN Cat Specialist Group.* International Union for Conservation of Nature and Natural Resources, Species Survival Commission, Cat Specialist Group.
Kozak, Marcin, Jan Bocianowski, Alina Liersch, Małgorzata Tartanus, Iwona Bartkowiak-Broda, Fernando A. Piotto, and Ricardo A. Azevedo. 2011. "Genetic Divergence Is Not the Same as Phenotypic Divergence." *Molecular Breeding* 28 (2): 277–80.
Krohn, P L, and S Zuckerman. 1937. "Water Metabolism in Relation to the Menstrual Cycle."
Kuhlwilm, M. et al. 2016. "Ancient Gene Flow from Early Modern Humans into Eastern Neanderthals." *Nature* 530 (7591): 429–33.
Kuhlwilm, Martin, Sojung Han, Vitor C. Sousa, Laurent Excoffier, and Tomas Marques-Bonet. 2019. "Ancient Admixture from an Extinct Ape Lineage into Bonobos." *Nature Ecology & Evolution*, 3(6):957-965.
Kumar, Prateek, Steven Henikoff, and Pauline C Ng. 2009. "Predicting the Effects of Coding Non-Synonymous Variants on Protein Function Using the SIFT Algorithm." *Nature Protocols* 4 (8): 1073–81.
Kuroda, Suehisa. 1979. "Grouping of the Pygmy Chimpanzees." *Primates* 20 (2): 161–83.
Kwon, Y. et al. 2005. "Use of SSR Markers to Complement Tests of Distinctiveness, Uniformity, and Stability (DUS) of Pepper (*Capsicum Annuum L.*) Varieties." *Molecules and Cells* 19 (3): 428–35.
Langergraber, Ke, and Kay Prüfer. 2012. "Generation Times in Wild Chimpanzees and Gorillas Suggest Earlier Divergence Times in Great Ape and Human Evolution." *Proceedings of the ...* 109: 15716–21.
Lawick-Goodall, Jane Van. 1968. "The Behaviour of Free-Living Chimpanzees in the Gombe Stream Reserve." *Animal Behaviour Monographs* 1 (January): 161-IN12.
Lehmann, Julia, and Christophe Boesch. 2008. "Sexual Differences in Chimpanzee Sociality." *International Journal of Primatology* 29 (1): 65–81.

Lehmann, Julia, Amanda H. Korstjens, and R. I. M. Dunbar. 2007. “Fission–Fusion Social Systems as a Strategy for Coping with Ecological Constraints: A Primate Case.” *Evolutionary Ecology* 21 (5): 613–34.

Li, Wen-Hsiung, Chung-I Wu, and Chi-Cheng Luo. 1984. “Nonrandomness of Point Mutation as Reflected in Nucleotide Substitutions in Pseudogenes and Its Evolutionary Implications.” *Journal of Molecular Evolution* 21 (1): 58–71.

Librado, Pablo, Cristina Gamba, Charleen Gaunitz, Clio Der Sarkissian, Mélanie Pruvost, Anders Albrechtsen, Antoine Fages, et al. 2017. “Ancient Genomic Changes Associated with Domestication of the Horse.” *Science (New York, N.Y.)* 356 (6336): 442–45.

Liu, Qingpo, Yongfeng Zhou, Peter L. Morrell, and Brandon S. Gaut. 2017. “Deleterious Variants in Asian Rice and the Potential Cost of Domestication.” *Molecular Biology and Evolution* 167 (4): msw296.

Lohmueller, Kirk E., Amit R. Indap, Steffen Schmidt, Adam R. Boyko, Ryan D. Hernandez, Melissa J. Hubisz, John J. Sninsky, et al. 2008. “Proportionally More Deleterious Genetic Variation in European than in African Populations.” *Nature* 451 (7181): 994–97.

Lukas, Dieter, and Elise Huchard. 2014. “Sexual Conflict. The Evolution of Infanticide by Males in Mammalian Societies.” *Science (New York, N.Y.)* 346 (6211): 841–44.

Lynch, Michael, and Wilfried Gabriel. 1990. “Mutation Load and the Survival of Small Populations.” *Evolution* 44 (7): 1725–37.

MacArthur, J. et al. 2017. “The New NHGRI-EBI Catalog of Published Genome-Wide Association Studies (GWAS Catalog).” *Nucleic Acids Research* 45 (D1): D896–901.

Mahmood, Khalid, Chol-hee Jung, Gayle Philip, Peter Georgeson, Jessica Chung, Bernard J. Pope, and Daniel J. Park. 2017. “Variant Effect Prediction Tools Assessed Using Independent, Functional Assay-Based Datasets: Implications for Discovery and Diagnostics.” *Human Genomics* 11 (1): 10.

Mallick, Swapan, Heng Li, Mark Lipson, Iain Mathieson, Melissa Gymrek, Fernando Racimo, Mengyao Zhao, et al. 2016. “The Simons Genome Diversity Project: 300 Genomes from 142 Diverse Populations.” *Nature* 538 (7624): 201–6.

de Manuel, M., M. Kuhlwilm, P. Frandsen, V.C. Sousa, T. Desai, J. Prado-Martinez, J. Hernandez-Rodriguez, et al. 2016. “Chimpanzee Genomic Diversity Reveals Ancient Admixture with Bonobos.” *Science* 354 (6311).

Margoliash, E. 1963. “Primary Structure and Evolution of Cytochrome C.” *Proceedings of the National Academy of Sciences of the United States of America* 50 (4): 672–79.

Margoliash, E, and E Smith. 1965. “Structural and Functional Aspects of Cytochrome c in Relation to Evolution.” *Evolving Genes and Proteins*, January, 221–42.

Marsden, C. D. et al. 2016. “Bottlenecks and Selective Sweeps during Domestication Have Increased Deleterious Genetic Variation in Dogs.” *Proceedings of the National Academy of Sciences of the United States of America* 113 (1): 152–57.

Martin, Simon H, and Chris D Jiggins. 2017. “Interpreting the Genomic Landscape of Introgression.” *Current Opinion in Genetics & Development* 47 (December): 69–74.

Mayr, E. 1942. *Systematics and the Origin of Species.* Harvard University Press.

———. 1963. *Animal Species and Evolution.* Cambridge MA: Belknap Press of Harvard University Press.

———. 1965. “Evolving Genes and Proteins.” In .

———. 1982. “Speciation and Macroevolution.” *Evolution* 36 (6): 1119–32.

McBrearty, S. and Jablonski, N. G. 2005. “First fossil chimpanzee.” *Nature* 437, 105–108.

McGrew, W. C. 1983. “Animal Foods in the Diets of Wild Chimpanzees (*Pan Troglodytes*): Why Cross-Cultural Variation?” *Journal of Ethology* 1 (1–2): 46–61.

Meyer, Rachel S, and Michael D Purugganan. 2013. “Evolution of Crop Species: Genetics of Domestication and Diversification.” *Nature Reviews. Genetics* 14 (12): 840–52.

Mitani, John C., and David P. Watts. 1999. “Demographic Influences on the Hunting Behavior of Chimpanzees.” *American Journal of Physical Anthropology* 109 (4): 439–54.

Mitani, John C., David P. Watts, and Sylvia J. Amsler. 2010. “Lethal Intergroup Aggression Leads to Territorial Expansion in Wild Chimpanzees.” *Current Biology* 20 (12): R507–8.

Miththapala, S, J Seidensticker, and SJ O’Brian. 1996. “Phylogeographic Subspecies Recognition in Leopards (Panthera Pardus): Molecular Genetic Variation.” Vol. 10.

Monroe, Burt L. 1982. “A Modern Concept of the Subspecies.” *The Auk: Ornithological Advances* 99 (3): 608–9.

Montgomery, Sy. 1991. *Walking with the Great Apes : Jane Goodall, Dian Fossey, Biruté Galdikas*. Houghton Mifflin Co.

Morell, V. 1993. “Called `Trimates,’ Three Bold Women Shaped Their Field.” *Science* 260 (5106): 420–25.

Mulavwa, M. et al. 2008. “Seasonal Changes in Fruit Production and Party Size of Bonobos at Wamba.” In *The Bonobos*, 121–34. New York, NY: Springer New York.

Nakamura, Masahiko. 1990. “Cloacal Protuberance and Copulatory Behavior of the Alpine Accentor (*Prunella Collaris*).” *The Auk* 107 (2): 284–95.

Nei, M. 2005. “Selectionism and Neutralism in Molecular Evolution.” *Molecular Biology and Evolution* 22 (12): 2318–42.

Nei, M, and G V Glazko. 2002. “Estimation of Divergence Times for a Few Mammalian and Several Primate Species.” *The American Genetic Association*. Vol. 93.

Nei, Masatoshi, Takeo Maruyama, and Ranajit Chakraborty. 1975. “The Bottleneck Effect and Genetic Variability in Populations.” *Evolution* 29 (1): 1–10.

Nielsen, R., and Ziheng Yang. 2003. “Estimating the Distribution of Selection Coefficients from Phylogenetic Data with Applications to Mitochondrial and Viral DNA.” *Molecular Biology and Evolution* 20 (8): 1231–39.

Nishida, T. 1983. “Alpha Status and Agonistic Alliance in Wild Chimpanzees (*Pan Troglodytes Schweinfurthii*).” *Primates* 24 (3): 318–36.

Nishida, Toshisada. 1968. “The Social Group of Wild Chimpanzees in the Mahali Mountains.” *Primates* 9 (3): 167–224.

Nishida, Toshisada, Nadia Corp, Miya Hamai, Toshikazu Hasegawa, Mariko Hiraiwa-Hasegawa, Kazuhiko Hosaka, Kevin D. Hunt, et al. 2003. “Demography, Female Life History, and Reproductive Profiles among the Chimpanzees of Mahale.” *American Journal of Primatology* 59 (3): 99–121.

Noordwijk, Maria A. von. 1985. “Sexual Behaviour of Sumatran Long-Tailed Macaques (Macaca Fascicularis).” *Zeitschrift Für Tierpsychologie* 70 (4): 277–96.

Nunn, Charles L. 1999. “The Evolution of Exaggerated Sexual Swellings in Primates and the Graded-Signal Hypothesis.” *Animal Behaviour* 58 (2): 229–46.

Ohta, T. 1972. “Population Size and Rate of Evolution.” *Journal of Molecular Evolution* 1 (4): 305–14.

———. 1973. “Slightly Deleterious Mutant Substitutions in Evolution.” *Nature* 246 (5428): 96–98.

———. 2013. “Molecular Evolution: Nearly Neutral Theory.” In *ELS*. Chichester, UK: John Wiley & Sons, Ltd.

Ohta, T, and JH Gillespie. 1996. “Development of Neutral and Nearly Neutral Theories.” *Theoretical Population Biology* 49 (2): 128–42.

Ozasa, Hiroshi, and Kenneth G. Gould. 1982. “Demonstration and Characterization of Estrogen Receptor in Chimpanzee Sex Skin: Correlation between Nuclear Receptor Levels and Degree of Swelling.” *Endocrinology* 111 (1): 125–31.

Pagel, Mark. 1994. “The Evolution of Conspicuous Oestrous Advertisement in Old World Monkeys.” *Animal Behaviour* 47 (6): 1333–41.

Palumbi, S R. 1994. “Genetic Divergence, Reproductive Isolation, and Marine Speciation.” *Annual Review of Ecology and Systematics* 25 (1): 547–72.

Parker, GA. 1970. “Sperm Competition and Its Evolutionary Consequences in the Insects.” *Biological Reviews* 45 (4): 525–67.

Patten, Michael A. 2015. “Subspecies and the Philosophy of Science.” *The Auk* 132 (2): 481–85.

Peischl, S., I. Dupanloup, M. Kirkpatrick, and L. Excoffier. 2013. “On the Accumulation of Deleterious Mutations during Range Expansions.” *Molecular Ecology* 22 (24): 5972–82.

Peter M. Kappeler. 2012. “Male Reproductive Strategies.” 2012.

Poulsen, John R., and Connie J. Clark. 2004. “Densities, Distributions, and Seasonal Movements of Gorillas and Chimpanzees in Swamp Forest in Northern Congo.” *International Journal of Primatology* 25 (2): 285–306.

Pozzi, Luca, Jason A. Hodgson, Andrew S. Burrell, Kirstin N. Sterner, Ryan L. Raaum, and Todd R. Disotell. 2014. “Primate Phylogenetic Relationships and Divergence Dates Inferred from Complete Mitochondrial Genomes.” *Molecular Phylogenetics and Evolution* 75 (June): 165–83.

Prado-Martinez, Javier, Peter H Sudmant, Jeffrey M Kidd, Heng Li, Joanna L Kelley, Belen Lorente-Galdos, Krishna R Veeramah, et al. 2013. “Great Ape Genetic Diversity and Population History.” *Nature* 499 (7459): 471–75.

Pruetz, Jill D., and Paco Bertolani. 2007. “Savanna Chimpanzees, *Pan Troglodytes Verus*, Hunt with Tools.” *Current Biology* 17 (5): 412–17.

Prüfer, Kay, Kasper Munch, Ines Hellmann, Keiko Akagi, Jason R Miller, Brian Walenz, Sergey Koren, et al. 2012. "The Bonobo Genome Compared with the Chimpanzee and Human Genomes." *Nature* 486 (7404): 527–31.

Prüfer, Kay, Fernando Racimo, Nick Patterson, Flora Jay, Sriram Sankararaman, Susanna Sawyer, Anja Heinze, et al. 2014. "The Complete Genome Sequence of a Neanderthal from the Altai Mountains." *Nature* 505 (7481): 43–49.

Rand, D M, and L M Kann. 1996. "Excess Amino Acid Polymorphism in Mitochondrial DNA: Contrasts among Genes from Drosophila, Mice, and Humans." *Molecular Biology and Evolution* 13 (6): 735–48.

Ritchie, Graham R S, Ian Dunham, Eleftheria Zeggini, and Paul Flicek. 2014. "Functional Annotation of Noncoding Sequence Variants." *Nature Methods* 11 (3): 294–96.

Roberts, S. C., J. Havlicek, J. Flegr, M. Hruskova, A. C. Little, B. C. Jones, D. I. Perrett, and M. Petrie. 2004. "Female Facial Attractiveness Increases during the Fertile Phase of the Menstrual Cycle." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271 (suppl_5).

Robertson, Alan. 1965. "The Interpretation of Genotypic Ratios in Domestic Animal Populations." *Animal Production* 7 (03): 319–24.

Rogers, Rebekah L., and Montgomery Slatkin. 2017. "Excess of Genomic Defects in a Woolly Mammoth on Wrangel Island." Edited by Gregory S. Barsh. *PLOS Genetics* 13 (3): e1006601.

Romero Zarco, Carlos Manuel. 2018. *The IUCN Red List of Threatened Species 2018*. International Union for Conservation of Nature - IUCN.

Rose, Kenneth David. 2006. *The Beginning of the Age of Mammals.* Johns Hopkins University Press.

Ryu, H, D Hill, and T Furuichi. 2015. "Prolonged Maximal Sexual Swelling in Wild Bonobos Facilitates Affiliative Interactions between Females." *Behaviour* 152 (3–4): 285–311.

Sakura, Osamu. 1994. "Factors Affecting Party Size and Composition of Chimpanzees (*Pan Troglodytes Verus*) Bossou, Guinea." *International Journal of Primatology* 15 (2): 167–83.

Sarich, V M, and A C Wilson. 1967. "Immunological Time Scale for Hominid Evolution." *Science (New York, N.Y.)* 158 (3805): 1200–1203.

Schwensow, Nina, Manfred Eberle, and Simone Sommer. 2007. "Compatibility Counts: MHC-Associated Mate Choice in a Wild Promiscuous Primate." *Proceedings of the Royal Society B: Biological Sciences* 275 (1634): 555–64.

Sequencing, The Chimpanzee, and Analysis Consortium. 2005. "Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome." *Nature* 437 (7055): 69–87.

Serckx, Adeline, Marie-Claude Huynen, Jean-François Bastin, Alain Hambuckers, Roseline C. Beudels-Jamar, Marie Vimond, Emilien Raynaud, and Hjalmar S. Kühl. 2014. "Nest Grouping Patterns of Bonobos (*Pan Paniscus*) in Relation to Fruit Availability in a Forest-Savannah Mosaic." Edited by Cédric Sueur. *PLoS ONE* 9 (4): e93742.

Setchell, JM, MJE Charpentier, KM Abbott, EJ Wickings, and LA Knapp. 2009. "Opposites Attract: MHC-Associated Mate Choice in a Polygynous Primate." *Journal of Evolutionary Biology* 23 (1): 136–48.

Shea, B T. 1983. “Paedomorphosis and Neoteny in the Pygmy Chimpanzee.” *Science (New York, N.Y.)* 222 (4623): 521–22.
Simpson, G. G. 1963. *The Meaning of Taxonomic Statements.*
Simpson, G. G. 1964. “Organisms and Molecules in Evolution.” *Science* 146: 1535–38.
Skov, L. et al. 2018. “Detecting archaic introgression using an unadmixed outgroup.” *PLoS Genet.* 14, e1007641.
Smith, John Maynard. 1991. “Theories of Sexual Selection.” *Trends in Ecology & Evolution* 6 (5): 146–51.
Stanford, Craig B. 1998. “The Social Behavior of Chimpanzees and Bonobos: Empirical Evidence and Shifting Assumptions.” *Current Anthropology* 39 (4): 399–420.
Stevens, N. J. et al. 2013. “Palaeontological Evidence for an Oligocene Divergence between Old World Monkeys and Apes.” *Nature* 497 (7451): 611–14.
Stoletzki, N., and Eyre-Walker, A. 2011. “Estimation of the Neutrality Index.” *Molecular Biology and Evolution* 28 (1): 63–70.
Subramanian, Sankar, and David M Lambert. 2012. “Selective Constraints Determine the Time Dependency of Molecular Rates for Human Nuclear Genomes.” *Genome Biology and Evolution* 4 (11): 1127–32.
Sugiyama, Yukimaru. 1984. “Some Aspects of Infanticide and Intermale Competition among Langurs, *Presbytis Entellus*, at Dharwar, India.” *Primates* 25 (4): 423–32.
Sugiyama, Yukimaru, and Jeremy Koman. 1987. “A Preliminary List of Chimpanzees’ Alimentation at Bossou, Guinea.” *Primates* 28 (1): 133–47.
Surbeck, Martin, Kevin E. Langergraber, Barbara Fruth, Linda Vigilant, and Gottfried Hohmann. 2017. “Male Reproductive Skew Is Higher in Bonobos than Chimpanzees.” *Current Biology* 27 (13): R640–41.
Takemoto, Hiroyuki, Yoshi Kawamoto, Shoko Higuchi, Emiko Makinose, John A Hart, Térese B Hart, Tetsuya Sakamaki, et al. 2017. “The Mitochondrial Ancestor of Bonobos and the Origin of Their Major Haplogroups.” *PloS One* 12 (5): e0174851.
Teixeira, J. C., C. de Filippo, A. Weihmann, J. R. Meneu, F. Racimo, M. Dannemann, B. Nickel, et al. 2015. “Long-Term Balancing Selection in *LAD1* Maintains a Missense Trans-Species Polymorphism in Humans, Chimpanzees, and Bonobos.” *Molecular Biology and Evolution* 32 (5): 1186–96.
Thompson-Handler, Nancy, Richard K. Malenky, and Noel Badrian. 1984. “Sexual Behavior of *Pan Paniscus* under Natural Conditions in the Lomako Forest, Equateur, Zaire.” In *The Pygmy Chimpanzee*, 347–68. Boston, MA: Springer US.
Thompson, M. 1997. “The History, Taxonomy and Ecology of the Bonobo (Pan Paniscus Schwarz, 1929) with a First Description of a Wild Population Living in a Forest/Savanna Mosaic Habitat.”
———. 2002. “The Status of Bonobos in Their Southernmost Geographic Range.” In *All Apes Great and Small*, 75–81. Boston: Kluwer Academic Publishers.

Thornhill, Randy. 1983. “Cryptic Female Choice and Its Implications in the Scorpionfly *Harpobittacus Nigriceps*.” *The American Naturalist*. The University of Chicago PressThe American Society of Naturalists.

Torkamani, Ali, Phillip Pham, Ondrej Libiger, Vikas Bansal, Guangfa Zhang, Ashley A Scott-Van Zeeland, Ryan Tewhey, Eric J Topol, and Nicholas J Schork. 2012. “Clinical Implications of Human Population Differences in Genome-Wide Rates of Functional Genotypes.” *Frontiers in Genetics* 3: 211.

Trivers, R. 1972. “Parental Investment and Sexual Selection.” In *Sexual Selection and the Descent of Man*.

Tsuneko, Onouchi, and Kato Junzo. 1983. “Estrogen Receptors and Estrogen-Inducible Progestin Receptors in the Sexual Skin of the Monkey.” *Journal of Steroid Biochemistry* 18 (2): 145–51.

Uphyrkina, Olga, Warren E. Johnson, Howard Quigley, Dale Miquelle, Laurie Marker, Mitchel Bush, and Stephen J. O’Brien. 2001. “Phylogenetics, Genome Diversity and Origin of Modern Leopard, Panthera Pardus.” *Molecular Ecology* 10 (11): 2617–33.

van der Valk, T, Díez-del-Molino D, Marques- Bonet T, Guschanski K, Dalén L. 2019. “Historical Genomes Reveal the Genomic Consequences of Recent Population Decline in Eastern Gorillas.” *Curr Biol.* 29: 165–170.

Valles-Ibáñez, Guillem de, Jessica Hernandez-Rodriguez, Javier Prado-Martinez, Pierre Luisi, Tomàs Marquès-Bonet, and Ferran Casals. 2016. “Genetic Load of Loss-of-Function Polymorphic Variants in Great Apes.” *Genome Biology and Evolution*.

Veeramah, K. R., A. E. Woerner, L. Johnstone, I. Gut, M. Gut, T. Marques-Bonet, L. Carbone, J. D. Wall, and M. F. Hammer. 2015. “Examining Phylogenetic Relationships among Gibbon Genera Using Whole Genome Sequence Data Using an Approximate Bayesian Computation Approach.” *Genetics* 200 (1): 295–308.

Vernot, B. and Akey, J. M. 2014. “Resurrecting surviving Neandertal lineages from modern human genomes.” *Science* 343, 1017–1021.

Vervaecke, H., H. de Vries, and L. van Elsacker. 2000. “Dominance and Its Behavioral Measures in a Captive Group of Bonobos (*Pan Paniscus*).” *International Journal of Primatology* 21 (1): 47–68.

de Waal, F. 1982. Chimpanzee Politics: Power and Sex among Apes. Vol. 10.

de Waal, F. 1987. “Tension Regulation and Nonreproductive Functions of Sex in Captive Bonobos *Pan Paniscus*.” *National Geographic Research* 3 (3): 318–35.

Wallis, Janette. 1992. “Chimpanzee Genital Swelling and Its Role in the Pattern of Sociosexual Behavior.” *American Journal of Primatology* 28 (2): 101–13.

Wayne, R K. 1993. “Molecular Evolution of the Dog Family.” *Trends in Genetics : TIG* 9 (6): 218–24.

West, Neal B, and Robert M Brenner. 1990. “Estrogen Receptor in the Ductuli Efferentes, Epididymis, and Testis of Rhesus and Cynomolgus Macaques.” *Biology of Reproduction*. Vol. 42.

White, Frances J. 1988. “Party Composition and Dynamics in *Pan Paniscus*.” *International Journal of Primatology* 9 (3): 179–93.

Whiten, A., J. Goodall, W. C. McGrew, T. Nishida, V. Reynolds, Y. Sugiyama, C. E. G. Tutin, R. W. Wrangham, and C. Boesch. 1999. “Cultures in Chimpanzees.” *Nature* 399 (6737): 682–85.

Williams, J. M., E. V. Lonsdorf, M. L. Wilson, J. Schumacher- Stankey, J. Goodall, and A. E. Pusey. 2008. “Causes of Death in the Kasekela Chimpanzees of Gombe National Park, Tanzania.” *American Journal of Primatology* 70 (8): 766–77.

Wilson, Don E., and DeeAnn M. Reeder. 2005. *Mammal Species of the World : A Taxonomic and Geographic Reference.* Johns Hopkins University Press.

Wilson, E. O., and W. L. Brown. 1953. “The Subspecies Concept and Its Taxonomic Application.” *Systematic Zoology* 2 (3): 97.

Wilson, M L, C Boesch, B Fruth, T Furuichi, I C. Gilby, C Hashimoto, C L Hobaiter, et al. 2014. “Lethal Aggression in *Pan* Is Better Explained by Adaptive Strategies than Human Impacts.” *Nature* 513 (7518): 414–17.

Wilson, M, and R Wrangham. 2003. “Intergroup Relations in Chimpanzees.” *Annu. Rev. Anthropol* 32: 363–92.

Won, Yong-Jin, and Jody Hey. 2005. “Divergence Population Genetics of Chimpanzees.” *Molecular Biology and Evolution* 22 (2): 297–307.

Wrangham, R. 1993. “The Evolution of Sexuality in Chimpanzees and Bonobos.” *Human Nature* 4 (1): 47–79.

———. 1996. *Chimpanzee Cultures.* Published by Harvard University Press in cooperation with the Chicago Academy of Sciences.

———. 2002. “The Cost of Sexual Attraction: Is There a Trade-off in Female *Pan* between Sex Appeal and Received Coercion?” In *Behavioural Diversity in Chimpanzees and Bonobos*, edited by Christophe Boesch and Gottfried Hohmann, 204–16. Cambridge: Cambridge University Press.

———. 2018. “Two Types of Aggression in Human Evolution.” *Proceedings of the National Academy of Sciences of the United States of America* 115 (2): 245–53.

Wrangham, R, and D Pilbeam. 2001. “African Apes as Time Machines.” In *All Apes Great and Small*, 5–17. Boston: Kluwer Academic Publishers.

Wright, S. 1931. “Evolution in Mendelian Populations.” *Genetics* 16 (2): 97–159.

Wroblewski, Emily E, Carson M Murray, Brandon F Keele, Joann C Schumacher-Stankey, Beatrice H Hahn, and Anne E Pusey. 2009. “Male Dominance Rank and Reproductive Success in Chimpanzees, *Pan Troglodytes Schweinfurthii*.” *Animal Behaviour* 77 (4): 873–85.

Yoder, Anne D, Jelmer W Poelstra, George P Tiley, and Rachel C Williams. 2018. “Neutral Theory Is the Foundation of Conservation Genetics.” Edited by Sudhir Kumar. *Molecular Biology and Evolution* 35 (6): 1322–26.

Zimmer, C, and D Emlen. 2013. *Evolution : Making Sense of Life.*

Zuckerkandl, E, and L Pauling. 1965. “Evolutionary Divergence and Convergence in Proteins.” In *Evolving Genes and Proteins*, 97–166. Elsevier.

Zuckerman, S, and A Parkes. 1939. “Observations on Secondary Sexual Characters in Monkeys.” *Journal of Endocrinology* 1 (1): 430–39.