Author/s:
Dannemann, M;Gallego Romero, I

Title:
Harnessing pluripotent stem cells as models to decipher human evolution

Date:
2022-06-01

Citation:
Dannemann, M. & Gallego Romero, I. (2022). Harnessing pluripotent stem cells as models to decipher human evolution. FEBS Journal, 289 (11), pp.2992-3010. https://doi.org/10.1111/febs.15885.

DR MICHAEL DANNEMANN (Orcid ID : 0000-0002-7076-8731)
DR IRENE GALLEGO ROMERO (Orcid ID : 0000-0003-1613-8998)

Received Date : 31-Jan-2021
Revised Date : 18-Mar-2021
Accepted Date : 16-Apr-2021
Article type : State-of-the-Art Review

Title: Harnessing pluripotent stem cells as models to decipher human evolution

Michael Dannemann[1], Irene Gallego Romero[1-4]

1. Center for Genomics, Evolution and Medicine, Institute of Genomics, University of Tartu, Riia 23b, 51010 Tartu, Estonia
2. Melbourne Integrative Genomics, University of Melbourne, Royal Parade, 3010, Parkville, Victoria, Australia
3. School of BioSciences, The University of Melbourne, Royal Parade, 3010, Parkville, Australia
4. The Centre for Stem Cell Systems, Faculty of Medicine, Dentistry and Health Sciences, The University of Melbourne, 30 Royal Parade Parkville, Victoria 3010, Australia

Correspondence:
michael.dannemann@ut.ee
irene.gallego@unimelb.edu.au

Running title: Pluripotent stem cell models for human evolution

Abbreviations:

bp: base-pair

eQTL: expression quantitative trait locus

ESC: embryonic stem cell

GWAS: genome-wide association study

iPSC: induced pluripotent stem cell

LD: linkage disequilibrium

MPRA: massively parallel reporter assay

SNP: single nucleotide polymorphism

Keywords:

Stem cells, Human evolution, Tissue models, development, gene expression

Conflicts of interest

The authors declare no competing interest

Abstract:

The study of human evolution, long constrained by a lack of experimental model systems, has been transformed by the emergence of the induced pluripotent stem cell (iPSC) field. iPSCs can be readily established from non-invasive tissues sources, both from humans and other primates; they can be maintained in the laboratory indefinitely and they can be differentiated into other tissue types. These qualities mean that iPSCs are rapidly becoming established as viable and powerful model systems with which it is possible to address questions in human evolution that were until now logistically and ethically intractable, especially in the quest to understand humans' place amongst the great apes, and the genetic basis of human uniqueness. In this review, we discuss the key lessons and takeaways of this nascent field; from the types of research iPSCs make possible to lingering challenges and likely future directions. We provide a comprehensive overview of how the seemingly unlikely combination of iPSCs and explicit evolutionary frameworks are transforming what is possible in our understanding of humanity's past and present.

Introduction

Thorough knowledge of humanity's evolutionary past is essential to understanding our present and future. But the full potential of human evolutionary biology studies has long been constrained by a lack of tractable model systems with which to explore this field. The publication of complete genome sequences from humans and great apes has demonstrated that the vast majority of DNA sequence differences separating humans from other primates occur outside the protein-coding

regions of the genome [1], and suggested that inter-species differences are likely to arise through effects on gene regulation [2,3]. Although the relative contributions of gene regulatory and protein coding change to evolutionary differences remains a matter of some debate [4–6], these observations have required substantial reconsideration—and creativity—on how best to approach studies of human evolution, whether at the species level, or when examining more subtle differences between present-day populations. Given the complexities that still surround *in silico* predictions of function from non-coding sequence, deciphering the genetic basis of human uniqueness remains an arduous process, one that frequently requires access torare tissue samples for experimental validation.

But this need is not easily met. Humans are one of eight extant great ape species [7]. All non-human great apes—chimpanzees, bonobos, gorillas and orangutans—are either endangered or critically endangered in the wild. Of these, only chimpanzees (*Pan troglodytes*, humanity's closest living relative) have ever been widely adopted as laboratory research animals, but numbers were already declining in 1995, when the USA's National Institutes of Health enacted a funding moratorium on their breeding [8]. Given their low numbers and complex cognitive abilities, invasively obtaining samples from wild apes, or capturing them for research purposes, is unethical. Until recently, molecular inter-species studies of human evolution were therefore chiefly reliant on post-mortem tissue samples collected from captive animals dying of unrelated causes in research centres. And while these samples have facilitated many significant insights (reviewed in [2]) they also suffered from some non-negligible shortcomings: limited sample availability, compounded by difficulties in matching or staging samples. Crucially, the fact that these samples are often collected post-mortem and can therefore not be experimentally manipulated, has largely restricted the field to observational insights. Additionally, existing samples are mainly derived from adult animals, limiting insights into differences between the species in earlier developmental time-points.

Likewise, achieving a thorough understanding of the phenotypic differences that exist between present-day human populations, and the genetic and biological mechanisms that underpin them, remains another significant challenge in the study of human evolution. The shorter time scales (roughly 60,000 years, reviewed below) that characterise humanity's expansion throughout the world and adaptation to all manner of local environments, or the contributions of introgression from archaic hominin groups such as Neandertals and Denisovans, have all received significant attention in the past decades. But as growing numbers of genome-wide studies have shown, the genetic complexity of human traits is not easily described [9,10]. Well characterised samples that span the

wealth and breadth of existing human genetic variation are essential to fully understanding how evolution has shaped our species.

Recent technological developments make these questions more tractable today. Here we review the potential and growing application of induced pluripotent stem cells (iPSCs) as an emerging model system with which to tackle questions of human evolution that would otherwise prove unanswerable. The ease with which they can be established, as well as their tantalisingly versatility to generate increasingly faithful and complex models from a growing number of tissues, in combination with the development of new high-throughput technologies and methodologies provide unprecedented avenues to study how genetic variation influences human-specific biology. These advances are transforming studies of human evolution, and shedding new light into the past and present of *Homo sapiens*.

The nature of human evolution

Approximately 7 million years, around 35 million base pair differences and countless genomic rearrangements separate humans from our closest living relatives, the chimpanzee (Figure 1A) [11]. Decades of work to reconstruct the human fossil record [12] have painted a comprehensive picture of the time since then, with the first anatomically modern humans—that is, recognisably *Homo sapiens*—emerging in Africa between 300,000 and 200,000 years ago [13]. Following successful expansion throughout the continent, humans then began expanding outwards more than 60,000 years ago [14,15]. Both the fossil record and genetic data demonstrate that *Homo sapiens* had reached Australia at least 55,000 years ago [16,17], and America sometime between 33,000 and 15,000 years ago [18,19], settling all continents except Antarctica by then. Complex population dispersals, migrations and replacements over the past 60,000 years have been revealed by large scale sequencing efforts of present-day individuals, as well as by a growing collection of DNA sequences generated from fossil samples [20].

Given these trends, studies of human evolution often focus on one of two questions: 1. human-unique traits, which are common to all humans, and emerged after the split of humans and chimpanzees, or 2. population-specific differences that are restricted to only some humans, such as the ability to digest the milk sugar lactose after weaning, or the ability of some populations to thrive at high altitude [21,22]. Many of the latter have emerged since humans expanded out of Africa, and therefore serve as examples of adaptation to new environments or cultural practises. Others are driven by the varying challenges presented by expansion into novel environments, for instance with regards to pathogen diversity [23].

Additional complexity comes from high-quality complete genome sequences from two extinct hominin groups, Neandertals [24–26] and Denisovans [27], both of whom last shared a common ancestor with humans roughly 700,000 years ago. These archaic genomes have conclusively demonstrated that gene flow between *H. sapiens* and the other two archaic groups occurred more than once following humanity's dispersal from Africa. All present-day individuals of non-African descent carry between 1-2% of Neandertal DNA in their genomes [28], while Denisovan DNA is more geographically structured [29] with its highest frequency in the present-day Indigenous Peoples of Papua New Guinea and Australia (4-6% of their genomes), but is also widespread throughout much of mainland Asia (Figure 1A) [30]. As introgressed DNA from both archaic human groups has been shown to be an additional source of functional variation that is contributing to local adaptation and phenotypic variation in people today [31,32], these findings raise a third question: what were the contributions of these archaic hominins to present-day humans?

No matter the time scale, the bulk of our insights on human evolutionary variation come from studies of DNA sequence. There is a wealth of well-established methods for identifying signatures of positive selection (Figure 1B) and evolutionary change at the genome level [33,34], understanding their implications has proven a lot more difficult—functional validation remains rare. This may partly stem from historical boundaries between population genetics and molecular biology, but it also reflects the difficulty of fully resolving the biological underpinnings of evolutionary change at the molecular level. Much phenotypic diversity appears to be driven not by a single change in a single locus, but by small changes at many loci across the genome that jointly contribute to a single complex phenotype [9,35]. Yet, to be successfully studied in a laboratory, complex traits observed at the organismal level have to be reduced to something both tractable and meaningful *in vitro*. It is this niche where iPSC-derived model systems will likely have the greatest impact.

A brief introduction to induced pluripotent stem cells

The development of induced pluripotent stem cells (iPSC) 15 years ago [36] has revolutionised fields ranging from regenerative medicine to comparative functional genomics [37,38] and stand poised to do the same for studies of human evolution. In 2006 Takahashi and Yamanaka first demonstrated that through the activation of four transcription factors, which today are collectively referred to as Yamanaka factors [39,40], terminally differentiated somatic cells could be forced to revert to a much earlier developmental state—pluripotency, the ability to give rise to any cell type in an animal. Indeed, iPSCs resemble embryonic stem cells (ESCs), one of the earliest cells

observed during development, in their transcriptional profiles and developmental potentials. This means that they can be maintained *in vitro* indefinitely, and differentiated into other cell types by the activation of key developmental pathways.

Multiple methods of generating iPSCs have been reported since then, but they all rely on the same fundamental mechanism: the activation of the Yamanaka factors. Crucially, the starting material for iPSCs includes cell types that can be collected in a minimally invasive manner, including dermal fibroblasts [36], blood cells [41,42] or even cells shed in urine [43], meaning that iPSCs can be readily generated from specific donors and circumvent the bulk of the ethical concerns associated with ESCs (Figure 2A). The differentiation potential of iPSCs is broadly similar to that of ESCs [44,45], and multiple studies have shown that they faithfully recapitulate the impact of genetic differences between individuals, making them valuable *in vitro* models [46–50].

Initially, iPSC generation was limited to either humans or laboratory animals. However, the establishment of iPSCs from non-model animals like the critically endangered northern white rhinoceros or the drill macaque, in both cases using the human Yamanaka factors sequences (although some ambiguity remains over whether optimal conditions and reprogramming factor cocktail are shared across taxa), conclusively demonstrated the robustness of cellular reprogramming [51]. Today, iPSC lines from mammals ranging from the platypus [52] to various commercially important livestock species [53] have been reported, with recent reports of urine-derived iPSC lines from gorilla and orangutan [54]. Potential uses for these iPSC lines range from conservation and recovery of endangered species to the production of genetically engineered animals, again showcasing their versatility [55].

Using iPSCs to model complex biology *in vitro*

However, the true potential of iPSCs stems from their ability to differentiate into other cell types. This makes them an invaluable resource, particularly in instances where relevant cell types are hard to obtain via other sources, including rare cell types or transient developmental intermediates. Fundamentally, iPSC differentiation involves the exogenous activation of specific transcription factors and gene expression modules that will drive the pluripotent cell towards the desired cell type [56], although protocols for iPSC differentiation vary vastly in their efficiency, duration, cost and in their ability to replicate the target cell type.

Early differentiation protocols gave rise to cells that grew 2-dimensionally along the bottom of tissue culture plates, and could therefore not fully represent the complexity of a developing tissue.

Nonetheless, many protocols sought to explicitly mirror *in vivo* development, progressing in discrete steps from pluripotency to commitment towards one of the three developmental germ layers and then further specification of terminal fate, a process that can span multiple weeks. By explicitly recapitulating *in utero* development, these protocols provide insights into otherwise unobservable processes in humans. However, the last decade has seen the development of protocols to differentiate stem cells into self-organising, organ-specific 3-dimensional structures, called organoids, which exhibit organ-like properties, and this has drastically advanced the ability to study developmental processes in a controlled environment. Protocols to generate more than 10 different types of iPCS-derived organoids have been established by now [57,58], and their complexity continues to evolve (Figure 2A). Organoid systems recapitulate the key transcriptomic features of early stages of organ development observed *in vivo* [57,59–64], and to date have been used mainly as disease models, particularly in cases with an early disease onset and cancer research [65–68]. Compared to 2D cultures, organoids provide a model system that more faithfully captures the physiological complexity of an organ and allows the study of developmental processes in a cell-type specific manner. For example, 2D culture does not provide control over cell shape and can induce atypical cell polarities that are usually not found *in vivo* [69]. Despite their differences, both 2D and 3D cultures provide valuable model systems to study cellular processes in a controlled environment, especially for cell types that are inaccessible through other means. iPSC-derived models can dramatically facilitate the study of morphological and transcriptomic profiles of differentiated cell types and establish links to possible phenotypic implications.

However, iPSC-derived cells also warrant some notes of caution. As a general rule, iPSC-derived tissues ressemble foetal, rather than adult, tissue in their gene expression patterns and metabolism. This phenomenon has been described in multiple cell types [70–75]; one of the clearest examples is iPSC-derived cardiomyocytes, which *in vivo* undergo a metabolic switch upon birth from using lactate to pyruvate for energy generation, while iPSC-derived cardiomyocytes thrive in lactate by default [76,77]. Additionally, some iPSC lines can refractory to differentiation into certain developmental lineages. Understanding why inter-individual differences manifest in this way, and whether they stem from effects acquired during the reprogramming process or other variables remains an open question [78–82].

Despite improvements, differentiation protocols consistently yield heterogeneous cell mixtures, due to both technical [83] and biological factors [84]. No two organoids or differentiation plates are identical, containing variable fractions of cells at different points along developmental trajectories, and these differences can have significant impacts on estimates of gene expression if aggregated

across cells (reviewed by [85]). Therefore, iPSC-derived cells are often analysed through single-cell sequencing, which captures cellular heterogeneity of the transcriptional profiles by generating data from individual cells [86]. Recent refinements that allow the simultaneous measurement of gene expression and gene regulatory activity—for example, chromatin accessibility or DNA methylation—from the same cell are driving the development of comprehensive regulatory profiles for each cell [87–91].

When working with non-model organisms, iPSCs therefore provide a suite of benefits. Many experimental covariates that are hard to address when working with opportunistically collected post-mortem samples can be successfully controlled in this setting. This includes ensuring a sex-balanced sample, similar passage number (instead of large age differences between donors), and the ability to limit inter-individual differences in exposures to environmental covariates such as diet. iPSC-derived cells can be purified via flow cytometry or manipulation of growth conditions to achieve comparable purity across samples [92]. Because they are alive, iPSCs and iPSC-differentiated cells can also be experimentally perturbed in a controlled setting. In combination with single-cell sequencing, cells from multiple individuals can even be pooled and cultured together in a single dish and computationally deconvoluted at the end of the experiment [93,94], thus minimising inter-individual differences during the experiment (such as differences in media acidification or handling batch effects) that may confound results.

Inter-species insights from iPSC-derived models for human evolution

Since its inception, the focus of the iPSC field has remained squarely biomedical. The advances detailed above were never developed with the explicit goal of testing evolutionary hypotheses, so it is the ease with which these approaches can be adopted that truly demonstrates the potential of iPSCs beyond the scope of regenerative medicine. Given the close evolutionary relationship between humans and chimpanzees, much of the focus of the field has been in establishing well-characterised lines from the latter, with cell lines from approximately 20 different chimpanzee individuals having been generated by researchers worldwide [95–103] and smaller numbers from other great apes [54,98,100,102,104,105] and other, more distantly related primates (for example, [51,106–108]). The establishment of these lines has allowed researchers to leverage their potential and incorporate them in functional studies explicitly designed to compare molecular processes between humans and other non-human primates.

For instance, multiple groups have described inter-species differences in gene expression and regulation at the pluripotent stage that could have bearing on adult phenotypes [95,103,104,109–

112]. Marchetto et al. [95] reported differences in activity in LINE-1 retrotransposons between 4 human iPSC lines and 2 chimpanzee and 2 bonobo lines. Likewise, comparing 7 human and 7 chimpanzee cell lines, Gallego Romero et al. [103] found that the transcription factor Rex1/ZFP42 appeared dispensable for maintaining pluripotency in chimpanzee iPSCs. By examining chromatin accessibility and gene expression across the two species, they also reported that a substantial fraction of differentially accessible regions of the genome overlapped known bivalently modified genes, which play fundamental roles in early development [110]. Differences in chromatin structure in iPSCs of humans and chimpanzees were also reported by Eres et al. [111], again suggesting that differences between the two species are established early in development. All of these studies additionally identified 1,000s of differentially expressed genes between the two species, although often the effect size of the difference was small, and their contribution to adult phenotypes remains undetermined.

Beyond the pluripotent state, a growing number of studies has compared iPSC-differentiated cells from humans and other evolutionarily relevant species (Figure 2B), again considering both gene expression and regulation — although there are significantly fewer datasets available that speak to the latter. While the phenotypic difference between humans and other great apes that has drawn the most focus is the advanced cognitive skills that are unique to humans [113,114], which we discuss below, comparative studies have focused on an array of organs and cell types across all developmental tissue layers [101,109,112,115–118]. Much of this work incorporates specific developmental questions or experimental manipulation that would be beyond the scope of non-iPSC models. For example, Ward et al. [115] exposed human and chimpanzee iPSC-derived cardiomyocytes to varying oxygen levels to simulate ischemic reperfusion injury, asking whether inter-species differences in the response could underlie the known differences in pathogenesis of cardiovascular disease in humans and the other great apes [119], but instead finding that it is broadly conserved across the two species. Meanwhile, Prescott et al. examined gene expression and regulation in cranial neural crest cells from humans and chimpanzees to identify differences in craniofacial morphology development across the two, and identified genome-wide changes in transcription factor activity at this early developmental stage of possible downstream consequence, as well as a novel DNA binding motif that could suggest rewiring of developmental networks across species [118]. Finally, Cugola et al. [120] examined the effect of infection with two different strains of Zika virus on human and chimpanzee cerebral organoids, and reported finding that a Brazilian strain adapted to human-to-human transmission was not able to infect chimpanzee organoids.

Recently, non-human primate iPSCs have also been used in combination with high-throughput screening technologies which allow the testing of the regulatory impact of large numbers of genetic variants at once. An early application of these has been to the study of human-accelerated regions (HARs), small genomic regions that are conserved throughout animals but divergent in humans, and thus suggestive of adaptive change. Two recent publications have made use of iPSC-derived cells to test the ability of these human-specific substitutions to drive gene expression compared to their ancestral states (Figure 3A, [121,122]), and both have identified multiple instances of differences in regulatory domains, and specific transcription factor sites, that were capable of altering gene expression levels significantly – at least *in vitro*. In another technological leap, two recent studies have reported the generation of human-chimpanzee hybrid tetraploid cell lines, and used this system to disentangle species-specific regulatory differences during craniofacial and neural development [123,124].

Specific insights into human cognitive differences

The genetic underpinnings of cognitive differences between humans and non-human primates are likely to be found in the brain, a tissue that is difficult to access in humans, and even more restricted in great apes. It is to this question, therefore, that inter-species iPSC models have been applied most frequently (see Mostajo-Radj et al. [125] for an extensive review of this field). In the absence of primary tissue, which is hard to obtain and stage, several studies have used iPSC-derived brain cell types using 2D culture differentiation protocols from different developmental time points in an attempt to explore differences that can be linked to human-specific cognition [122,126–133]. For example, Otani et al. made elegant use of the staged nature of a 2D cortical development protocol [134] to examine cortical cell proliferation across humans, chimpanzees and macaques, finding differences in the amount of time cortical progenitor cells remained proliferative across the three species [129] that could potentially be linked to differences in brain cell numbers between them. Likewise, in a more recent study, Kitajima et al. established a culture system of free floating clusters of neural stem cells, also referred to as neurosphere, from chimpanzee-derived iPSCs [126]. Using this system the authors were able to study early neural developmental processes and the underlying regulatory network involved in the formation of radial glia.

In recent years these efforts have been complemented by studies that have analyzed 3D brain organoids derived from human and chimpanzee iPSCs, comparing developmental processes at single cell resolutions to study human-specific brain biology [98,99,123,135–141]. For example, by studying the transcriptomic and chromatin accessibility landscapes in human, chimpanzee and macaque cerebral organoids, Kanton et al. showed that human neuronal development is delayed

compared to the other two primates, and that this difference could be partially attributed to differences in chromatin state between the two species [98]. In addition, the authors identified human-specific expression changes that they show to persist into adulthood. Using a similar approach Pollen et al. found differential expression patterns between human and chimpanzee organoids, some of which they link to the activation of the AKT-mTOR pathway via two up-regulated receptors in human radial glia [99].

The majority of these studies have focused on inter-species differences in gene expression, or fundamental cellular traits like proliferation or growth, reasoning that these have direct bearing on human cognitive skills (Figure 2B). Some of the mechanisms and observations we discuss above may well underlie these differences in cognitive complexity between humans and other primates. But it bears remembering that these are complex phenotypes that do not easily lend themselves to *in vitro* modelling. Interpretation of this first wave of results therefore remains tentative. While improvements in differentiation protocols will continue to come closer to replicating actual brain biology, there may come a time when the faithfulness of our models may necessitate a different type of caution: consideration of the ethical issues arising from increasingly complex cerebral organoid systems [142].

Examining evolutionary processes within human populations

Contrary to the above, the application of iPSC-based models to investigate intra-species variation in present-day humans in an explicitly evolutionary context has so far been limited. The main application of human iPSCs remains disease modelling, where small numbers of patient-derived iPSC lines are used. Often, this approach is complemented with genome editing techniques to test for phenotypic effects of disease-associated variants, both in 2D and 3D organoid models (see [143–145] for examples).

In parallel, the last few years have seen the generation of multiple population-scale iPSC panels. HipSci [49] contains nearly 500 cell lines of primarily European ancestry; iPSCORE [48] 222 lines from a diverse set of US-based donors, although still predominantly European. Banovich et al. [146] generated iPSC lines from 58 Yoruba individuals from the HapMap Project [147], a set of extremely well characterised individuals of Bantu ancestry. Other efforts have focused on generating iPSC banks from donors that carry the most common HLA haplotypes in the local population (for example, [148,149]), to facilitate donor compatibility in future medical applications. Of these, HipSci and iPSCORE are the most readily accessible by other researchers, being available through commercial cell line repositories.

The scope of these panels enables not only the mechanistic dissection of specific molecular pathways or patient-specific condition, but also the linking of genotype and phenotype through mapping of expression quantitative trait loci (eQTLs, Figure 3B), SNPs where differences in genotype across individuals are robustly associated to differences in gene expression levels [49,146,150–153]. For instance, Knowles et al. used iPSC-derived cardiomyocytes from 45 donors to identify a set of eQTLs that could predict cardiotoxicity upon exposure to anthracyclines [154], commonly used chemotherapeutic agents. Warren et al., meanwhile, established iPSC lines from 68 donors and used them to validate multiple candidate regions identified in a metabolic disease genome-wide association study (GWAS) [155]. Both Cuomo et al,. and Strober et al., leveraged the ability to use iPSCs to observe early human development [156,157], again identifying eQTLs active during endoderm and mesoderm differentiation. It has been proposed that these eQTLs, many of which are not active in adult tissues, may underlie some inter-individual differences in disease predisposition later in life [158]. Similar approaches to study variation in modern humans for their developmental and regulatory effects in an evolutionary context using organoid models alongside single cell sequencing (Figure 3C) have been proposed to be feasible model systems [159].

Although the motivation for most of these studies is not explicitly evolutionary, insights from them can nonetheless be interpreted through the lens of human population genetics or evolutionary change. This might appear counterintuitive, but ultimately, all of these studies seek to link differences in genotype to differences in phenotype. Changes in response to positive selection are only one of the mechanisms by which evolution proceeds; genetic drift, the gradual shift in allele frequencies between populations that occurs simply due to neutral processes, is a powerful force of itself, and the accumulated legacy of these evolutionary processes has biomedical implications today [160]. Therefore, instances where differences in genotype can be linked to differences in traits like disease prevalence provide a basis to study their evolutionary history, and possibly the environmental forces that have contributed in shaping the underlying genetic variation.

Remaining challenges and arising opportunities

Coming years will hopefully see growing use of iPSCs to address outstanding questions in human evolution. Continued improvement in differentiation strategies, diminishing costs, and emergent functional genomics technologies are all poised to come together and dramatically expand the range of questions within research scope. In this final section we outline areas where challenges and opportunities to the successful establishment of iPSCs as a widely adopted model system abound,

and their potential to be truly transformative in the study of questions about human-specific biology.

What can be modelled? What can we afford to model?

Differentiation of large quantities of mature cells from iPSCs is one of the current focuses of the differentiation field (for example, [161]), as is determining how faithfully these cells can model late-onset traits without being maintained in the laboratory for months or years (Figure 4A). Increasing organoid complexity is another priority, not only with regards to vascularisation or incorporation of tissue-resident macrophages; recent developments have seen multiple tissue types and organ systems integrated into single "assembloids" [162,163]. However, many evolutionary questions remain out of bounds even now. Beyond cognitive complexity, questions such as, which of the DNA differences that separate humans and chimpanzees contribute to human-uniqueness, or, how have different human populations adapted to environmental conditions such as variable pathogenic loads, encompass multiple organ systems and tissue types are likely not answerable with current methods. Reducing complex traits to phenotypes that can be robustly investigated in a laboratory setting requires not only additional technological developments, but also consideration of two other limitations: cost, and study design.

Even when protocols are available, the expertise and cost requirements associated with iPSC research are often prohibitive for all but a small number of research groups or consortia. A single organoid differentiation experiment can cost thousands of dollars—before sequencing. Recent advances that promise, for example, minimal cost media for iPSC maintenance [164] are welcome steps, but differentiation protocols often require highly specialised media as well as multiple recombinant proteins or synthetic small molecules. And even as prices continue to drop, other limitations, such as the need to culture 10s or 100s of organoids at a time are becoming more relevant.

Context-dependent study design

Concerns about cost cannot be divorced from sound study design. Robustly identifying instances of selective change between humans and chimpanzees at the transcriptome level can be done with as few as 6-10 samples from each species (Figure 4B) [165,166]. Yet even in the case of comparative great ape studies, the limited number of globally available cell lines means that it is difficult for any particular researcher to gain access to a sufficiently diverse number of them—often not for a want of trying. The international distribution of non-human ape iPSC lines is governed by CITES, a

global agreement on trade of endangered species, which adds substantial regulatory complexity to resource sharing.

The study of more recent evolutionary change presents different, but equally substantial challenges: Most methods for identifying recent (population-specific) positive selection single out candidate genetic regions that can span 100,000s of bases. Even when selection tests focus on individual SNPs [22], results are analogous to those from GWAS, with peaks of high-scoring SNPs in linkage disequilibrium (LD) with one another [167,168], and no assurance that the most statistically significant one is also the biologically relevant one [169]. This is an unavoidable consequence of existing LD structures in the human genome [167,168], exacerbated by the fact that when an advantageous allele rapidly rises in frequency, so do other SNPs in high linkage with it (Figure 1B, 4C). This problem is particularly pronounced for the functional study of introgressed archaic DNA. Due to the recent nature of admixture of humans with Neandertals and Denisovans, introgressed DNA is still segregating on haplotypes over tens of thousands of base pairs long. These haplotypes often carry tens or even hundreds of archaic variants [170], making prioritisation a complex problem. And while increasing sample sizes will help reduce the size of association peaks to address some of these issues, it is unlikely that we will soon reach the point of pinpointing single variants for testing simply via these existing methods.

In addition, GWAS have revealed that complex human traits like height or weight are highly polygenic [9], with associated loci numbered in the 1,000s and distributed throughout the genome. The majority of these make small contributions to the trait of interest. While there are stand-out examples of human evolutionary differences that are encoded by one or a handful of loci (reviewed in [171]), this genetic architecture means that when selection acts on a complex trait, it most likely does so through concerted small shifts in allele frequency across many of the loci that contribute to that trait [35]. Thus, while it is relatively straightforward to isolate and test the role of a single DNA variant using isogenic iPSC lines in combination with genome editing, the success of this approach is contingent on variant effect sizes and trait architecture. Although beyond the scope of this review, multiple studies have followed this path, sometimes referred to as ancestralisation [172]. While this allows for very fine-scale examination of variant effects, it bears remembering that these kinds of study ask a subtly different question — how does a specific variant (fixed in chimpanzees, for example) behave in this particular (human, for example) genetic context? — and raises an additional one: can it be confidently assumed that observations in this system truly mimic the effects of that variant in its native genetic context? In addition, the limitations of current genome

editing methods, which only allow the simultaneous editing of a handful of bases at once [173], mean that this approach is not well-suited for genome-wide approaches on its own.

Integrating functional genomics and population-scale iPSC studies for human evolution

Given these caveats, a promising approach is to leverage the flexibility and power of iPSC approaches with the scalability of high-throughput functional reporter assays such as massively parallel reporter assays (MPRAs, Figure 3A). These approaches combine traditional reporter assay designs with molecular barcoding, allowing the simultaneous testing of 10,000s of variants in a single experiment [174]. Above we highlighted two applications of MPRAs to the study of human-chimpanzee DNA differences [121,122]; another example is the work of Weiss et al., who combined human iPSCs and their differentiated progeny with an MPRA to test for differences in gene regulatory potential of 14,042 fixed DNA differences between Neandertals, Denisovans and humans, identifying 407 where the modern and archaic alleles had a significantly different ability to drive reporter gene expression [175]. Some of these variants impact genes implicated in cerebellum size, or cranial anatomy, although others were not as easily parsed.

In parallel, we are likely to see a combination of the eQTL mapping approaches, discussed above, where expression differences in either iPSCs or differentiated cells are linked to genetic variation, with more overtly evolutionary study designs. But in order to succeed, this approach must overcome a significant challenge—genetic diversity. A well-powered genome-wide eQTL study that can associate changes in DNA sequence to differences in gene expression between humans populations that diverged only 50,000 - 10,000 years ago requires a minimum of ~50-70 individuals to be sufficiently powered (smaller designs are possible, but drastically underpowered to detect small effects sizes, Figure 3B, 4B; see [176]). And while existing iPSC panels like HipSci or iPSCORE far exceed those sizes, and therefore represent useful assets to study evolutionary questions [159,177], they are far from representative of the breadth of existing human genetic diversity. With donors in both panels being predominantly of European descent (Figure 4D), their use in investigations of the architecture of complex traits excludes a substantial fraction of the global population [178,179].

These approaches can be further enhanced by combining them with experimental designs that take full advantage of the potential to perturb iPSCs in laboratory settings. Large-scale differentiation of iPSCs from a truly representative and global set of individuals into immune cells such as macrophages, followed by exposure to a set of pathogens or pathogen-mimicking molecules, would

allow for the identification of eQTLs associated with differences in immune response between individuals of different populations, complementing existing studies on whole blood [180–184]. Likewise, differentiation into hepatocytes would allow for high-throughput screening of drug compounds that vary significantly in efficacy across genotypes and populations.

Questions like these, and others that once appeared equally intractable, are within scope — if not now, then in the next five or ten years. iPSCs have a vast and clear potential to facilitate the study of human evolution at unprecedented molecular depth, bringing the field closer to grasping the biological roots of human uniqueness. The challenge now is chiefly one of imagination.

Author contributions

MD conceived the topic and figures, and wrote and edited the paper. IGR conceived the topic and figures, and wrote and edited the paper.

Acknowledgments

All figures were created with Biorender.com. Both authors were supported by European Union through Horizon 2020 Research and Innovation Programme under grant no 810645 and the European Union through the European Regional Development Fund project no. MOBEC008. IGR was additionally supported by Australian Research Council Discovery Project DP200101552.

References

1 Varki A & Altheide TK (2005) Comparing the human and chimpanzee genomes: searching for needles in a haystack. *Genome Res* **15**, 1746–1758.

2 Gallego Romero I, Ruvinsky I & Gilad Y (2012) Comparative studies of gene expression and the evolution of gene regulation. *Nature Reviews Genetics* **13**, 505–516.

3 King M & Wilson A (1975) Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116.

4 Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25–36.

5 Hoekstra HE & Coyne JA (2007) The locus of evolution: evo devo and the genetics of adaptation. *Evolution* **61**, 995–1016.

6 Carroll SB (2003) Genetics and the making of Homo sapiens. *Nature* **422**, 849–857.

7 Wilson DE & Reeder DM (2005) *Mammal Species of the World: A Taxonomic and Geographic Reference* JHU Press.

8 Knight A (2008) The beginning of the end for chimpanzee experiments? *Philos Ethics Humanit Med* **3**, 16.

9 Boyle EA, Li YI & Pritchard JK (2017) An Expanded View of Complex Traits: From Polygenic to

552 Omnigenic. *Cell* **169**, 1177–1186.
553 10 Loos RJF (2020) 15 years of genome-wide association studies and no signs of slowing down. *Nat*
554 *Commun* **11**, 5900.
555 11 Chimpanzee Sequencing and Analysis Consortium (2005) Initial sequence of the chimpanzee genome and
556 comparison with the human genome. *Nature* **437**, 69–87.
557 12 Galway-Witham J, Cole J & Stringer C (2019) Aspects of human physical and behavioural evolution
558 during the last 1 million years. *J Quat Sci* **34**, 355–378.
559 13 Hublin J-J, Ben-Ncer A, Bailey SE, Freidline SE, Neubauer S, Skinner MM, Bergmann I, Le Cabec A,
560 Benazzi S, Harvati K & Gunz P (2018) Author Correction: New fossils from Jebel Irhoud, Morocco and
561 the pan-African origin of Homo sapiens. *Nature* **558**, E6.
562 14 Stringer C (2003) Out of Ethiopia. *Nature* **423**, 693–695.
563 15 Henn BM, Cavalli-Sforza LL & Feldman MW (2012) The great human expansion. *Proc Natl Acad Sci U*
564 *S A* **109**, 17758–17764.
565 16 Bowler JM, Johnston H, Olley JM, Prescott JR, Roberts RG, Shawcross W & Spooner NA (2003) New
566 ages for human occupation and climatic change at Lake Mungo, Australia. *Nature* **421**, 837–840.
567 17 Malaspinas A-S, Westaway MC, Muller C, Sousa VC, Lao O, Alves I, Bergström A, Athanasiadis G,
568 Cheng JY, Crawford JE, Heupink TH, Macholdt E, Peischl S, Rasmussen S, Schiffels S, Subramanian
569 S, Wright JL, Albrechtsen A, Barbieri C, Dupanloup I, Eriksson A, Margaryan A, Moltke I, Pugach I,
570 Korneliussen TS, Levkivskyi IP, Moreno-Mayar JV, Ni S, Racimo F, Sikora M, Xue Y, Aghakhanian
571 FA, Brucato N, Brunak S, Campos PF, Clark W, Ellingvåg S, Fourmile G, Gerbault P, Injie D, Koki G,
572 Leavesley M, Logan B, Lynch A, Matisoo-Smith EA, McAllister PJ, Mentzer AJ, Metspalu M,
573 Migliano AB, Murgha L, Phipps ME, Pomat W, Reynolds D, Ricaut F-X, Siba P, Thomas MG, Wales
574 T, Wall CM, Oppenheimer SJ, Tyler-Smith C, Durbin R, Dortch J, Manica A, Schierup MH, Foley RA,
575 Lahr MM, Bowern C, Wall JD, Mailund T, Stoneking M, Nielsen R, Sandhu MS, Excoffier L, Lambert
576 DM & Willerslev E (2016) A genomic history of Aboriginal Australia. *Nature* **538**, 207–214.
577 18 Becerra-Valdivia L & Higham T (2020) The timing and effect of the earliest human arrivals in North
578 America. *Nature* **584**, 93–97.
579 19 Ardelean CF, Becerra-Valdivia L, Pedersen MW, Schwenninger J-L, Oviatt CG, Macías-Quintero JI,
580 Arroyo-Cabrales J, Sikora M, Ocampo-Díaz YZE, Rubio-Cisneros II, Watling JG, de Medeiros VB, De
581 Oliveira PE, Barba-Pingarón L, Ortiz-Butrón A, Blancas-Vázquez J, Rivera-González I, Solís-Rosales
582 C, Rodríguez-Ceja M, Gandy DA, Navarro-Gutierrez Z, De La Rosa-Díaz JJ, Huerta-Arellano V,
583 Marroquín-Fernández MB, Martínez-Riojas LM, López-Jiménez A, Higham T & Willerslev E (2020)
584 Evidence of human occupation in Mexico around the Last Glacial Maximum. *Nature* **584**, 87–92.
585 20 Skoglund P & Mathieson I (2018) Ancient Genomics of Modern Humans: The First Decade. *Annu Rev*
586 *Genomics Hum Genet* **19**, 381–404.
587 21 Ilardo M & Nielsen R (2018) Human adaptation to extreme environmental conditions. *Current Opinion in*
588 *Genetics & Development* **53**, 77–82.
589 22 Fan S, Hansen MEB, Lo Y & Tishkoff SA (2016) Going global by adapting local: A review of recent

590 human adaptation. *Science* **354**, 54–59.
591 23 Quintana-Murci L (2019) Human Immunology through the Lens of Evolutionary Genetics. *Cell* **177**, 184–
592 199.
593 24 Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, Heinze A, Renaud G, Sudmant PH,
594 de Filippo C, Li H, Mallick S, Dannemann M, Fu Q, Kircher M, Kuhlwilm M, Lachmann M, Meyer M,
595 Ongyerth M, Siebauer M, Theunert C, Tandon A, Moorjani P, Pickrell J, Mullikin JC, Vohr SH, Green
596 RE, Hellmann I, Johnson PLF, Blanche H, Cann H, Kitzman JO, Shendure J, Eichler EE, Lein ES,
597 Bakken TE, Golovanova LV, Doronichev VB, Shunkov MV, Derevianko AP, Viola B, Slatkin M,
598 Reich D, Kelso J & Pääbo S (2014) The complete genome sequence of a Neanderthal from the Altai
599 Mountains. *Nature* **505**, 43–49.
600 25 Prüfer K, de Filippo C, Grote S, Mafessoni F, Korlević P, Hajdinjak M, Vernot B, Skov L, Hsieh P,
601 Peyrégne S, Reher D, Hopfe C, Nagel S, Maricic T, Fu Q, Theunert C, Rogers R, Skoglund P,
602 Chintalapati M, Dannemann M, Nelson BJ, Key FM, Rudan P, Kućan Ž, Gušić I, Golovanova LV,
603 Doronichev VB, Patterson N, Reich D, Eichler EE, Slatkin M, Schierup MH, Andrés AM, Kelso J,
604 Meyer M & Pääbo S (2017) A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science*
605 **358**, 655–658.
606 26 Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH-
607 Y, Hansen NF, Durand EY, Malaspinas A-S, Jensen JD, Marques-Bonet T, Alkan C, Prüfer K, Meyer
608 M, Burbano HA, Good JM, Schultz R, Aximu-Petri A, Butthof A, Höber B, Höffner B, Siegemund M,
609 Weihmann A, Nusbaum C, Lander ES, Russ C, Novod N, Affourtit J, Egholm M, Verna C, Rudan P,
610 Brajkovic D, Kucan Ž, Gušic I, Doronichev VB, Golovanova LV, Lalueza-Fox C, de la Rasilla M,
611 Fortea J, Rosas A, Schmitz RW, Johnson PLF, Eichler EE, Falush D, Birney E, Mullikin JC, Slatkin M,
612 Nielsen R, Kelso J, Lachmann M, Reich D & Pääbo S (2010) A draft sequence of the Neandertal
613 genome. *Science* **328**, 710–722.
614 27 Meyer M, Kircher M, Gansauge M-T, Li H, Racimo F, Mallick S, Schraiber JG, Jay F, Prüfer K, de
615 Filippo C, Sudmant PH, Alkan C, Fu Q, Do R, Rohland N, Tandon A, Siebauer M, Green RE, Bryc K,
616 Briggs AW, Stenzel U, Dabney J, Shendure J, Kitzman J, Hammer MF, Shunkov MV, Derevianko AP,
617 Patterson N, Andrés AM, Eichler EE, Slatkin M, Reich D, Kelso J & Pääbo S (2012) A high-coverage
618 genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226.
619 28 Sankararaman S, Mallick S, Dannemann M, Prüfer K, Kelso J, Pääbo S, Patterson N & Reich D (2014)
620 The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**, 354–357.
621 29 Vernot B, Tucci S, Kelso J, Schraiber JG, Wolf AB, Gittelman RM, Dannemann M, Grote S, McCoy RC,
622 Norton H, Scheinfeldt LB, Merriwether DA, Koki G, Friedlaender JS, Wakefield J, Pääbo S & Akey JM
623 (2016) Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals.
624 *Science* **352**, 235–239.
625 30 Qin P & Stoneking M (2015) Denisovan Ancestry in East Eurasian and Native American Populations.
626 *Molecular Biology and Evolution* **32**, 2665–2674.
627 31 Gittelman RM, Schraiber JG, Vernot B, Mikacenic C, Wurfel MM & Akey JM (2016) Archaic Hominin

Admixture Facilitated Adaptation to Out-of-Africa Environments. *Curr Biol* **26**, 3375–3382.

32 Dannemann M & Racimo F (2018) Something old, something borrowed: admixture and adaptation in human evolution. *Curr Opin Genet Dev* **53**, 1–8.

33 Booker TR, Jackson BC & Keightley PD (2017) Detecting positive selection in the genome. *BMC Biol* **15**, 1–10.

34 Vitti JJ, Grossman SR & Sabeti PC (2013) Detecting natural selection in genomic data. *Annu Rev Genet* **47**, 97–120.

35 Wellenreuther M & Hansson B (2016) Detecting Polygenic Evolution: Problems, Pitfalls, and Promises. *Trends Genet* **32**, 155–164.

36 Takahashi K & Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676.

37 Shi Y, Inoue H, Wu JC & Yamanaka S (2017) Induced pluripotent stem cell technology: a decade of progress. *Nat Rev Drug Discov* **16**, 115–130.

38 Malik N & Rao MS (2013) A Review of the Methods for Human iPSC Derivation. *Methods in Molecular Biology*, 23–33.

39 Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II & Thomson JA (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**.

40 Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K & Yamanaka S (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872.

41 Loh Y-H, Agarwal S, Park I-H, Urbach A, Huo H, Heffner GC, Kim K, Miller JD, Ng K & Daley GQ (2009) Generation of induced pluripotent stem cells from human blood. *Blood* **113**, 5476–5479.

42 Ye Z, Zhan H, Mali P, Dowey S, Williams DM, Jang Y-Y, Dang CV, Spivak JL, Moliterno AR & Cheng L (2009) Human-induced pluripotent stem cells from blood cells of healthy donors and patients with acquired blood disorders. *Blood* **114**, 5473–5480.

43 Zhou T, Benda C, Duzinger S, Huang Y, Li X, Li Y, Guo X, Cao G, Chen S, Hao L, Chan Y-C, Ng K-M, Ho JC, Wieser M, Wu J, Redl H, Tse H-F, Grillari J, Grillari-Voglauer R, Pei D & Esteban MA (2011) Generation of induced pluripotent stem cells from urine. *J Am Soc Nephrol* **22**, 1221–1228.

44 Bilic J & Izpisua Belmonte JC (2012) Concise review: Induced pluripotent stem cells versus embryonic stem cells: close enough or yet too far apart? *Stem Cells* **30**, 33–41.

45 Tomer Halevy AU (2014) Comparing ESC and iPSC—Based Models for Human Genetic Disorders. *J Clin Med Res* **3**, 1146.

46 Burrows CK, Banovich NE, Pavlovic BJ, Patterson K, Gallego Romero I, Pritchard JK & Gilad Y (2016) Genetic Variation, Not Cell Type of Origin, Underlies the Majority of Identifiable Regulatory Differences in iPSCs. *PLoS Genet* **12**, e1005793.

47 Rouhani F, Kumasaka N, de Brito MC, Bradley A, Vallier L & Gaffney D (2014) Genetic background drives transcriptional variation in human induced pluripotent stem cells. *PLoS Genet* **10**, e1004432.

48 Panopoulos AD, D'Antonio M, Benaglio P, Williams R, Hashem SI, Schuldt BM, DeBoever C, Arias

AD, Garcia M, Nelson BC, Harismendy O, Jakubosky DA, Donovan MKR, Greenwald WW, Farnam K, Cook M, Borja V, Miller CA, Grinstein JD, Drees F, Okubo J, Diffenderfer KE, Hishida Y, Modesto V, Dargitz CT, Feiring R, Zhao C, Aguirre A, McGarry TJ, Matsui H, Li H, Reyna J, Rao F, O'Connor DT, Yeo GW, Evans SM, Chi NC, Jepsen K, Nariai N, Müller F-J, Goldstein LSB, Izpisua Belmonte JC, Adler E, Loring JF, Berggren WT, D'Antonio-Chronowska A, Smith EN & Frazer KA (2017) iPSCORE: A Resource of 222 iPSC Lines Enabling Functional Characterization of Genetic Variation across a Variety of Cell Types. *Stem Cell Reports* **8**, 1086–1100.

49 Kilpinen H, Goncalves A, Leha A, Afzal V, Alasoo K, Ashford S, Bala S, Bensaddek D, Casale FP, Culley OJ, Danecek P, Faulconbridge A, Harrison PW, Kathuria A, McCarthy D, McCarthy SA, Meleckyte R, Memari Y, Moens N, Soares F, Mann A, Streeter I, Agu CA, Alderton A, Nelson R, Harper S, Patel M, White A, Patel SR, Clarke L, Halai R, Kirton CM, Kolb-Kokocinski A, Beales P, Birney E, Danovi D, Lamond AI, Ouwehand WH, Vallier L, Watt FM, Durbin R, Stegle O & Gaffney DJ (2017) Common genetic variation drives molecular heterogeneity in human iPSCs. *Nature* **546**, 370–375.

50 Kyttälä A, Moraghebi R, Valensisi C, Kettunen J, Andrus C, Pasumarthy KK, Nakanishi M, Nishimura K, Ohtaka M, Weltner J, Van Handel B, Parkkonen O, Sinisalo J, Jalanko A, Hawkins RD, Woods N-B, Otonkoski T & Trokovic R (2016) Genetic Variability Overrides the Impact of Parental Cell Type and Determines iPSC Differentiation Potential. *Stem Cell Reports* **6**, 200–212.

51 Ben-Nun IF, Montague SC, Houck ML, Tran HT, Garitaonandia I, Leonardo TR, Wang Y-C, Charter SJ, Laurent LC, Ryder OA & Loring JF (2011) Induced pluripotent stem cells from highly endangered species. *Nat Methods* **8**, 829–831.

52 Whitworth DJ, Limnios IJ, Gauthier M-E, Weeratunga P, Ovchinnikov DA, Baillie G, Grimmond SM, Graves JAM & Wolvetang EJ (2019) Platypus Induced Pluripotent Stem Cells: The Unique Pluripotency Signature of a Monotreme. *Stem Cells Dev* **28**, 151–164.

53 Ogorevc J, Orehek S & Dovč P (2016) Cellular reprogramming in farm animals: an overview of iPSC generation in the mammalian farm animal species. *J Anim Sci Biotechnol* **7**, 1–9.

54 Geuder J, Wange LE, Janjic A, Radmer J, Janssen P, Bagnoli JW, Müller S, Kaul A, Ohnuki M & Enard W (2021) A non-invasive method to generate induced pluripotent stem cells from primate urine. *Sci Rep* **11**, 3516.

55 Pessôa LV de F, Bressan FF & Freude KK (2019) Induced pluripotent stem cells throughout the animal kingdom: Availability and applications. *World J Stem Cells* **11**, 491–505.

56 Yujeong Oh JJ (2019) Directed Differentiation of Pluripotent Stem Cells by Transcription Factors. *Mol Cells* **42**, 200.

57 Kim J, Koo B-K & Knoblich JA (2020) Human organoids: model systems for human biology and medicine. *Nat Rev Mol Cell Biol* **21**, 571–584.

58 Andersen J, Revah O, Miura Y, Thom N, Amin ND, Kelley KW, Singh M, Chen X, Thete MV, Walczak EM, Vogel H, Fan HC & Paşca SP (2020) Generation of Functional Human 3D Cortico-Motor Assembloids. *Cell* **183**, 1913–1929.e26.

59 Lancaster MA, Renner M, Martin C-A, Wenzel D, Bicknell LS, Hurles ME, Homfray T, Penninger JM, Jackson AP & Knoblich JA (2013) Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373–379.

60 Sato T, Stange DE, Ferrante M, Vries RGJ, van Es JH, van den Brink S, van Houdt WJ, Pronk A, van Gorp J, Siersema PD & Clevers H (2011) Long-term Expansion of Epithelial Organoids From Human Colon, Adenoma, Adenocarcinoma, and Barrett's Epithelium. *Gastroenterology* **141**, 1762–1772.

61 Fujii M, Matano M, Toshimitsu K, Takano A, Mikami Y, Nishikori S, Sugimoto S & Sato T (2018) Human Intestinal Organoids Maintain Self-Renewal Capacity and Cellular Diversity in Niche-Inspired Culture Condition. *Cell Stem Cell* **23**, 787–793.e6.

62 Takasato M, Er PX, Chiu HS, Maier B, Baillie GJ, Ferguson C, Parton RG, Wolvetang EJ, Roost MS, Lopes SMC de S & Little MH (2016) Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. *Nature* **536**, 238.

63 Lee J, Rabbani CC, Gao H, Steinhart MR, Woodruff BM, Pflum ZE, Kim A, Heller S, Liu Y, Shipchandler TZ & Koehler KR (2020) Hair-bearing human skin generated entirely from pluripotent stem cells. *Nature* **582**, 399–404.

64 Hu H, Gehart H, Artegiani B, LÖpez-Iglesias C, Dekkers F, Basak O, van Es J, Chuva de Sousa Lopes SM, Begthel H, Korving J, van den Born M, Zou C, Quirk C, Chiriboga L, Rice CM, Ma S, Rios A, Peters PJ, de Jong YP & Clevers H (2018) Long-Term Expansion of Functional Mouse and Human Hepatocytes as 3D Organoids. *Cell* **175**, 1591–1606.e19.

65 Modeling Development and Disease with Organoids (2016) *Cell* **165**, 1586–1597.

66 Clevers H & Tuveson DA (2019) Organoid Models for Cancer Research. *Annu Rev Cancer Biol* **3**, 223–234.

67 Kaushik G, Ponnusamy MP & Batra SK (2018) Concise Review: Current Status of Three-Dimensional Organoids as Preclinical Models. *Stem Cells* **36**, 1329–1340.

68 Tuveson D & Clevers H (2019) Cancer modeling meets human organoid technology. *Science* **364**, 952–955.

69 Duval K, Grover H, Han L-H, Mou Y, Pegoraro AF, Fredberg J & Chen Z (2017) Modeling Physiological Events in 2D vs. 3D Cell Culture. *Physiology* **32**, 266–277.

70 Touboul T, Hannan NRF, Corbineau S, Martinez A, Martinet C, Branchereau S, Mainot S, Strick-Marchand H, Pedersen R, Di Santo J, Weber A & Vallier L (2010) Generation of functional hepatocytes from human embryonic stem cells under chemically defined conditions that recapitulate liver development. *Hepatology* **51**, 1754–1765.

71 Si-Tayeb K, Noto FK, Nagaoka M, Li J, Battle MA, Duris C, North PE, Dalton S & Duncan SA (2010) Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. *Hepatology* **51**, 297–305.

72 Song Z, Cai J, Liu Y, Zhao D, Yong J, Duo S, Song X, Guo Y, Zhao Y, Qin H, Yin X, Wu C, Che J, Lu S, Ding M & Deng H (2009) Efficient generation of hepatocyte-like cells from human induced pluripotent stem cells. *Cell Res* **19**, 1233–1242.

73 Wu H, Uchimura K, Donnelly EL, Kirita Y, Morris SA & Humphreys BD (2018) Comparative Analysis and Refinement of Human PSC-Derived Kidney Organoid Differentiation with Single-Cell Transcriptomics. *Cell Stem Cell* **23**, 869–881.e8.

74 Camp JG, Sekine K, Gerber T, Loeffler-Wirth H, Binder H, Gac M, Kanton S, Kageyama J, Damm G, Seehofer D, Belicova L, Bickle M, Barsacchi R, Okuda R, Yoshizawa E, Kimura M, Ayabe H, Taniguchi H, Takebe T & Treutlein B (2017) Multilineage communication regulates human liver bud development from pluripotency. *Nature* **546**, 533–538.

75 Lancaster MA & Knoblich JA (2014) Generation of cerebral organoids from human pluripotent stem cells. *Nature Protocols* **9**, 2329–2340.

76 van den Berg CW, Okawa S, Chuva de Sousa Lopes SM, van Iperen L, Passier R, Braam SR, Tertoolen LG, del Sol A, Davis RP & Mummery CL (2015) Transcriptome of human foetal heart compared with cardiomyocytes from pluripotent stem cells. *Development* **142**, 3231–3238.

77 Yang X, Pabon L & Murry CE (2014) Engineering adolescence: maturation of human pluripotent stem cell-derived cardiomyocytes. *Circ Res* **114**, 511–523.

78 D'Antonio M, Woodruff G, Nathanson JL, D'Antonio-Chronowska A, Arias A, Matsui H, Williams R, Herrera C, Reyna SM, Yeo GW, Goldstein LSB, Panopoulos AD & Frazer KA (2017) High-Throughput and Cost-Effective Characterization of Induced Pluripotent Stem Cells. *Stem Cell Reports* **8**, 1101–1111.

79 Koyanagi-Aoi M, Ohnuki M, Takahashi K, Okita K, Noma H, Sawamura Y, Teramoto I, Narita M, Sato Y, Ichisaka T, Amano N, Watanabe A, Morizane A, Yamada Y, Sato T, Takahashi J & Yamanaka S (2013) Differentiation-defective phenotypes revealed by large-scale analyses of human pluripotent stem cells. *Proc Natl Acad Sci U S A* **110**, 20569–20574.

80 Keller A, Dziedzicka D, Zambelli F, Markouli C, Sermon K, Spits C & Geens M (2018) Genetic and epigenetic factors which modulate differentiation propensity in human pluripotent stem cells. *Hum Reprod Update* **24**, 162–175.

81 Yoon S-J, Elahi LS, Pașca AM, Marton RM, Gordon A, Revah O, Miura Y, Walczak EM, Holdgate GM, Fan HC, Huguenard JR, Geschwind DH & Pașca SP (2019) Reliability of human cortical organoid generation. *Nat Methods* **16**, 75–78.

82 Velasco S, Kedaigle AJ, Simmons SK, Nash A, Rocha M, Quadrato G, Paulsen B, Nguyen L, Adiconis X, Regev A, Levin JZ & Arlotta P (2019) Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. *Nature* **570**, 523–527.

83 Volpato V, Smith J, Sandor C, Ried JS, Baud A, Handel A, Newey SE, Wessely F, Attar M, Whiteley E, Chintawar S, Verheyen A, Barta T, Lako M, Armstrong L, Muschet C, Artati A, Cusulin C, Christensen K, Patsch C, Sharma E, Nicod J, Brownjohn P, Stubbs V, Heywood WE, Gissen P, De Filippis R, Janssen K, Reinhardt P, Adamski J, Royaux I, Peeters PJ, Terstappen GC, Graf M, Livesey FJ, Akerman CJ, Mills K, Bowden R, Nicholson G, Webber C, Cader MZ & Lakics V (2018) Reproducibility of Molecular Phenotypes after Long-Term Differentiation to Human iPSC-Derived Neurons: A Multi-Site Omics Study. *Stem Cell Reports* **11**, 897–911.

780 84 Volpato V & Webber C (2020) Addressing variability in iPSC-derived models of human disease:
781 guidelines to promote reproducibility. *Dis Model Mech* **13**.
782 85 Wen L & Tang F (2016) Single-cell sequencing in stem cell biology. *Genome Biol* **17**, 71.
783 86 Brazovskaja A, Treutlein B & Camp JG (2019) High-throughput single-cell transcriptomics on organoids.
784 *Curr Opin Biotechnol* **55**, 167–171.
785 87 Kashima Y, Sakamoto Y, Kaneko K, Seki M, Suzuki Y & Suzuki A (2020) Single-cell sequencing
786 techniques from individual to multiomics analyses. *Exp Mol Med* **52**, 1419–1427.
787 88 Lee J, Hyeon DY & Hwang D (2020) Single-cell multiomics: technologies and data analysis methods.
788 *Exp Mol Med* **52**, 1428–1442.
789 89 Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, Satija R
790 & Smibert P (2017) Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods*
791 **14**, 865–868.
792 90 Clark SJ, Argelaguet R, Kapourani C-A, Stubbs TM, Lee HJ, Alda-Catalinas C, Krueger F, Sanguinetti G,
793 Kelsey G, Marioni JC, Stegle O & Reik W (2018) scNMT-seq enables joint profiling of chromatin
794 accessibility DNA methylation and transcription in single cells. *Nat Commun* **9**, 781.
795 91 Cao J, Cusanovich DA, Ramani V, Aghamirzaie D, Pliner HA, Hill AJ, Daza RM, McFaline-Figueroa JL,
796 Packer JS, Christiansen L, Steemers FJ, Adey AC, Trapnell C & Shendure J (2018) Joint profiling of
797 chromatin accessibility and gene expression in thousands of single cells. *Science* **361**, 1380–1385.
798 92 Tohyama S, Hattori F, Sano M, Hishiki T, Nagahata Y, Matsuura T, Hashimoto H, Suzuki T, Yamashita
799 H, Satoh Y, Egashira T, Seki T, Muraoka N, Yamakawa H, Ohgino Y, Tanaka T, Yoichi M, Yuasa S,
800 Murata M, Suematsu M & Fukuda K (2013) Distinct metabolic flow enables large-scale purification of
801 mouse and human pluripotent stem cell-derived cardiomyocytes. *Cell Stem Cell* **12**, 127–137.
802 93 Kang HM, Subramaniam M, Targ S, Nguyen M, Maliskova L, McCarthy E, Wan E, Wong S, Byrnes L,
803 Lanata CM, Gate RE, Mostafavi S, Marson A, Zaitlen N, Criswell LA & Ye CJ (2018) Multiplexed
804 droplet single-cell RNA-sequencing using natural genetic variation. *Nat Biotechnol* **36**, 89–94.
805 94 Huang Y, McCarthy DJ & Stegle O (2019) Vireo: Bayesian demultiplexing of pooled single-cell RNA-
806 seq data without genotype reference. *Genome Biol* **20**, 273.
807 95 Marchetto MCN, Narvaiza I, Denli AM, Benner C, Lazzarini TA, Nathanson JL, Paquola ACM, Desai
808 KN, Herai RH, Weitzman MD, Yeo GW, Muotri AR & Gage FH (2013) Differential L1 regulation in
809 pluripotent stem cells of humans and apes. *Nature* **503**, 525–529.
810 96 Wunderlich S, Kircher M, Vieth B, Haase A, Merkert S, Beier J, Göhring G, Glage S, Schambach A,
811 Curnow EC, Pääbo S, Martin U & Enard W (2014) Primate iPS cells as tools for evolutionary analyses.
812 *Stem Cell Res* **12**, 622–629.
813 97 Fujie Y, Fusaki N, Katayama T, Hamasaki M, Soejima Y, Soga M, Ban H, Hasegawa M, Yamashita S,
814 Kimura S, Suzuki S, Matsuzawa T, Akari H & Era T (2014) New type of Sendai virus vector provides
815 transgene-free iPS cells derived from chimpanzee blood. *PLoS One* **9**, e113052.
816 98 Kanton S, Boyle MJ, He Z, Santel M, Weigert A, Sanchís-Calleja F, Guijarro P, Sidow L, Fleck JS, Han
817 D, Qian Z, Heide M, Huttner WB, Khaitovich P, Pääbo S, Treutlein B & Camp JG (2019) Organoid

single-cell genomic atlas uncovers human-specific features of brain development. *Nature* **574**, 418–422.

99 Pollen AA, Bhaduri A, Andrews MG, Nowakowski TJ, Meyerson OS, Mostajo-Radji MA, Di Lullo E, Alvarado B, Bedolli M, Dougherty ML, Fiddes IT, Kronenberg ZN, Shuga J, Leyrat AA, West JA, Bershteyn M, Lowe CB, Pavlovic BJ, Salama SR, Haussler D, Eichler EE & Kriegstein AR (2019) Establishing Cerebral Organoids as Models of Human-Specific Brain Evolution. *Cell* **176**, 743–756.e17.

100 Field AR, Jacobs FMJ & Fiddes IT (2019) Structurally Conserved Primate LncRNAs Are Transiently Expressed during Human Cortical. *Stem cell reports, 12 (2)*.

101 Blake LE, Thomas SM, Blischak JD, Hsiao CJ, Chavarria C, Myrthil M, Gilad Y & Pavlovic BJ (2018) A comparative study of endoderm differentiation in humans and chimpanzees. *Genome Biol* **19**, 162.

102 Momeni B, Waite AJ & Shou W (2013) Spatial self-organization favors heterotypic cooperation over cheating. *Elife* **2**, e00960.

103 Gallego Romero I, Pavlovic BJ, Hernando-Herraez I, Zhou X, Ward MC, Banovich NE, Kagan CL, Burnett JE, Huang CH, Mitrano A, Chavarria CI, Friedrich Ben-Nun I, Li Y, Sabatini K, Leonardo TR, Parast M, Marques-Bonet T, Laurent LC, Loring JF & Gilad Y (2015) A panel of induced pluripotent stem cells from chimpanzees: a resource for comparative functional genomics. *Elife* **4**, e07103.

104 Ramsay L, Marchetto MC, Caron M, Chen S-H, Busche S, Kwan T, Pastinen T, Gage FH & Bourque G (2017) Conserved expression of transposon-derived non-coding transcripts in primate stem cells. *BMC Genomics* **18**, 214.

105 Ramaswamy K, Yik WY, Wang X-M, Oliphant EN, Lu W, Shibata D, Ryder OA & Hacia JG (2015) Derivation of induced pluripotent stem cells from orangutan skin fibroblasts. *BMC Res Notes* **8**, 577.

106 Navara CS, Chaudhari S & McCarrey JR (2018) Optimization of culture conditions for the derivation and propagation of baboon (Papio anubis) induced pluripotent stem cells. *PLoS One* **13**, e0193195.

107 Wunderlich S, Haase A, Merkert S, Beier J, Schwanke K, Schambach A, Glage S, Göhring G, Curnow EC & Martin U (2012) Induction of pluripotent stem cells from a cynomolgus monkey using a polycistronic simian immunodeficiency virus-based vector, differentiation toward functional cardiomyocytes, and generation of stably expressing reporter lines. *Cell Reprogram* **14**, 471–484.

108 Hemmi JJ, Mishra A & Hornsby PJ (2017) Overcoming barriers to reprogramming and differentiation in nonhuman primate induced pluripotent stem cells. *Primate Biol* **4**, 153–162.

109 Ward MC, Zhao S, Luo K, Pavlovic BJ, Karimi MM, Stephens M & Gilad Y (2018) Silencing of transposable elements may not be a major driver of regulatory evolution in primate iPSCs. *eLife* **7**.

110 Gallego Romero I, Gopalakrishnan S & Gilad Y (2018) Widespread conservation of chromatin accessibility patterns and transcription factor binding in human and chimpanzee induced pluripotent stem cells. *bioRxiv*.

111 Eres IE, Luo K, Hsiao CJ, Blake LE & Gilad Y (2019) Reorganization of 3D genome structure may contribute to gene regulatory evolution in primates. *PLoS Genet* **15**, e1008278.

112 Lin ZY-C, Nakai R, Hirai H, Kozuka D, Katayama S, Nakamura S-I, Okada S, Kitajima R, Imai H, Okano H & Imamura M (2020) Reprogramming of chimpanzee fibroblasts into a multipotent cancerous but not fully pluripotent state by transducing iPSC factors in 2i/LIF culture. *Differentiation* **112**, 67–76.

113 Herrmann E, Call J, Hernàndez-Lloreda MV, Hare B & Tomasello M (2007) Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* **317**, 1360–1366.

114 Dean LG, Kendal RL, Schapiro SJ, Thierry B & Laland KN (2012) Identification of the social and cognitive processes underlying human cumulative culture. *Science* **335**, 1114–1118.

115 Ward MC & Gilad Y (2019) A generally conserved response to hypoxia in iPSC-derived cardiomyocytes from humans and chimpanzees. *Elife* **8**.

116 Pavlovic BJ, Blake LE, Roux J, Chavarria C & Gilad Y (2018) A Comparative Assessment of Human and Chimpanzee iPSC-derived Cardiomyocytes with Primary Heart Tissues. *Sci Rep* **8**, 15312.

117 Roodgar M, Suchy FP, Bajpai V, Viches-Moure JG, Bhadury J, Oikonomopoulos A, Wu JC, Mankowski JL, Loh KM, Nakauchi H, VandeVoort CA & Snyder MP (2019) Cross-species blastocyst chimerism between nonhuman primates using iPSCs. *bioRxiv*.

118 Prescott SL, Srinivasan R, Marchetto MC, Grishina I, Narvaiza I, Selleri L, Gage FH, Swigut T & Wysocka J (2015) Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest. *Cell* **163**, 68–83.

119 Varki N, Anderson D, Herndon JG, Pham T, Gregg CJ, Cheriyan M, Murphy J, Strobert E, Fritz J, Else JG & Varki A (2009) Heart disease is common in humans and chimpanzees, but is caused by different pathological processes. *Evol Appl* **2**, 101–112.

120 Cugola FR, Fernandes IR, Russo FB, Freitas BC, Dias JLM, Guimarães KP, Benazzato C, Almeida N, Pignatari GC, Romero S, Polonio CM, Cunha I, Freitas CL, Brandão WN, Rossato C, Andrade DG, Faria D de P, Garcez AT, Buchpigel CA, Braconi CT, Mendes E, Sall AA, Zanotto PM de A, Peron JPS, Muotri AR & Beltrão-Braga PCB (2016) The Brazilian Zika virus strain causes birth defects in experimental models. *Nature* **534**, 267–271.

121 Uebbing S, Gockley J, Reilly SK, Kocher AA, Geller E, Gandotra N, Scharfe C, Cotney J & Noonan JP (2021) Massively parallel discovery of human-specific substitutions that alter enhancer activity. *Proc Natl Acad Sci U S A* **118**.

122 Ryu H, Inoue F, Whalen S, Williams A, Kircher M, Martin B, Alvarado B, Samee MAH, Keough K, Thomas S, Kriegstein A, Shendure J, Pollen A, Ahituv N & Pollard KS (2018) Massively parallel dissection of human accelerated regions in human and chimpanzee neural progenitors. *bioRxiv*.

123 Agoglia RM, Sun D, Birey F, Yoon S-J, Miura Y, Sabatini K, Pașca SP & Fraser HB (2021) Primate cell fusion disentangles gene regulatory divergence in neurodevelopment. *Nature*.

124 Gokhman D, Agoglia RM, Kinnebrew M, Gordon W, Sun D, Bajpai VK, Naqvi S, Chen C, Chan A, Chen C, Petrov DA, Ahituv N, Zhang H, Mishina Y, Wysocka J, Rohatgi R & Fraser HB (2021) Publisher Correction: Human-chimpanzee fused cells reveal cis-regulatory divergence underlying skeletal evolution. *Nat Genet*.

125 Mostajo-Radji MA, Schmitz MT, Montoya ST & Pollen AA (2020) Reverse engineering human brain evolution using organoid models. *Brain Res* **1729**, 146582.

126 Kitajima R, Nakai R, Imamura T, Kameda T, Kozuka D, Hirai H, Ito H, Imai H & Imamura M (2020) Modeling of early neural development in vitro by direct neurosphere formation culture of chimpanzee

induced pluripotent stem cells. *Stem Cell Res* **44**, 101749.

127 Marchetto MC, Hrvoj-Mihic B, Kerman BE, Yu DX, Vadodaria KC, Linker SB, Narvaiza I, Santos R, Denli AM, Mendes AP, Oefner R, Cook J, McHenry L, Grasmick JM, Heard K, Fredlender C, Randolph-Moore L, Kshirsagar R, Xenitopoulos R, Chou G, Hah N, Muotri AR, Padmanabhan K, Semendeferi K & Gage FH (2019) Species-specific maturation profiles of human, chimpanzee and bonobo neural cells. *Elife* **8**.

128 Grassi DA, Brattås PL, Valdés JG, Rezeli M, Jönsson ME, Nolbrant S, Parmar M, Marko-Varga G & Jakobsson J (2019) Post-transcriptional mechanisms distinguish human and chimp forebrain progenitor cells. *bioRxiv*.

129 Otani T, Marchetto MC, Gage FH, Simons BD & Livesey FJ (2016) 2D and 3D Stem Cell Models of Primate Cortical Development Identify Species-Specific Differences in Progenitor Behavior Contributing to Brain Size. *Cell Stem Cell* **18**, 467–480.

130 Schörnig M, Ju X, Fast L, Ebert S, Weigert A, Kanton S, Schaffer T, Nadif Kasri N, Treutlein B, Peter BM, Hevers W & Taverna E (2021) Comparison of induced neurons reveals slower structural and functional maturation in humans than in apes. *Elife* **10**.

131 Zintel TM, Pizzollo J, Claypool CG & Babbitt CC (2020) Astrocytes drive divergent metabolic gene expression in humans and chimpanzees. *bioRxiv*.

132 Johansson PA, Brattås PL, Douse CH, Hsieh P, Pontis J, Grassi D, Garza R, Jönsson ME, Atacho DAM, Pircs K, Eren F, Sharma Y, Johansson J, Trono D, Eichler EE & Jakobsson J (2020) A human-specific structural variation at the ZNF558 locus controls a gene regulatory network during forebrain development. *bioRxiv*.

133 Jang J, Wang Y, Lalli MA, Guzman E, Godshalk SE, Zhou H & Kosik KS (2016) Primary Cilium-Autophagy-Nrf2 (PAN) Axis Activation Commits Human Embryonic Stem Cells to a Neuroectoderm Fate. *Cell* **165**, 410–420.

134 Shi Y, Kirwan P & Livesey FJ (2012) Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nat Protoc* **7**, 1836–1846.

135 Mora-Bermúdez F, Badsha F, Kanton S, Camp JG, Vernot B, Köhler K, Voigt B, Okita K, Maricic T, He Z, Lachmann R, Pääbo S, Treutlein B & Huttner WB (2016) Differences and similarities between human and chimpanzee neural progenitors during cerebral cortex development. *Elife* **5**.

136 Field AR, Jacobs FMJ, Fiddes IT, Phillips APR, Reyes-Ortiz AM, LaMontagne E, Whitehead L, Meng V, Rosenkrantz JL, Olsen M, Hauessler M, Katzman S, Salama SR & Haussler D (2019) Structurally Conserved Primate LncRNAs Are Transiently Expressed during Human Cortical Differentiation and Influence Cell-Type-Specific Genes. *Stem Cell Reports* **12**, 245–257.

137 Fischer J, Peters J, Namba T, Huttner WB & Heide M (2020) Human-specific ARHGAP11B is necessary and sufficient for human-type basal progenitor levels in primate brain organoids. *bioRxiv*.

138 Kronenberg ZN, Fiddes IT, Gordon D, Murali S, Cantsilieris S, Meyerson OS, Underwood JG, Nelson BJ, Chaisson MJP, Dougherty ML, Munson KM, Hastie AR, Diekhans M, Hormozdiari F, Lorusso N, Hoekzema K, Qiu R, Clark K, Raja A, Welch AE, Sorensen M, Baker C, Fulton RS, Armstrong J,

Graves-Lindsay TA, Denli AM, Hoppe ER, Hsieh P, Hill CM, Pang AWC, Lee J, Lam ET, Dutcher SK, Gage FH, Warren WC, Shendure J, Haussler D, Schneider VA, Cao H, Ventura M, Wilson RK, Paten B, Pollen A & Eichler EE (2018) High-resolution comparative analysis of great ape genomes. *Science* **360**.

139 Fiddes IT, Lodewijk GA, Mooring M, Bosworth CM, Ewing AD, Mantalas GL, Novak AM, van den Bout A, Bishara A, Rosenkrantz JL, Lorig-Roach R, Field AR, Haeussler M, Russo L, Bhaduri A, Nowakowski TJ, Pollen AA, Dougherty ML, Nuttle X, Addor M-C, Zwolinski S, Katzman S, Kriegstein A, Eichler EE, Salama SR, Jacobs FMJ & Haussler D (2018) Human-Specific NOTCH2NL Genes Affect Notch Signaling and Cortical Neurogenesis. *Cell* **173**, 1356–1369.e22.

140 Trujillo CA, Rice ES, Schaefer NK, Chaim IA, Wheeler EC, Madrigal AA, Buchanan J, Preissl S, Wang A, Negraes PD, Szeto RA, Herai RH, Huseynov A, Ferraz MSA, Borges FS, Kihara AH, Byrne A, Marin M, Vollmers C, Brooks AN, Lautz JD, Semendeferi K, Shapiro B, Yeo GW, Smith SEP, Green RE & Muotri AR (2021) Reintroduction of the archaic variant of in cortical organoids alters neurodevelopment. *Science* **371**.

141 Benito-Kwiecinski S, Giandomenico SL, Sutcliffe M, Riis ES, Freire-Pritchett P, Kelava I, Wunderlich S, Martin U, Wray GA, McDole K & Lancaster MA (2021) An early cell shape transition drives evolutionary expansion of the human forebrain. *Cell*.

142 Sawai T, Sakaguchi H, Thomas E, Takahashi J & Fujita M (2019) The Ethics of Cerebral Organoid Research: Being Conscious of Consciousness. *Stem Cell Reports* **13**, 440–447.

143 Fujii M, Clevers H & Sato T (2019) Modeling Human Digestive Diseases With CRISPR-Cas9-Modified Organoids. *Gastroenterology* **156**, 562–576.

144 Roper J & Yilmaz ÖH (2019) Breakthrough Moments: Genome Editing and Organoids. *Cell Stem Cell* **24**, 841–842.

145 Klaus J, Kanton S, Kyrousi C, Ayo-Martin AC, Di Giaimo R, Riesenberg S, O'Neill AC, Camp JG, Tocco C, Santel M, Rusha E, Drukker M, Schroeder M, Götz M, Robertson SP, Treutlein B & Cappello S (2019) Altered neuronal migratory trajectories in human cerebral organoids derived from individuals with neuronal heterotopia. *Nat Med* **25**, 561–568.

146 Banovich NE, Li YI, Raj A, Ward MC, Greenside P, Calderon D, Tung PY, Burnett JE, Myrthil M, Thomas SM, Burrows CK, Romero IG, Pavlovic BJ, Kundaje A, Pritchard JK & Gilad Y (2018) Impact of regulatory variation across human iPSCs and differentiated cells. *Genome Res* **28**, 122–131.

147 International HapMap 3 Consortium, Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Peltonen L, Dermitzakis E, Bonnen PE, Altshuler DM, Gibbs RA, de Bakker PIW, Deloukas P, Gabriel SB, Gwilliam R, Hunt S, Inouye M, Jia X, Palotie A, Parkin M, Whittaker P, Yu F, Chang K, Hawes A, Lewis LR, Ren Y, Wheeler D, Gibbs RA, Muzny DM, Barnes C, Darvishi K, Hurles M, Korn JM, Kristiansson K, Lee C, McCarrol SA, Nemesh J, Dermitzakis E, Keinan A, Montgomery SB, Pollack S, Price AL, Soranzo N, Bonnen PE, Gibbs RA, Gonzaga-Jauregui C, Keinan A, Price AL, Yu F, Anttila V, Brodeur W, Daly MJ, Leslie S, McVean G, Moutsianas L, Nguyen H, Schaffner SF, Zhang Q, Ghori MJR, McGinnis R, McLaren W, Pollack S,

970 Price AL, Schaffner SF, Takeuchi F, Grossman SR, Shlyakhter I, Hostetter EB, Sabeti PC, Adebamowo
971 CA, Foster MW, Gordon DR, Licinio J, Manca MC, Marshall PA, Matsuda I, Ngare D, Wang VO,
972 Reddy D, Rotimi CN, Royal CD, Sharp RR, Zeng C, Brooks LD & McEwen JE (2010) Integrating
973 common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58.
974 148 Umekage M, Sato Y & Takasu N (2019) Overview: an iPS cell stock at CiRA. *Inflammation and*
975 *Regeneration* **39**.
976 149 Lee S, Huh JY, Turner DM, Lee S, Robinson J, Stein JE, Shim SH, Hong CP, Kang MS, Nakagawa M,
977 Kaneko S, Nakanishi M, Rao MS, Kurtz A, Stacey GN, Marsh SGE, Turner ML & Song J (2018)
978 Repurposing the cord blood bank for haplobanking of HLA-homozygous iPSCs and their usefulness to
979 multiple populations. *Stem Cells* **36**, 1552–1566.
980 150 DeBoever C, Li H, Jakubosky D, Benaglio P, Reyna J, Olson KM, Huang H, Biggs W, Sandoval E,
981 D'Antonio M, Jepsen K, Matsui H, Arias A, Ren B, Nariai N, Smith EN, D'Antonio-Chronowska A,
982 Farley EK & Frazer KA (2017) Large-Scale Profiling Reveals the Influence of Genetic Variation on
983 Gene Expression in Human Induced Pluripotent Stem Cells. *Cell Stem Cell* **20**, 533–546.e7.
984 151 Carcamo-Orive I, Hoffman GE, Cundiff P, Beckmann ND, D'Souza SL, Knowles JW, Patel A,
985 Papatsenko D, Abbasi F, Reaven GM, Whalen S, Lee P, Shahbazi M, Henrion MYR, Zhu K, Wang S,
986 Roussos P, Schadt EE, Pandey G, Chang R, Quertermous T & Lemischka I (2017) Analysis of
987 Transcriptional Variability in a Large Human iPSC Library Reveals Genetic and Non-genetic
988 Determinants of Heterogeneity. *Cell Stem Cell* **20**, 518–532.e9.
989 152 Pashos EE, Park Y, Wang X, Raghavan A, Yang W, Abbey D, Peters DT, Arbelaez J, Hernandez M,
990 Kuperwasser N, Li W, Lian Z, Liu Y, Lv W, Lytle-Gabbin SL, Marchadier DH, Rogov P, Shi J, Slovik
991 KJ, Stylianou IM, Wang L, Yan R, Zhang X, Kathiresan S, Duncan SA, Mikkelsen TS, Morrisey EE,
992 Rader DJ, Brown CD & Musunuru K (2017) Large, Diverse Population Cohorts of hiPSCs and Derived
993 Hepatocyte-like Cells Reveal Functional Genetic Variation at Blood Lipid-Associated Loci. *Cell Stem*
994 *Cell* **20**, 558–570.e10.
995 153 Bonder MJ, Smail C, Gloudemans MJ, Frésard L, Jakubosky D, D'Antonio M, Li X, Ferraro NM,
996 Carcamo-Orive I, Mirauta B, Seaton DD, Cai N, Vakili D, Horta D, Zhao C, Zastrow DB, Bonner DE,
997 HipSci Consortium, iPSCORE consortium, Undiagnosed Diseases Network, PhLiPS consortium,
998 Wheeler MT, Kilpinen H, Knowles JW, Smith EN, Frazer KA, Montgomery SB & Stegle O (2021)
999 Identification of rare and common regulatory variants in pluripotent cells using population-scale
1000 transcriptomics. *Nat Genet* **53**, 313–321.
1001 154 Knowles DA, Burrows CK, Blischak JD, Patterson KM, Serie DJ, Norton N, Ober C, Pritchard JK &
1002 Gilad Y (2018) Determining the genetic basis of anthracycline-cardiotoxicity by molecular response
1003 QTL mapping in induced cardiomyocytes. *Elife* **7**.
1004 155 Warren CR, O'Sullivan JF, Friesen M, Becker CE, Zhang X, Liu P, Wakabayashi Y, Morningstar JE,
1005 Shi X, Choi J, Xia F, Peters DT, Florido MHC, Tsankov AM, Duberow E, Comisar L, Shay J, Jiang X,
1006 Meissner A, Musunuru K, Kathiresan S, Daheron L, Zhu J, Gerszten RE, Deo RC, Vasan RS,
1007 O'Donnell CJ & Cowan CA (2017) Induced Pluripotent Stem Cell Differentiation Enables Functional

Validation of GWAS Variants in Metabolic Disease. *Cell Stem Cell* **20**, 547–557.e7.

156 Strober BJ, Elorbany R, Rhodes K, Krishnan N, Tayeb K, Battle A & Gilad Y (2019) Dynamic genetic regulation of gene expression during cellular differentiation. *Science* **364**, 1287–1290.

157 Cuomo ASE, Seaton DD, McCarthy DJ, Martinez I, Bonder MJ, Garcia-Bernardo J, Amatya S, Madrigal P, Isaacson A, Buettner F, Knights A, Natarajan KN, HipSci Consortium, Vallier L, Marioni JC, Chhatriwala M & Stegle O (2020) Publisher Correction: Single-cell RNA-sequencing of differentiating iPS cells reveals dynamic genetic effects on gene expression. *Nat Commun* **11**, 1572.

158 Umans BD, Battle A & Gilad Y (2020) Where Are the Disease-Associated eQTLs? *Trends Genet*.

159 Dannemann M, He Z, Heide C, Vernot B, Sidow L, Kanton S, Weigert A, Treutlein B, Pääbo S, Kelso J & Camp JG (2020) Human Stem Cell Resources Are an Inroad to Neandertal DNA Functions. *Stem Cell Reports* **15**, 214–225.

160 Benton ML, Abraham A, LaBella AL, Abbot P, Rokas A & Capra JA (2021) The influence of evolutionary history on human health and disease. *Nat Rev Genet*.

161 Giandomenico SL, Sutcliffe M & Lancaster MA (2020) Generation and long-term culture of advanced cerebral organoids for studying later stages of neural development. *Nat Protoc*.

162 Marton RM & Pașca SP (2020) Organoid and Assembloid Technologies for Investigating Cellular Crosstalk in Human Brain Development and Disease. *Trends Cell Biol* **30**, 133–143.

163 Pașca SP (2019) Assembling human brain organoids. *Science* **363**, 126–127.

164 Kuo H-H, Gao X, DeKeyser J-M, Fetterman KA, Pinheiro EA, Weddle CJ, Fonoudi H, Orman MV, Romero-Tejeda M, Jouni M, Blancard M, Magdy T, Epting CL, George AL Jr & Burridge PW (2020) Negligible-Cost and Weekend-Free Chemically Defined Human iPSC Culture. *Stem Cell Reports* **14**, 256–270.

165 Blekhman R, Marioni JC, Zumbo P, Stephens M & Gilad Y (2010) Sex-specific and lineage-specific alternative splicing in primates. *Genome Res* **20**, 180–189.

166 Kelley JL & Gilad Y (2020) Effective study design for comparative functional genomics. *Nat Rev Genet* **21**, 385–386.

167 Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farhadian SF, Ward R & Lander ES (2001) Linkage disequilibrium in the human genome. *Nature* **411**, 199–204.

168 Wall JD & Pritchard JK (2003) Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet* **4**, 587–597.

169 Schaid DJ, Chen W & Larson NB (2018) From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat Rev Genet* **19**, 491–504.

170 Sankararaman S, Patterson N, Li H, Pääbo S & Reich D (2012) The date of interbreeding between Neandertals and modern humans. *PLoS Genet* **8**, e1002947.

171 Scheinfeldt LB & Tishkoff SA (2013) Recent human adaptation: genomic approaches, interpretation and insights. *Nat Rev Genet* **14**, 692–702.

172 Riesenberg S, Maricic T & Pääbo S (2018) "Ancestralization" of human pluripotent stem cells by multiplexed precise genome editing. *bioRxiv*.

173 Pickar-Oliver A & Gersbach CA (2019) The next generation of CRISPR–Cas technologies and applications. *Nat Rev Mol Cell Biol* **20**, 490–507.

174 Inoue F & Ahituv N (2015) Decoding enhancers using massively parallel reporter assays. *Genomics* **106**, 159–164.

175 Weiss CV, Harshman L, Inoue F, Fraser HB, Petrov DA, Ahituv N & Gokhman D (2020) The cis-regulatory effects of modern human-specific variants. *bioRxiv*.

176 GTEx Consortium (2020) The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330.

177 Warren CR & Cowan CA (2018) Humanity in a Dish: Population Genetics with iPSCs. *Trends Cell Biol* **28**, 46–57.

178 Sirugo G, Williams SM & Tishkoff SA (2019) The Missing Diversity in Human Genetic Studies. *Cell* **177**, 1080.

179 Mogil LS, Andaleon A, Badalamenti A, Dickinson SP, Guo X, Rotter JI, Johnson WC, Im HK, Liu Y & Wheeler HE (2018) Genetic architecture of gene expression traits across diverse populations. *PLoS Genet* **14**, e1007586.

180 Quach H, Rotival M, Pothlichet J, Loh Y-HE, Dannemann M, Zidane N, Laval G, Patin E, Harmant C, Lopez M, Deschamps M, Naffakh N, Duffy D, Coen A, Leroux-Roels G, Clément F, Boland A, Deleuze J-F, Kelso J, Albert ML & Quintana-Murci L (2016) Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. *Cell* **167**, 643–656.e17.

181 Nédélec Y, Sanz J, Baharian G, Szpiech ZA, Pacis A, Dumaine A, Grenier J-C, Freiman A, Sams AJ, Hebert S, Pagé Sabourin A, Luca F, Blekhman R, Hernandez RD, Pique-Regi R, Tung J, Yotova V & Barreiro LB (2016) Genetic Ancestry and Natural Selection Drive Population Differences in Immune Responses to Pathogens. *Cell* **167**, 657–669.e21.

182 Kim-Hellmuth S, Bechheim M, Pütz B, Mohammadi P, Nédélec Y, Giangreco N, Becker J, Kaiser V, Fricker N, Beier E, Boor P, Castel SE, Nöthen MM, Barreiro LB, Pickrell JK, Müller-Myhsok B, Lappalainen T, Schumacher J & Hornung V (2017) Genetic regulatory effects modified by immune activation contribute to autoimmune disease associations. *Nat Commun* **8**, 266.

183 Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, Ha B, Altay G, Greenbaum JA, McVicker G, Seumois G, Rao A, Kronenberg M, Peters B & Vijayanand P (2018) Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. *Cell* **175**, 1701–1715.e16.

184 Lee MN, Ye C, Villani A-C, Raj T, Li W, Eisenhaure TM, Imboywa SH, Chipendo PI, Ran FA, Slowikowski K, Ward LD, Raddassi K, McCabe C, Lee MH, Frohlich IY, Hafler DA, Kellis M, Raychaudhuri S, Zhang F, Stranger BE, Benoist CO, De Jager PL, Regev A & Hacohen N (2014) Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science* **343**, 1246980.

185 Moorjani P, Amorim CEG, Arndt PF & Przeworski M (2016) Variation in the molecular clock of primates. *Proc Natl Acad Sci U S A* **113**, 10607–10612.

Figure legends

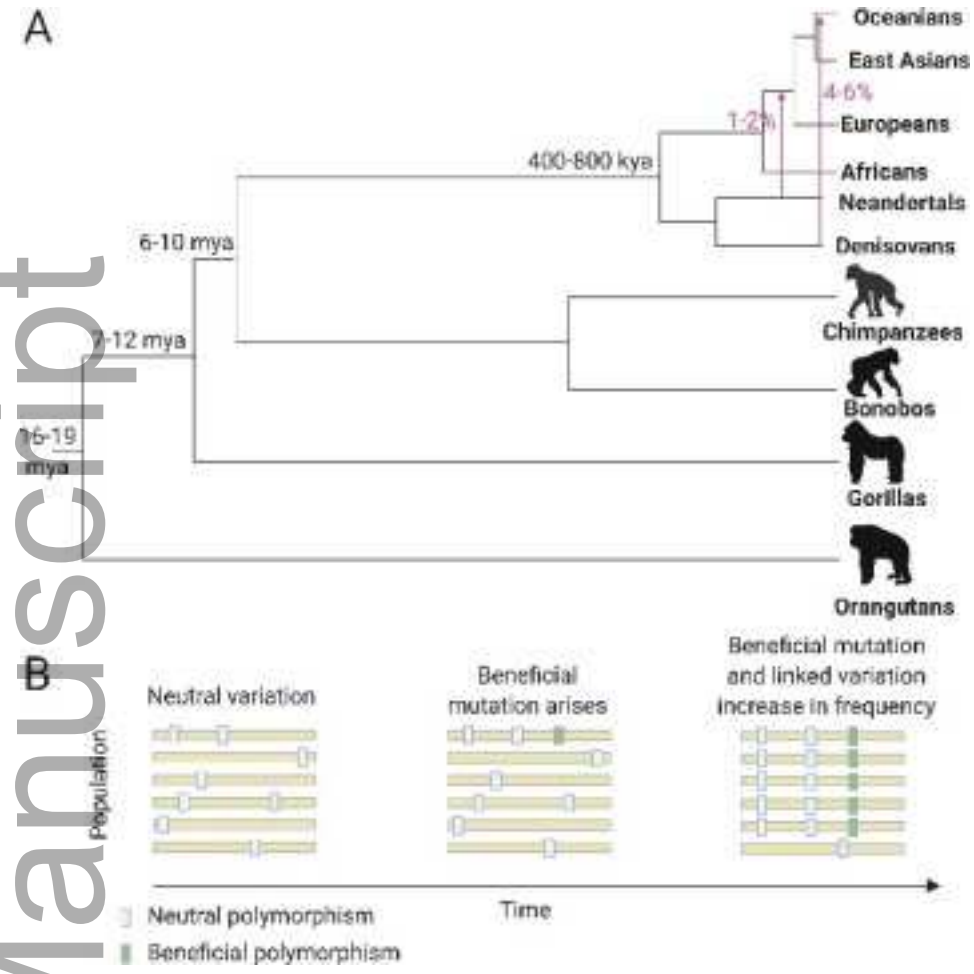
Figure 1: Adaptive change and human evolution (A) Genetic relationship of modern and archaic humans with other great ape species. A phylogenetic tree displaying the genetic relationship between different modern human populations, Neandertals, Denisovans and four great ape species is illustrated. Split times between species and modern and archaic humans are shown together with two major admixture events (admixture levels displayed in percent) between archaic and modern humans [185]. **(B) Schematic illustration of the process of positive selection over time.** Left: Haplotypes in a population (yellow) carrying neutral polymorphisms (light blue) are displayed. Middle: Emergence of a beneficial polymorphism (red) on one haplotype. Right: Haplotypes carrying the beneficial polymorphism increase in frequency in the population.

Figure 2: Application of iPSC differentiations to inter-species studies (A) The generation of induced pluripotent stem cells (iPSCs) and their potential to differentiate into several cell types and 3D organ models. Upper part: Three resources for the reprogramming of iPSCs and their ability for self-renewal are displayed in the upper part. Lower part: Differentiation protocols enable the generation of various cell types and organoids through the ectoderm, mesoderm and endoderm germ layers. For 3D organoid models the organs the corresponding organoids are resembling are illustrated. **(B) Application of iPSC-derived models to study human evolution.** Publications that have applied 2D and 3D models in evolutionary studies to compare humans with other great apes are displayed.

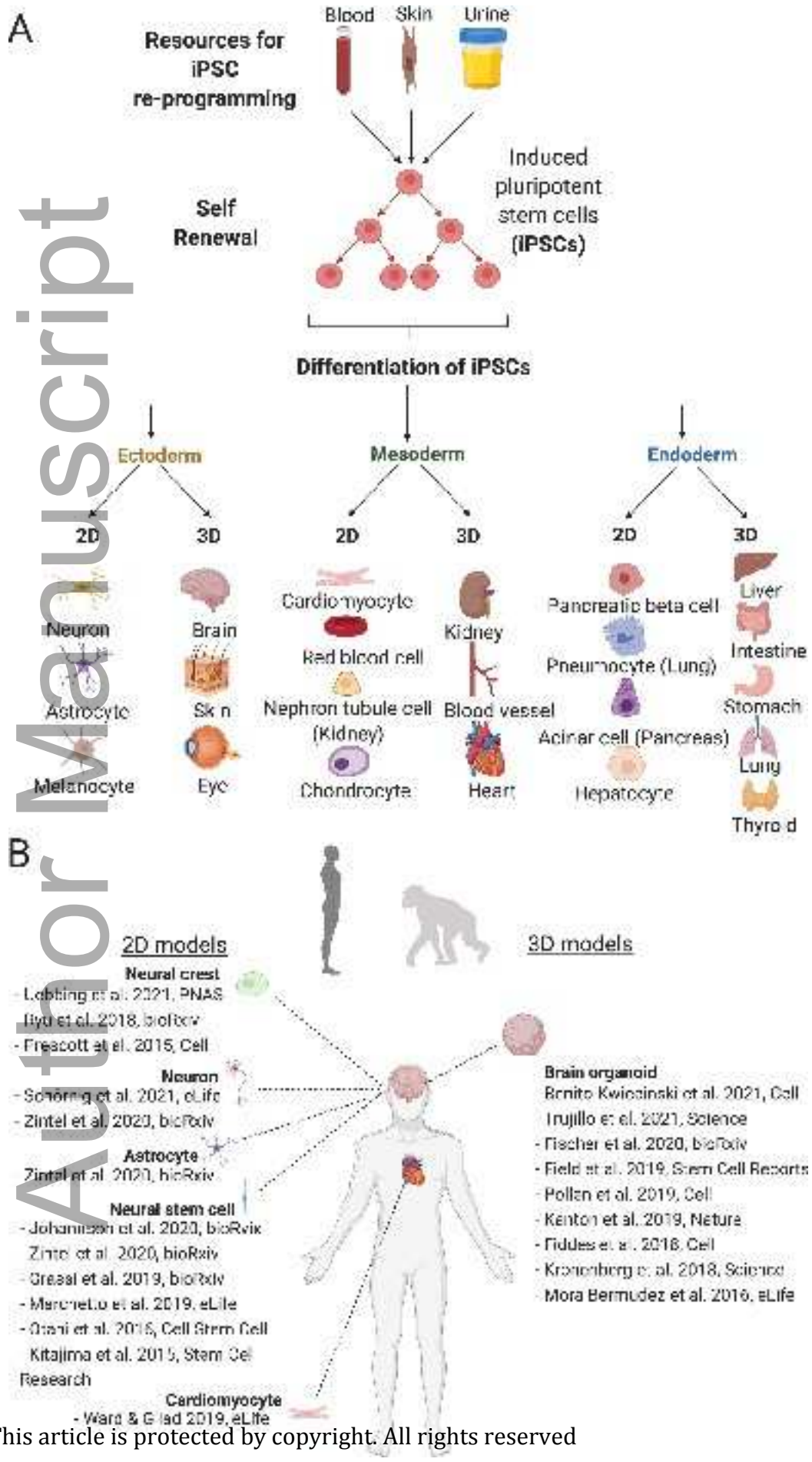
Figure 3: Frequently applied technologies in the context of iPSC research. (A) Massively parallel reporter assays. A high-throughput methodology that allows the simultaneous testing of large numbers of DNA variants for their gene regulatory potential. Thousands of small oligonucleotides (~150-200 bp) each containing a single DNA variant of interest are synthesised in bulk and assembled into a traditional reporter assay plasmid construct, so that their ability to drive expression of a reporter marker can be tested. They are then associated with unique DNA barcodes and transfected in bulk into cells. By quantifying the abundance of each barcode through RNA-sequencing, each variant's regulatory potential can be quantified. **(B) Expression quantitative trait loci.** An association between genotype at a given SNP and expression levels of a nearby gene. Individuals are stratified by genotype at the SNP of interest, and then a simple test compares expression levels of a gene across the three possible genotype classes (each homozygote and the heterozygote). Variants that are associated with significant expression differences between genotype classes are referred to as expression quantitative trait loci (eQTL). **(C) Single-cell sequencing**

technologies. In instances where tissues represent a heterogeneous mix of different cell types, single-cell technologies are often used to measure the expression profiles in individual cells and quantify the variability in a cell-type specific manner. Tissues must first be dissociated into a single cell suspension, followed by the sequencing of the transcriptomes of single cells. The reconstructed expression profiles based on the sequencing data can then be used to group cells from the same cell type using clustering approaches such as t-distributed stochastic neighbor embedding (t-SNE) or Uniform Manifold Approximation and Projection (UMAP).

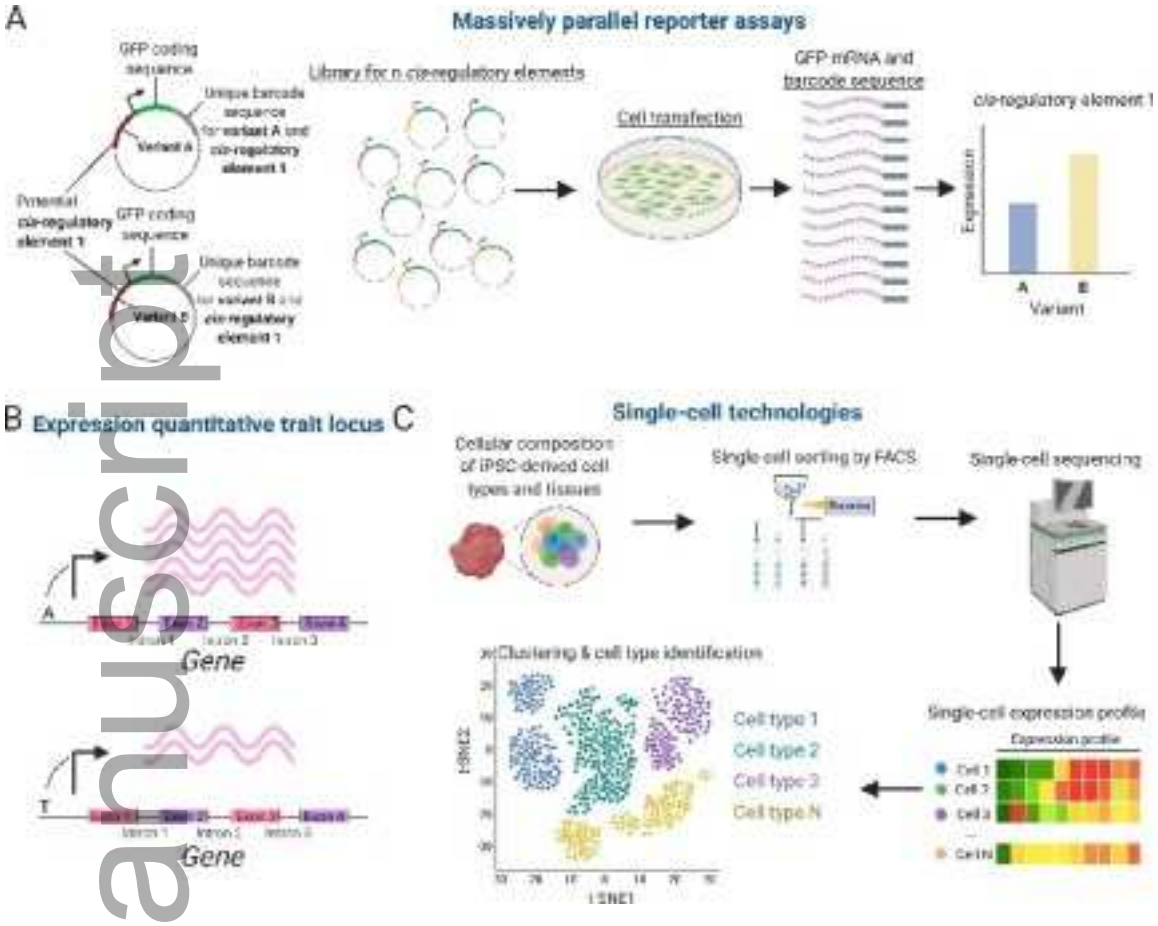
Figure 4: Future challenges in the application of iPSC models to study human evolution. (A) The developmental range represented by iPSC-derived tissue models. The level of maturation of iPSC-derived 2D and 3D models is almost exclusively limited to mirror the human prenatal stage. An example of brain organoids shows that their morphological and transcriptional profile is most similar to the human fetal brain in parts of the first and second trimester. **(B) Statistical power differences between inter and intra-species comparisons.** The number of sufficient samples needed for a robust differential expression analysis is an order of magnitude higher in human intra-species analyses compared to differential expression analyses between humans and other great apes. **(C) Schematic illustration of the concept of linkage disequilibrium.** A genomic phenomena in which genetic variants with a shared history co-occur until recombination separates them. LD between genetic variants hinders the prediction of causal variants in association studies. Evolutionary processes that are associated with increased levels of LD include positive and negative selection, as well as recent admixture. **(D) Limited genetic diversity in human stem cell banks.** The ancestry of stem cell donors in two large stem cell banks, the HipSci [49] and iPSCORE [48] is displayed.



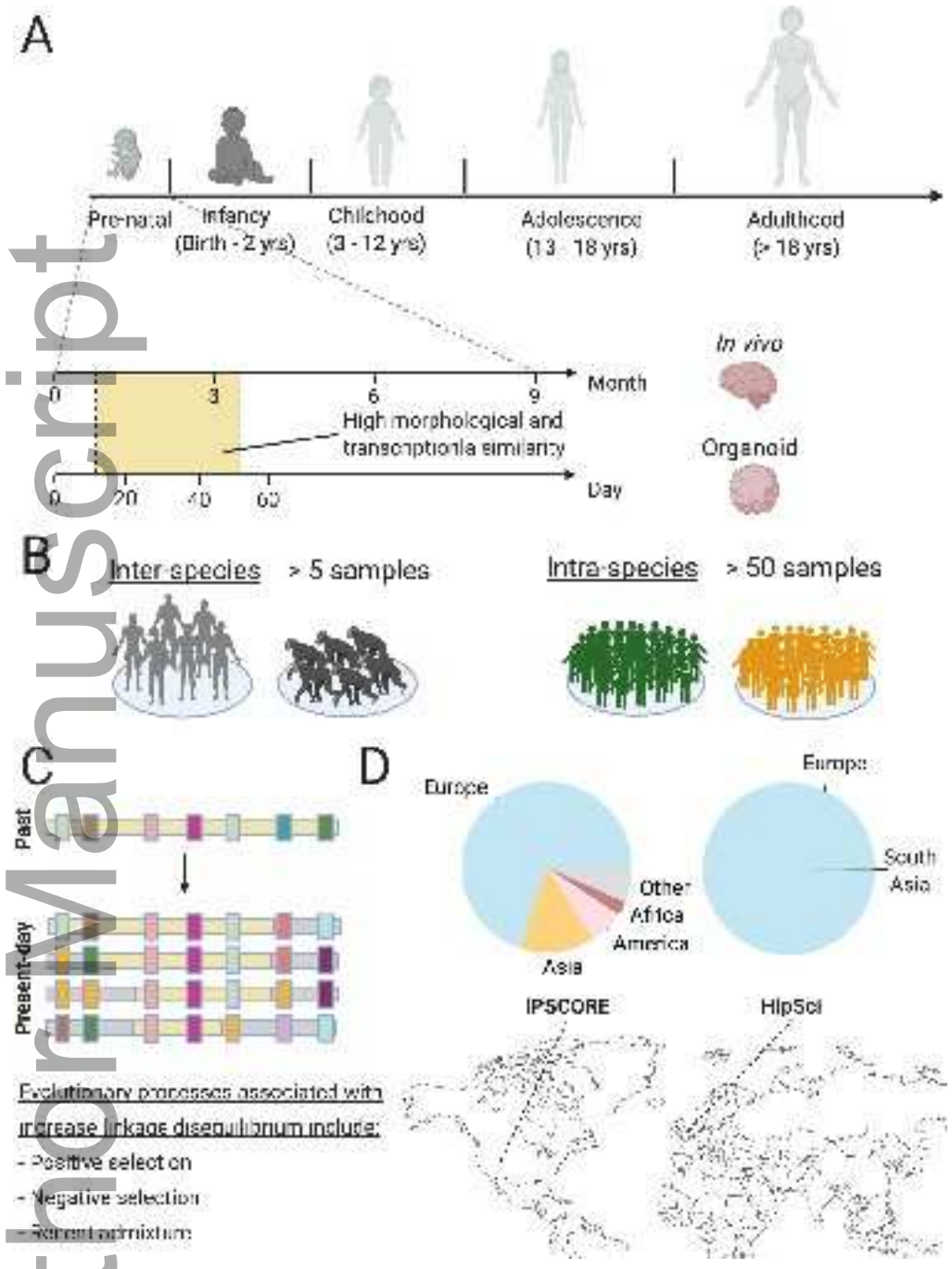
febs_15885_f1.png



febs_15885_f2.png



febs_15885_f3.png



febs_15885_f4.png