

REVIEW ARTICLE **OPEN**

Grave-to-cradle: human embryonic lineage tracing from the postmortem body

Seock Hwan Choi^{1,2,6}, Eu Jeong Ku^{3,6}, Yujin Angelina Choi⁴ and Ji Won Oh⁵ ^{4,5} 

© The Author(s) 2022

Curiosity concerning the process of human creation has been around for a long time. Relevant questions seemed to be resolved with the knowledge of how cells divide after fertilization obtained through in vitro fertilization experiments. However, we still do not know how human life is created at the cellular level. Recently, the value of cadavers as a resource from which to obtain “normal” cells and tissues has been established, and human research using postmortem bodies has attracted growing scientific attention. As the human genome can be analyzed at the level of nucleotides through whole-genome sequencing, individual cells in a postmortem body can be traced back to determine what developmental processes have transpired from fertilization. These retrospective lineage tracing studies have answered several unsolved questions on how humans are created. This review covers the methodologies utilized in lineage tracing research in a historical context and the conceptual basis for reconstructing the division history of cells in a retrospective manner using postzygotic somatic variants in postmortem tissue. We further highlight answers that postmortem research could potentially address and discuss issues that wait to be solved in the future.

Experimental & Molecular Medicine (2023) 55:13–21; <https://doi.org/10.1038/s12276-022-00912-y>

INTRODUCTION

How a person is created and how individual organs develop have long remained unanswered questions. Leonardo da Vinci left a sketch of how an embryo was formed in 1510 (Fig. 1a). Andreas Vesalius drew an intact fetus with umbilical cord and placenta in his book, *De humani corporis fabrica libri septem* (“On the Fabric of the Human Body in Seven Books”), in both the original (1543) and second editions (1555) (Fig. 1c, d). In 1554, based on the teachings of Aristotle, the fetus was believed to be the product of the coagulum formed by blood and semen, an idea that was passed down until the 18th century (Fig. 1b). Uncertainty and assumptions about how a person is created were clarified, to some extent, macroscopically through the work of Swiss embryologist Wilhelm His¹ (Fig. 1e), the Carnegie Stage study in 1942 by Streeter² and later in 1987 by O’Rahilly and Müller³. Therein, the stages of embryo development were divided, in detail, into 23 stages, covering fertilization to the gestational age of 10 weeks.

Advances in modern embryologic research have been bolstered by advances in in vitro fertilization (IVF) technology, which Robert Edward developed for the treatment of infertility⁴. IVF involves fertilization of human eggs with sperm outside the body and the artificial insertion of the embryo into a uterus to give birth to a human being. The inherent nature of IVF has made it possible for scientists to observe fertilized eggs externally. From this, detailed knowledge of developmental processes after fertilization, especially cleavage, and the processes of making morulae and blastocysts have been obtained. However, the idea that individuals can be created through these artificial processes, as well as

the concept that IVF cells can become human beings, have raised ethical concerns. Accordingly, a 14-day rule, permitting the study of IVF embryos only up to 14 days after fertilization or until a primitive streak occurs, was established in 1979⁵. This 14-day rule is agreed upon between the scientific community and social ethics leaders and is enforced in all IRB regulations. Thus, human embryo research into developmental processes occurring later than 2 weeks after fertilization, including gastrulation, organogenesis, somite development, neurulation, and the development of limbs and genitalia, has had to seek other experimental directions.

In this paper, we intend to discuss lineage tracing and the process of conducting human embryology research using recently developed next-generation sequencing (NGS) technology, with a focus on postmortem human tissues. In particular, we cover the development of lineage tracing tools in embryology, including the use of somatic variants, in detail. Additionally, we address the biological and scientific background of how this development was possible, as well as challenges and hypotheses that need to be solved in the future.

Lineage tracing using somatic variants

Lineage tracing is defined as the identification of all progenies of a single cell and how each individual cell divides and determines its fate⁶. In animal models, experimental interventions have made it possible to conduct sophisticated and meticulous lineage tracking with the support of many repetitive experiments. For example, Conklin’s work in 1905 showed how a multicellular organism (ascidian egg) develops through several cell divisions⁷. In

¹Department of Urology, School of Medicine, Kyungpook National University, Daegu, Republic of Korea. ²Bio-Medical Research Institute, Kyungpook National University Hospital, Daegu, Republic of Korea. ³Department of Internal Medicine, Chungbuk National University Hospital and Chungbuk National University, College of Medicine, Cheongju, Republic of Korea. ⁴Department of Anatomy, Yonsei University College of Medicine, Seoul, Republic of Korea. ⁵Brain Korea 21 PLUS Project for Medical Science, Yonsei University College of Medicine, Seoul, Republic of Korea. ⁶These authors contributed equally: Seock Hwan Choi, Eu Jeong Ku. [✉]email: G1@yuhs.ac

Received: 4 April 2022 Revised: 2 September 2022 Accepted: 13 October 2022

Published online: 4 January 2023

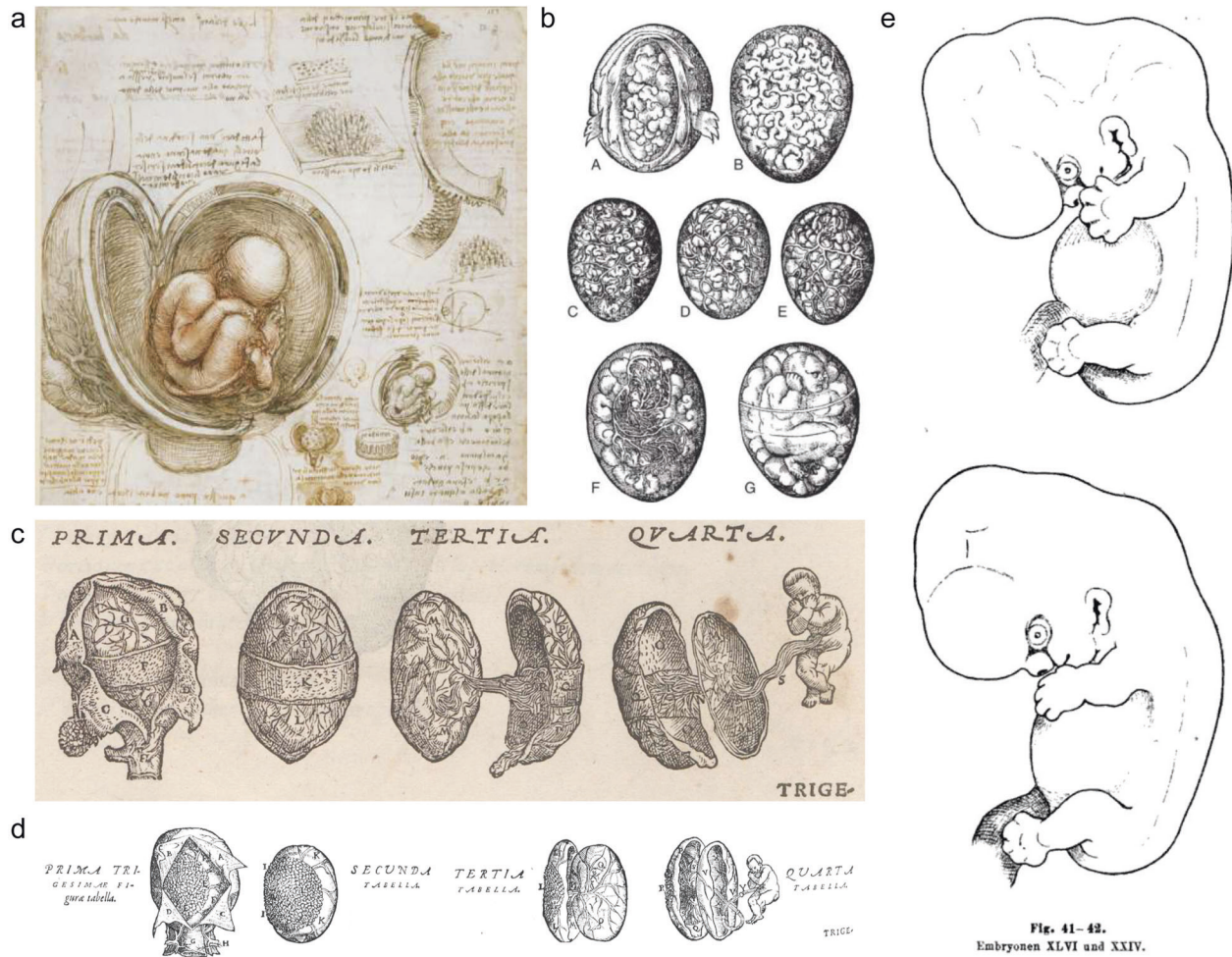


Fig. 1 Old drawings of the human embryo. **a** Fetus in the womb from Leonardo da Vinci's sketches and notes, 1510. **b** The fetal creation process from Jacob Rueff's *De Conceptu et Generatione Hominis*, 1554. **c** The placenta and the fetus drawn from *De humani corporis fabrica libri septem*, original edition, 1543, and **(d)** second edition, 1555. **e** Drawing of the embryo from His in *Anatomie menschlicher Embryonen*, 1880.

particular, he investigated the development of early embryonic cells by tracking the intrinsic yellow crescent cytoplasm observed in muscle and mesenchymal lineage cells in ascidian eggs under specific optical conditions⁸. Developmental processes from a single fertilized egg to gastrulation were tracked, revealing time points for the differentiation of the endoderm, ectoderm, and mesoderm. He also proved the concept of "cell lineage" based on the study of annelids by Edmund Wilson, who suggested that specific cells have their individual "fates" before their division in succession⁹.

In 1983, a milestone study traced the entire process from fertilization to adulthood through lineage tracing. Sir John Sulston and his colleagues reported how the early development of *Caenorhabditis elegans* proceeded and when cells' fates were determined^{10,11}. They found that cells can spontaneously disappear, and upon confirming that these were not errors of the experimental observation, they arrived at the concept of programmed cell death for the first time.

Traditional studies of cancer have indicated that mutations in tumor suppressor genes and oncogenes can lead to cancer. Mutations in these genes are collectively called driver mutations. This concept of driver mutations accelerated the study of how normal cells change to tumor cells in cancer¹². However, when various cancers were analyzed through NGS, researchers discovered, in addition to mutations in major cancer-related genes, mutations throughout the entire genome: in 1989, mutations were

found to be present in various areas of the noncoding region of chromosome 17 in astrocytoma¹³. Although the function of these mutations has not been clearly elucidated thus far, these mutations have been called "passenger" mutations, in contrast to driver mutations¹⁴.

The fact that such mutations exist in cancer cells has enabled cancer biologists to conduct retrospective lineage tracing in patients with multiple cancers. Taking advantage of the simple fact that mutations are inherited by daughter cells, it seems possible to reconstruct the "life history" of cancer cells through mutations. The challenge of this was first reported in 2012¹⁵. To investigate intratumor heterogeneity, the authors utilized somatic variants in individual cancer cells, making a phylogenetic tree of cancers based on patterns of shared and private mutations with temporal dynamics of the somatic mutations. The tree visualized how and when metastases evolved at the cellular level and traced them back to the progression from normal tissue to cancer. As this methodology requires only dissected cancer tissues from patients without any experimental intervention, retrospective lineage tracing of human cells is possible at the in vivo level without ethical concerns. Similar methodologies have been applied to reconstruct the history of human cancer cells using somatic variants underlying cancer cell heterogeneity¹⁶ in many different cancers, including leukemia¹⁷, glioblastoma¹⁸, breast cancer^{19,20}, myeloproliferative neoplasms²¹, muscle-invasive bladder cancer²², kidney tumor²³, and adenoma²⁴.

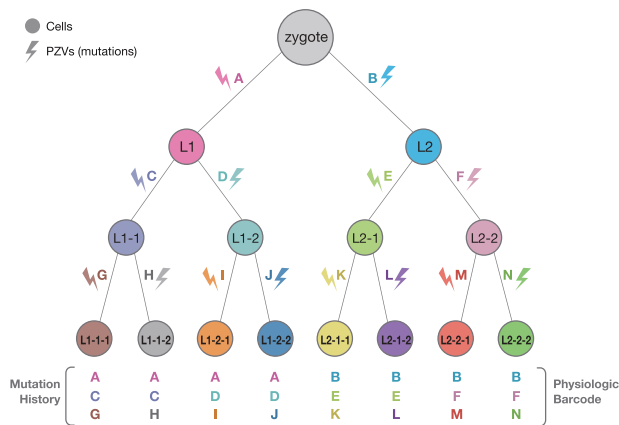


Fig. 2 A theoretical example of sequential mutation events of the first three cell divisions. When the zygote divides into two progenies, each progeny having a single different mutation, A and B, these mutations of lineages 1 (L1) and 2 (L2) can be distinguished. After two more divisions, each cell has different combinations of sequential mutation histories. Based on shared and private individual mutations, the phylogenetic tree of each cell can be reconstructed.

In general, the frequency of somatic mutations in cancer cells increases because of driver mutations^{25–27}. However, studies of cancer and other sporadic diseases have revealed that many cancer-free controls exhibit somatic mosaicism, even under normal conditions²⁸. Eventually, researchers discovered postzygotic variants (PZVs) in early embryonic stages. For example, long interspersed nuclear element-1 (LINE-1) retrotransposons can be altered in terms of copy number during embryogenesis²⁹, resulting in mosaicism of neuronal progenitor cells and adult brain tissue³⁰. Structural variations, such as copy number variations, have also been reported in normal tissues³¹. Additionally, during cleavage of the human embryo, chromosomal instability is relatively common, both pathologically and physiologically^{32,33}.

Mosaicism is a state in which different genotypes (DNA) exist within an individual³⁴. It is mainly known to be involved in pathological diseases of the human body. In the case of skin mosaicism, due to a significant phenotype compared to other normal skin tissues, it becomes much more visible. Indeed, somatic mosaicism, or PZVs, can occur throughout the entirety of human life³⁵. Research has confirmed that the somatic mutations that create these mosaics are randomly distributed across the entire human genome³⁶. The average base-substitution somatic mutation rate in normal tissues is estimated at 3.05 mutations per cell (diploid genome) per cell division (pcpcd) in retinal development, 0.88 mutations pcpcd in intestinal epithelial cells, 4.54 mutations pcpcd in B and T lymphocytes, and 1.05 mutations pcpcd in fibroblasts³⁷. Our previous study reported 3.8 (range, 1.4 to 6.3) mutations pcpcd in the earliest cell division up to four cell stages, which is similar to other reports, but 1.2 (range, 0.8 to 1.9) mutations pcpcd after four cell stages in early human embryogenesis^{38,39}. In normal tissue, the estimated rate of mutation accumulation is 33 per year per diploid genome in the liver,⁴⁰ 36 per year per diploid genome in adult stem cells of the colon, small intestine, and liver;⁴¹ 14.2 per year per diploid genome in hematopoietic stem cells and multipotent progenitor cells in bone marrow;⁴² and 43.6 per year per diploid genome in colorectal epithelial cells⁴³.

Theoretically, if a single mutation accumulates at each cell division, as cell division continues and other mutations accumulate during each division, all of the cells of an individual can have different DNA (Fig. 2). Since these cells divide thousands of times over a lifetime, it can be inferred that the

mutations accumulated during each division ultimately result in each cell having its own unique DNA pattern. In the case of PZVs occurring in early embryogenesis, since a large number of progeny cells contain those mutations, it is possible to completely reconstruct the initial developmental processes of these cells depending on whether they have shared mutations or not. That is, with each division throughout the life of an individual, genetic mutations accumulate, suggesting that the history of cell division can be tracked⁴⁴. This concept was first validated in mice using clones obtained from the organoid technique⁴⁵. Eventually, PZVs were found to enable retrospective lineage tracing in the study of human embryogenesis^{38,46–51}.

Postmortem tissue for advanced research

In most animals, including humans, a fertilized egg keeps dividing and eventually differentiates into the ectoderm, mesoderm, and endoderm during development. Thus, to correctly reconstruct a cell's lifetime, cells from these three germ layers must be obtained within an individual to reconstruct their processes. For example, ectoderm-derived tissue, such as corneal epithelial cells and skin epithelium; mesoderm-derived tissue, such as blood cells and fibroblasts; and endoderm-derived tissue, such as colon crypt cells and hepatocytes, should be simultaneously obtained from one individual for the proper reconstruction of embryogenesis. This experimental strategy can determine the timing of divergence from each germ layer. However, for samples from a surgical patient obtained in general medical research with informed consent, it is quite challenging to obtain cells of the above three germ layers at the same time. On the other hand, in the case of a donated postmortem body, it is possible to secure enough cells derived from these three germ layers in one individual.

Most human cadavers in the past were utilized for anatomy education, understanding structure, and solving superficial pathologies, while in modern times, advanced biological methodologies have been applied to better understand the *in vivo* state of the human body^{52,53}. Efforts such as the Genotype-Tissue Expression (GTEx) project, initiated by the National Institutes of Health to collect various postmortem normal tissues that reflect the *in vivo* state and cannot be obtained from patients, have increased our understanding of molecular biological aspects, specifically at the transcriptome level⁵⁴. The GTEx project has gathered 17,382 samples across 54 nondiseased tissues from 948 individuals, including data from gene expression profiling, whole-genome sequencing (WGS), and RNA-seq⁵⁵. They also evaluated the usefulness of postmortem intervals at the level of the transcriptome. Data from 1 h to 27 h after death are reported in detail concerning how tissues change over time at the molecular level and how transcriptomic signals are changed⁵⁶. According to studies of gene expression profiles in postmortem tissue, the postmortem interval (PMI) of RNA degradation differs between tissues, genotypes, and specific genes; for example, the cerebellum and pituitary gland show relative stability in terms of PMI-associated genes, while esophageal mucosa and whole blood exhibit many altered expression patterns of PMI-associated genes.

Transcriptome data fluctuate with postmortem RNA degradation, which requires strict control over the "normal" structure, but DNA in postmortem tissues is preserved safely. However, in lineage tracing of embryogenesis, it is necessary to secure a sufficient amount of intracellular DNA originating from a single cell or to cultivate a clonal cell from postmortem tissues *in vitro*. Since error rates for nonreference genotype cells range from 0.1% to 0.6%, depending on the depth of coverage or the platform^{57,58}, finding a somatic variant that occurs less frequently than the error rate of NGS requires the use of specialized experimental strategies.

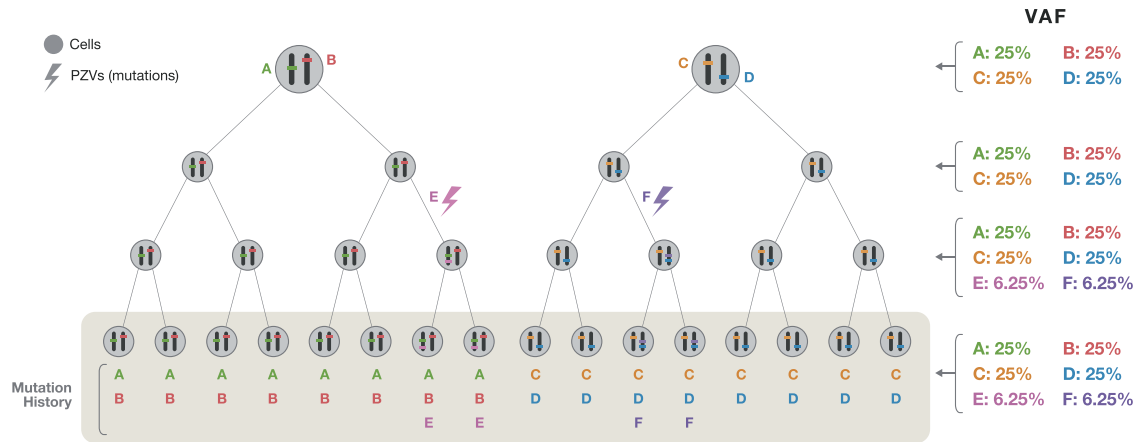


Fig. 3 A theoretical example of expansion from two different lineage cells. When cells in different lineages divide equally, the variant allele frequency is much lower than 50%. In this example, 25% of all genuine mutations are represented, except for culture-induced mutations (E and F), which are lower than 25%. In this case, it is relatively difficult to determine in which cell the mutation originated. If the origin of cells is polyclonal in LCM, the situation is much more complicated.

METHODS FOR THE DETECTION OF EARLY EMBRYONIC MUTATIONS (EEMS) FOR RETROSPECTIVE LINEAGE TRACING OF POSTMORTEM TISSUES

Two main challenges hinder embryonic lineage tracing using the postmortem body. First, the genome of single cells must be obtained from a variety of different anatomical areas that accurately reflect the developmental processes in the three germ layers. Second, because the incidence rate of errors in NGS exceeds the incidence rate of PZV events *in vivo*, which is from 0.1 to 3 mutations per cell per cell division³⁹, it is very important to set up an experimental strategy to pinpoint the somatic variants accurately in wet experimental designs and bioinformatical analysis. Currently, to proceed with retrospective lineage tracing of postmortem tissues, somatic variants are explored in three different ways.

Clonal expansion *in vitro* from various anatomical locations and organs

Since clonal expansion originates from a single cell, “expansion as a replication tool” retains all mutations that reflect the history of a cell (Fig. 3). Clonal expansion for retrospective lineage tracing requires not only well-trained developmental biologists who have expertise in culture but also anatomists who can dissect and quickly preserve proper organs. Precise primary culture methods and advanced clonal expansion culture techniques are also essential. At the same time, to obtain a sufficient level of DNA from a clone, it is necessary to culture cells for >2 months. These clones produced by clonal expansion methods can provide relatively pure signals of PZVs with minimal artifacts in NGS. To date, protocols have been established for obtaining clonal lines of mesodermal origin, such as skin fibroblasts^{38,49} and muscle myocytes from head to extremities³⁸, hair dermal papilla stem cells⁵⁹, and hematopoietic stem cells;^{60,61} endodermal origin, such as organoids of the stomach⁴⁵, liver⁶², small intestine, colon⁴¹, and prostate;⁶³ and ectodermal origin, such as cornea keratinocytes³⁸ and neuronal progenitor cells⁴⁶ of the fetus. Although it is very challenging and labor-intensive to obtain large numbers of clones from an individual, especially a cadaver, it is believed that a clone can provide the purest somatic mutation signals of the entire history of cellular division for WGS, compared to other methods, such as laser capture microdissection (LCM) and single-cell DNA sequencing.

Recently, an attempt has been made to use induced pluripotent stem cells (iPSCs) to obtain clonal cultures of fibroblasts⁴⁹. The researchers used transient episomal vectors expressing human Oct4, Sox2, Lin-28, Klf4, and L-myc to culture the iPSCs of

fibroblasts⁴⁹. Although iPSC clones were created using fibroblasts for the proof of concept for the detection of early PZVs, iPSC techniques could be utilized for cells with limited proliferative potential. Based on our experience, it can be difficult to obtain clones from ectodermal origins with a sufficient amount of DNA for WGS. Though the success rate of primary culture per location of the epithelium is 43%–62%, which is similar to that of fibroblasts, the success rate of clonal expansion per location of the epithelium is relatively very low at approximately 5%–8%, while that of fibroblasts is 30%–45%. These empirical results indicate the need to establish different strategies for the epithelium, where there is a low potential of establishing clones. In this respect, iPSC techniques may play a pivotal role in minimized artifact DNA replication machinery for low-potential proliferative cells, such as epithelial cells or differentiated cells, not for pluripotency purposes.

Clonal expansion *in vitro* has some disadvantages: it takes tremendous initial effort, and only cells that can grow for a long time are obtained due to selection bias. In addition, it is a prospective experiment that must be carried out within 36 hours after death. However, once obtained, all the mutations a cell has developed in its lifetime can be investigated, including those in mitochondrial DNA. Clonal expansion may generate a few (<10) culture-induced mutations that are shared by most of the cultured cells from individual clones. We have found that the vast majority of mutations identified in clones are genuine³⁸.

Laser capture microdissection (LCM) from various anatomical locations and organs

The second method is LCM^{64,65}. LCM is commonly used in retrospective lineage tracing in the liver^{40,66}, colorectal epithelial cells⁴³, colon⁶⁷, endometrial epithelium⁶⁸, placenta^{51,69}, diverse glands⁴⁸, and normal tissue mutational profiles⁷⁰. The scientific rationale is that certain stem cell populations grow clonally within a specific niche *in vivo*^{71–73}. LCM can physically separate these natural cellular clusters of dominant clones and provide solid information on somatic mutations among WGS clones.

Unlike clonal expansion, this method has some advantages: it can be performed histologically, it requires less labor, and it can be used for various *in vivo* clonal expansion tissues. Additionally, LCM methods have the advantages of preserved spatial distribution and histopathological features of cells within a tissue⁶⁵. At the same time, since the LCM experiment can be carried out after storage of the tissue, it has the strength of being able to be performed electively at the desired time.

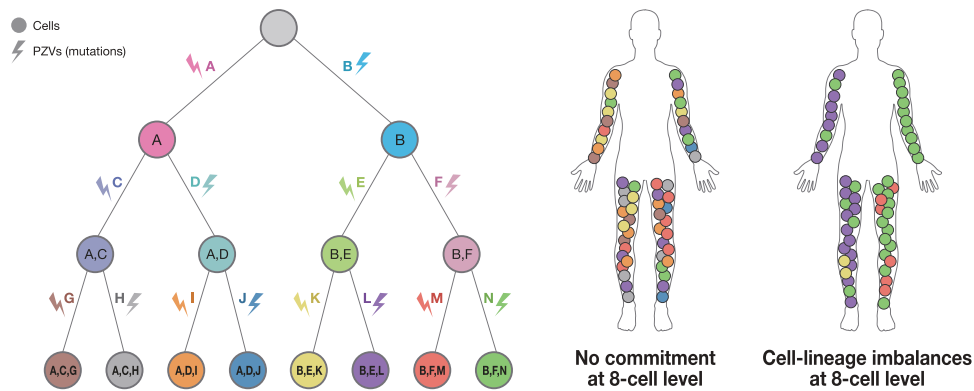


Fig. 4 A theoretical example of cell-lineage imbalance. If a specific cell contributes more when making a certain tissue or organ, the proportion of that cell in the tissue will be relatively larger than that of other cells. Conversely, if every cell contributes equally to making a particular tissue, it can be inferred that the cell's fate has not yet been decided. In this case, since the fate is determined as division progresses, the division stage of the cell is also important. This principle makes it possible to infer when each cell's fate has been decided.

However, LCM cannot be applied to several types of tissues or cells, such as fibroblasts and brain and heart tissue, that do not naturally create clone structures *in vivo*. Although an apparent clone structure may be seen on histology, this could be the result of contamination by other cells, resulting in polyclonal areas that waste money or require sophisticated bioinformatical analysis (Fig. 3). In addition, the quality of DNA could be relatively low and bioinformatically difficult to analyze due to signal noise. Nevertheless, LCM methods are the leading methods for the study of embryogenesis using somatic variants and are expected to be developed for almost every internal organ and endodermal lineage for higher resolution of retrospective lineage tracing maps.

Single-cell DNA WGS using whole-genome amplification (WGA) with or without high-depth bulk sequencing (>200X)

A third method is WGA of a single cell^{47,74}. This method has the advantage of being able to acquire a large number of cells simultaneously and is relatively easy to perform without selection bias. At the same time, it provides flexibility for a diverse range of cells, including nondividing cells, such as neurons. However, it is difficult to distinguish between genuine PZVs and *de novo* mutations occurring in the early stage of the amplification process⁷⁵. At the same time, "allele dropout" and intrinsic errors during amplification are relatively common⁷⁶, and therefore it is not preferred as a single method for detecting PZVs.

To overcome these limitations, a recent study combined high-depth bulk sequencing data to detect shared PZVs and then confirmed the results via single-cell target sequencing for the reconstruction of early embryogenesis^{50,77}. Using the fact that early EEMs are shared in many cells of the body, resulting in a high variant allele frequency (VAF) in bulk tissue in WGS, bulk sequencing is performed at a depth of 200X or more to find EEMs with statistical confidence^{78,79}. Then, the exact locus positions of the found EEMs are validated in a large number of cortical single cells, leading to the precise embryonic lineages of each of these cells.

This bulk sequencing approach cannot lead to the discovery of PZVs as physiological barcodes of the entire lifespan. However, anatomically refined high-resolution bulk sequencing of more than two samples from one individual can complement clonal strategies and provide information on early lineage⁷⁹. One study sought to determine shared PZVs for the reconstruction of early lineages, which are relatively frequent⁸⁰. Indeed, WGS at a depth of approximately 100X can detect somatic variants during early embryogenesis, which have different mutational profiles, such as replication timing and chromatin status of the later tissue-specific clonal expansion⁸¹.

Several single-cell WGA methods have been developed thus far⁷⁶, and various computational approaches have been applied to

identify somatic variants⁸². Among them, the recently developed isothermal WGA method, primary template-directed amplification (PTA), is attracting attention⁸³. Since PTA is relatively more uniform and accurate than the existing single-cell WGA methods, such as multiple displacement amplification (MDA), the accuracy and sensitivity of somatic variant calling are improved^{83,84}. As a result, its use is increasing in neuron and brain research, where there is no histological niche to perform LCM or *in vitro* clonal expansion^{78,85}. PTA is expected to be applied in other organs in various experimental strategies, along with bulk WGS sequencing.

Findings of Human Embryogenesis from Postmortem Lineage Tracing

Proof-of-concept studies for retrospective lineage tracing, which used PZVs as physiological barcodes to reflect cellular lineage history, have provided potential answers to unsolved questions in human embryology. Recent studies have presented evidence for various human embryologic hypotheses concerning when the fates of the trophoblast and inner cell mass (ICM) are determined, embryologic asymmetry in the human body, and in what order each cell determines their fate in terms of the three germ layers, anatomical right-left commitment, and organogenesis.

Based on our data with mathematical simulations, the time point for the specification between the trophoblast and ICM occurs following the third or fourth division of a zygote: five of eight cells at the third division (or 10 of 16 cells at the fourth division) are considered to be of the trophoblast lineage, which forms the placenta, while the remaining three cells contribute to the fetus proper³⁸. These three cells eventually develop into the ICM, differentiating into every organ of the human body. Because of an odd number of founder cells in the fetus, the cellular contribution to the whole body is considered to be asymmetric (approximately 1:2 at the first division of the zygote). Based on seven postmortem bodies in our study³⁸ and other related embryogenesis studies, including five fetal post-mortem bodies^{46,69}, six postmortem bodies^{48,50}, and three living individuals^{49,61}, we can statistically generalize the scientific evidence of this asymmetric contribution of the founder cells to the human body according to the patterns of shared and unique PZVs in each clone within an individual.

Based on the analysis of PZVs of clones within an individual, the shared patterns of PZVs among clones as a marker for early embryonic lineage could be utilized for further analysis. For example, when mutually exclusive PZVs are identified, they could be used to identify the specific lineages of clones carrying these PZVs. Because the specific genomic position of PZVs can be defined for specific lineages after the analysis of clones (capture

taking part in yolk sac hematopoiesis. The epiblast is divided into three germ layers and an amnion, which will later become a partial source of PGCs. Among the three germ layers, the fate of the ectoderm is determined first, after which the mesoderm and endoderm seem to separate. This process is expected to occur approximately in the 3rd–7th cell generation, after which organogenesis begins and regional imbalance occurs. The composition of the early embryonic lineages seems to be changed by external stimuli and intrinsic properties, leading to cellular stochasticity.

Retrospective embryological research using postmortem bodies has been conducted on approximately 25 people. Intrinsic mutation rates are very important in studying and precisely reconstructing phylogenetic embryologic lineages. In other words, a key point to ensuring that an experimental plan is successful is that mutations must occur at a sufficiently high frequency. Our experience and other studies have indicated that mutations in early human development occur at an average of 0.1 to 3 mutations pcpd. However, some individuals have many early mutations, while others have very few. Theoretically, in instances of mutation rates less than 1 pcpd, the full reconstruction of embryogenesis would be very challenging due to the lack of PZVs to mark the cellular lineages, resulting in polytomy, a single node with more than two descendent branches, not dichotomy, a node with only two descendent branches. Indeed, many cases exhibit multifurcation, not bifurcation, in the early stages due to the low mutation rate, leading to uninformative division. Although instances of lower mutation rates can be compensated for by solid statistical methods, future studies of a more precise estimate of specific commitment at the level of cellular generation require special experimental strategies.

Future directions and unresolved questions

The cellular lineage field of human developmental biology is still in its infancy. Although the commitment trajectory of 40–277 clones with precise phylogenetic lineage trees was determined from various individuals, the number of cells was insufficient to reconstruct the commitment fate of human blastomeres to the entire human body. In addition, other studies have revealed the developmental processes of humans by partially focusing on specific anatomical organs either from living individuals or patients, such as the liver^{40,66}, esophagus⁹⁰, endometrium⁶⁸, blood⁶⁰, and brain⁵⁰, with similar experimental strategies, although they explain the process only partially. The more we understand about early human embryogenesis, the more apparent the necessity of a higher resolution reference trajectory map becomes.

Although many questions have been answered, there are still many problems to be solved. First, it is necessary to pinpoint the exact time of commitment of each lineage in different developmental stages. This kind of work will present a clear map of the time when stochasticity is eliminated, resulting in commitment to a specific destiny. For example, a fibroblast is a unique mesodermal cell that originates from the nascent mesoderm at gastrulation. Since fibroblasts are transferrable during developmental stages and are located throughout the body in an adult, it is still inconclusive whether fibroblasts adjacent to other lineages are deterministic at the developmental stage or stochastic depending on the anatomical locations.

Second, as the behavior of “normal cells” changes with age, the exact evolutionary processes of cell division and the intrinsic differences in mutational rate in different cell lineages need to be investigated. Specifically, studies on which normal cells become cancerous and, conversely, which cells are more resistant to various mutations and maintain a normal state need to be conducted. In addition, it is necessary to study whether mutations that initially appear to have no function actually play a role. With age, some cells accumulate more mutations, and others carry relatively few mutations. It is crucial to ascertain whether these

differences originate from a stage of stochastic development or are a property of chance acquired during development.

Third, questions concerning organogenesis should be addressed in more detail. Organogenesis is the phase of development that starts at the end of gastrulation after the commitment of the three germ layers. If sufficient clones are isolated from a specifically targeted organ, combined with other blastomere lineages from the same individual, researchers can determine when the most recent common ancestor (MRCA) commits to an organ and the minimum cell numbers of the MRCA at that stage. As the minimum pluripotent cell number is critical to mimic a physiological artificial human organ in vitro with a scaffold in regenerative medicine, this can provide insight into human embryology and practical evidence for the newly emerging field of mini organs.

Finally, a study of approximately 25 people thus far has yielded some generalizable knowledge, but it may still be an insufficient number to generalize to the human species. There is a need to study a much larger number of postmortem bodies to produce statistically significant data. In addition, to conclude the fate of each lineage commitment and to explore the unique developmental processes of the human species, it seems necessary to secure and analyze a larger number of clones of the various anatomical tissues from an individual, who contains more than 30 trillion cells.

CONCLUSION

In this review, we summarized retrospective lineage tracing methods applied to the whole genome using postmortem human bodies. A massive number of experimental and bioinformatical methods are being developed for this field, necessitating understanding of the advantages and disadvantages of different experimental approaches, such as in vitro clonal expansion, LCM, and single-cell WGA with high-depth bulk sequencing for the detection of EEMs. In the future, utilization of a combination of two or more methods should be attempted for a better understanding of human embryology. Different groups use different terms with the same or similar meaning in this field, for example, somatic mutations, somatic variants, mosaic mutations (MMs), early mosaic mutations (EMMs), postzygotic variants (PZVs), and embryonic variants. Although there are subtle differences in the meaning of the terms, efforts to unify the terminology for this field will be necessary.

Postmortem body research, which involves obtaining different intact human organs simultaneously, is becoming a new means by which to study human embryology along with various NGS methods. It is paradoxical to be able to trace the birth process, which is the starting point of life, through the irreversible cessation of life in death. We are deeply indebted to those who donated their bodies, cells, and tissues for the advancement of this field.

REFERENCES

1. His, W. *Anatomie menschlicher Embryonen* (Vogel, 1880).
2. Streeter, G. L. Developmental horizons in human embryos. *Contrib. Embryol.* **30**, 213–230 (1942).
3. Müller, F. & O’Rahilly, R. (eds) *Developmental Stages in Human Embryos* (Carnegie Institution of Washington, 1987).
4. Steptoe, P. C. & Edwards, R. G. Birth after the reimplantation of a human embryo. *Lancet* **2**, 366 (1978).
5. Cavaliere, G. A 14-day limit for bioethics: the debate over human embryo research. *BMC Med. Ethics* **18**, 38 (2017).
6. Kretschmar, K. & Watt, F. M. Lineage tracing. *Cell* **148**, 33–45 (2012).
7. Conklin, E. G. *The Organization and Cell-lineage of the Ascidian Egg* (Academy of Natural Sciences, 1905).
8. Jeffery, W. R., Tomlinson, C. R., Brodeur, R. D. & Meier, S. in *Molecular Aspects of Early Development* (eds Malacinski, G. M. & Klein, W. H.) 1–38 (Springer US, 1984).

9. Wilson, E. B. The cell-lineage of Nereis. A contribution to the cytogeny of the annelid body. *J. Morphol.* **6**, 361–480 (1892).
10. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119 (1983).
11. Sulston, J. E. & Horvitz, H. R. Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev. Biol.* **56**, 110–156 (1977).
12. Vogelstein, B. et al. Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
13. James, C. D., Carlom, E., Nordenskjold, M., Collins, V. P. & Cavenee, W. K. Mitotic recombination of chromosome 17 in astrocytomas. *Proc. Natl Acad. Sci. USA* **86**, 2858–2862 (1989).
14. McFarland, C. D. et al. The damaging effect of passenger mutations on cancer progression. *Cancer Res.* **77**, 4763–4772 (2017).
15. Gerlinger, M. et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
16. Kreso, A. & Dick, J. E. Evolution of the cancer stem cell model. *Cell Stem Cell* **14**, 275–291 (2014).
17. Landau, D. A. et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **152**, 714–726 (2013).
18. Sottoriva, A. et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc. Natl Acad. Sci. U. S. A.* **110**, 4009–4014 (2013).
19. Aceto, N. et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* **158**, 1110–1122 (2014).
20. Wang, Y. et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**, 155–160 (2014).
21. Ortmann, C. A. et al. Effect of mutation order on myeloproliferative neoplasms. *N. Engl. J. Med.* **372**, 601–612 (2015).
22. Li, Y. et al. Single-cell sequencing analysis characterizes common and cell-lineage-specific mutations in a muscle-invasive bladder cancer. *Gigascience* **1**, 12 (2012).
23. Xu, X. et al. Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* **148**, 886–895 (2012).
24. Humphries, A. et al. Lineage tracing reveals multipotent stem cells maintain human adenomas and the pattern of clonal expansion in tumor evolution. *Proc. Natl Acad. Sci. USA* **110**, E2490–E2499 (2013).
25. McGranahan, N. & Swanton, C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell* **168**, 613–628 (2017).
26. Garraway, L. A. & Lander, E. S. Lessons from the cancer genome. *Cell* **153**, 17–37 (2013).
27. Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nat. Med.* **10**, 789–799 (2004).
28. Forsberg, L. A., Gisselsson, D. & Dumanski, J. P. Mosaicism in health and disease—clones picking up speed. *Nat. Rev. Genet.* **18**, 128–142 (2017).
29. van den Hurk, J. A. et al. L1 retrotransposition can occur early in human embryonic development. *Hum. Mol. Genet.* **16**, 1587–1592 (2007).
30. Coufal, N. G. et al. L1 retrotransposition in human neural progenitor cells. *Nature* **460**, 1127–1131 (2009).
31. Piotrowski, A. et al. Somatic mosaicism for copy number variation in differentiated human tissues. *Hum. Mutat.* **29**, 1118–1124 (2008).
32. Bielanska, M., Tan, S. L. & Ao, A. Chromosomal mosaicism throughout human preimplantation development in vitro: incidence, type, and relevance to embryo outcome. *Hum. Reprod.* **17**, 413–419 (2002).
33. Vanneste, E. et al. Chromosome instability is common in human cleavage-stage embryos. *Nat. Med.* **15**, 577–583 (2009).
34. De, S. Somatic mosaicism in healthy human tissues. *Trends Genet.* **27**, 217–223 (2011).
35. Thorpe, J., Osei-Owusu, I. A., Avigdor, B. E., Tupler, R. & Pevsner, J. Mosaicism in human health and disease. *Annu. Rev. Genet.* **54**, 487–510 (2020).
36. Kübler, K. et al. Tumor mutational landscape is a record of the pre-malignant state. *bioRxiv* <https://www.biorxiv.org/content/10.1101/517565v1> (2019).
37. Lynch, M. Rate, molecular spectrum, and consequences of human mutation. *Proc. Natl Acad. Sci. USA* **107**, 961–968 (2010).
38. Park, S. et al. Clonal dynamics in early human embryogenesis inferred from somatic mutation. *Nature* **597**, 393–397 (2021).
39. Ju, Y. S. et al. Somatic mutations reveal asymmetric cellular dynamics in the early human embryo. *Nature* **543**, 714–718 (2017).
40. Brunner, S. F. et al. Somatic mutations and clonal dynamics in healthy and cirrhotic human liver. *Nature* **574**, 538–542 (2019).
41. Blokzijl, F. et al. Tissue-specific mutation accumulation in human adult stem cells during life. *Nature* **538**, 260–264 (2016).
42. Osorio, F. G. et al. Somatic mutations reveal lineage relationships and age-related mutagenesis in human hematopoiesis. *Cell Rep.* **25**, 2308–2316.e2304 (2018).
43. Lee-Six, H. et al. The landscape of somatic mutation in normal colorectal epithelial cells. *Nature* **574**, 532–537 (2019).
44. Woodworth, M. B., Girsakis, K. M. & Walsh, C. A. Building a lineage from single cells: genetic techniques for cell lineage tracking. *Nat. Rev. Genet.* **18**, 230–244 (2017).
45. Behjati, S. et al. Genome sequencing of normal cells reveals developmental lineages and mutational processes. *Nature* **513**, 422–425 (2014).
46. Bae, T. et al. Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis. *Science* **359**, 550–555 (2018).
47. Lodato, M. A. et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* **359**, 555–559 (2018).
48. Coorens, T. H. H. et al. Extensive phylogenies of human development inferred from somatic mutations. *Nature* **597**, 387–392 (2021).
49. Fasching, L. et al. Early developmental asymmetries in cell lineage trees in living individuals. *Science* **371**, 1245–1248 (2021).
50. Bizzotto, S. et al. Landmarks of human embryonic development inscribed in somatic mutations. *Science* **371**, 1249–1253 (2021).
51. Coorens, T. H. H. et al. Inherent mosaicism and extensive mutation of human placentas. *Nature* **592**, 80–85 (2021).
52. Ferreira, P. G. et al. The effects of death and post-mortem cold ischemia on human tissue transcriptomes. *Nat. Commun.* **9**, 490 (2018).
53. Palmer, T. D. et al. Cell culture. Progenitor cells from human brain after death. *Nature* **411**, 42–43 (2001).
54. Lonsdale, J. et al. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
55. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
56. Zhu, Y., Wang, L., Yin, Y. & Yang, E. Systematic analysis of gene expression patterns associated with postmortem interval in human tissues. *Sci. Rep.* **7**, 5435 (2017).
57. Wall, J. D. et al. Estimating genotype error rates from high-coverage next-generation sequence data. *Genome Res.* **24**, 1734–1739 (2014).
58. Paskov, K. et al. Estimating sequencing error rates using families. *BioData Min.* **14**, 27 (2021).
59. Kwack, M. H. et al. Establishment and characterization of five immortalized human scalp dermal papilla cell lines. *Biochem. Biophys. Res. Commun.* **496**, 346–351 (2018).
60. Mitchell, E. et al. Clonal dynamics of haematopoiesis across the human lifespan. *Nature* **606**, 343–350 (2022).
61. Lee-Six, H. et al. Population dynamics of normal human blood inferred from somatic mutations. *Nature* **561**, 473–478 (2018).
62. Jager, M. et al. Measuring mutation accumulation in single human adult stem cells by whole-genome sequencing of organoid cultures. *Nat. Protoc.* **13**, 59–78 (2018).
63. Drost, J. et al. Organoid culture systems for prostate epithelial and cancer tissue. *Nat. Protoc.* **11**, 347–358 (2016).
64. Espina, V. et al. Laser-capture microdissection. *Nat. Protoc.* **1**, 586–603 (2006).
65. Ellis, P. et al. Reliable detection of somatic mutations in solid tissues by laser-capture microdissection and low-input DNA sequencing. *Nat. Protoc.* **16**, 841–871 (2021).
66. Ng, S. W. K. et al. Convergent somatic mutations in metabolism genes in chronic liver disease. *Nature* **598**, 473–478 (2021).
67. Olafsson, S. et al. Somatic evolution in non-neoplastic IBD-affected colon. *Cell* **182**, e611 (2020).
68. Moore, L. et al. The mutational landscape of normal human endometrial epithelium. *Nature* **580**, 640–646 (2020).
69. Spencer Chapman, M. et al. Lineage tracing of human development through somatic mutations. *Nature* **595**, 85–90 (2021).
70. Moore, L. et al. The mutational landscape of human somatic and germline cells. *Nature* **597**, 381–386 (2021).
71. Blanpain, C. & Simons, B. D. Unravelling stem cell dynamics by lineage tracing. *Nat. Rev. Mol. Cell Biol.* **14**, 489–502 (2013).
72. Simons, B. D. & Clevers, H. Strategies for homeostatic stem cell self-renewal in adult tissues. *Cell* **145**, 851–862 (2011).
73. Snippet, H. J. et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134–144 (2010).
74. Lodato, M. A. et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science* **350**, 94–98 (2015).
75. Biezuner, T. et al. Comparison of seven single cell whole genome amplification commercial kits using targeted sequencing. *Sci. Rep.* **11**, 17171 (2021).
76. Gawad, C., Koh, W. & Quake, S. R. Single-cell genome sequencing: current state of the science. *Nat. Rev. Genet.* **17**, 175–188 (2016).
77. Huang August, Y. et al. Parallel RNA and DNA analysis after deep sequencing (PRDD-seq) reveals cell type-specific lineage patterns in human brain. *Proc. Natl Acad. Sci. USA* **117**, 13886–13895 (2020).
78. Bae, T. et al. Analysis of somatic mutations in 131 human brains reveals aging-associated hypermutability. *Science* **377**, 511–517 (2022).
79. Breuss, M. W. et al. Somatic mosaicism reveals clonal distributions of neocortical development. *Nature* **604**, 689–696 (2022).
80. Wang, Y. et al. Comprehensive identification of somatic nucleotide variants in human brain tissue. *Genome Biol.* **22**, 92 (2021).

81. Huang, A. Y. et al. Distinctive types of postzygotic single-nucleotide mosaicisms in healthy individuals revealed by genome-wide profiling of multiple organs. *PLoS Genet.* **14**, e1007395 (2018).
82. Valecha, M. & Posada, D. Somatic variant calling from single-cell DNA sequencing data. *Comput. Struct. Biotechnol. J.* **20**, 2978–2985 (2022).
83. Gonzalez-Pena, V. et al. Accurate genomic variant detection in single cells with primary template-directed amplification. *Proc. Natl. Acad. Sci. USA* **118**, e2024176118 (2021).
84. Luquette, L. J. et al. Ultraspecific somatic SNV and indel detection in single neurons using primary template-directed amplification. *bioRxiv* <https://www.biorxiv.org/content/10.1101/2021.04.30.442032v1> (2021).
85. Miller, M. B. et al. Somatic genomic changes in single Alzheimer's disease neurons. *Nature* **604**, 714–722 (2022).
86. Sadler, T. W. *Langman's Medical Embryology* (Wolters Kluwer, 2018).
87. Barresi, M. J. F. & Gilbert, S. F. *Developmental Biology*. 12th. edn (Sinauer Associates, Inc., Publishers, 2019).
88. Kwon, G. S., Viotti, M. & Hadjantonakis, A. K. The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. *Dev. Cell* **15**, 509–520 (2008).
89. Stern, C. D. Aptly named visceral endoderm. *Dev. Cell* **15**, 493–494 (2008).
90. Martincorena, I. et al. Somatic mutant clones colonize the human esophagus with age. *Science* **362**, 911–917 (2018).

ACKNOWLEDGEMENTS

This work was supported by a Biomedical Research Institute grant, Kyungpook National University Hospital (2020) and by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI15C0001). We thank Areum Cho and Jung Min Park for their technical assistance.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Ji Won Oh.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022