

Methods in clonal analysis and applications

Anne-Cécile PETIT**, Emilie LEGUÉ**, Jean-François NICOLAS*

Unité de Biologie moléculaire du Développement, Institut Pasteur, 25 rue du Docteur Roux,
75724 Paris Cedex 15, France

Abstract – During development, embryonic cells display a large variety of behaviors that lead to the formation of embryonic structures that are frequently transient. Simultaneously, cells progress towards a specific fate. The current challenge for embryologists is to resolve how these two distinct aspects of development co-exist. As cell behaviors (including elementary cellular operations such as motility, adhesiveness, polarization, change in shape, division and death) and their control are much less well understood than the genetic aspects of cell fate determination, there is currently much interest in the study of cell behaviors. This mainly consists of labeling groups of cells or, less frequently, single cells and observing their descendants. In this review, we describe a few techniques for labeling groups of cells and we discuss prospective and retrospective clonal analysis, in particular the LaacZ system, in detail. We examine the information generated by these approaches.

cell lineage / *laacZ* / cell fate assignment / developmental patterning / change in shape / cell adhesiveness / cell motility

1. INTRODUCTION

During development, embryonic cells display a large variety of behaviors based on various combinations of elementary cellular operations: motility, adhesiveness, polarization, change in shape, division and death [1]. This leads to the formation of embryonic structures that are frequently transient. Simultaneously, cell fate is progressively determined. The current challenge for embryologists is to determine how these two distinct aspects of development co-exist. Cell behavior and assignment of

cell fate could be two separate aspects controlled independently (in parallel). However, progression towards a specific cell fate and selection of a particular behavior could depend on the same signal. Although simple, this co-control theory does not take into account the historical and contingent characteristics of evolution that constrain development. From an evolutionary viewpoint, the development of increasingly elaborate cell behaviors raises questions about the associated selective advantage [2]. It is possible that cell behaviors authorize novel cell-cell interactions, thus enhancing

* Corresponding author: jfnicola@pasteur.fr

** These authors contributed equally to this work.

developmental plasticity. This is particularly plausible in vertebrates, unlike in organisms with invariant lineages and poorly elaborated cell behaviors.

There is currently much interest in cell behaviors and their control as these aspects are less well understood than the genetic aspects of cell fate determination. The movements that occur during gastrulation, which have always fascinated embryologists [3], are being intensively studied in amphibians [4], chicks [5] and zebrafish [6], although very little is known about the molecular mechanisms involved. The description of cell behaviors should make it possible to determine how they are involved in the formation of embryonic structures and, more importantly, to understand how cell behaviors and genetic aspects of cell fate are coordinated during development.

A powerful way to study many aspects of cell behavior during embryogenesis would be to film development at the cell level by videomicroscopy (with the new powerful imaging techniques) and to reconstitute the life of the embryo in 4D. However, it is not yet possible to generate complete 4D reconstructions of a vertebrate embryo because of the large number of cells involved and because development is a long-term process. In addition, embryos are frequently mobile and it is difficult to interfere with mobility without interfering with development. Finally, mammals develop in utero and their in vitro culture is only possible for short periods of time and at certain stages. For the mouse, in vitro culture is limited to 48 h during early development (from E6 to E7) and even less at later stages. Therefore, although extremely informative [7], videomicroscopy is restricted to small regions of the embryo and to short-term analysis. As these approaches are, so far, limited by severe technological problems, the analysis of cell behaviors is mainly performed by labeling groups of cells or (less frequently) single cells and the observing their descendants [8]. The labeling of groups of cells allows global analysis of the

fate, the proliferation and the movements of tissues, but it does not give any information about individual behaviors, which is often necessary to define the precise parameters of development.

Clonal analysis is based on the study of clones derived from single cells. It gives very precious information about cell characteristics and makes it possible to generate lineages of embryonic cells. It can detect apoptosis [9] and provide information on the rate and mode of growth (either oriented or isotropic). Cell dispersal and cell movements can be defined. Furthermore, clonal analysis makes it easy to detect the regions [10] or compartments [11] of a structure and to study the involvement of stem cells in their formation [12].

In this review, we briefly describe a few techniques for labeling groups of cells and we discuss clonal analysis in detail, in particular the LaacZ system [13, 14], and the information provided by these novel approaches.

2. METHODS FOR ANALYZING OF CELL BEHAVIORS

2.1. The labeling of groups of cells

The labeling of groups of cells provides information about their fate, their proliferation and their movements during development. There are several different labeling methods [15]:

(1) Delivery of fluorescent lipophilic dyes (DiI, DiO) using iontophoresis or pressure.

(2) Grafting. In quail chick transplantation, quail cells are grafted into a chick embryo. The nuclei of quail cells can be distinguished from chick nuclei [16]. Grafts of cells expressing GFP or LacZ or previously labeled in vitro with a fluorescent dye are also used.

(3) Electroporation of DNA encoding a reporter gene.

(4) In the mouse, tetraparental embryos have been used [17]. This approach involves combining two 2 to 8 cell stage embryos from different lines to produce a chimera. Cells are identified due to marker polymorphism (e.g. cis-glucose-6-phosphate-isomerase GPI) or differences in the satellite DNA [18]. The expression of a reporter gene (LacZ, GFP, alkaline phosphatase) by one line is also possible. This technique is limited because chimeras can only be generated at preimplantation stages.

Dye labeling has been used to produce detailed fate maps at various stages of mouse, chick and zebrafish embryogenesis. Quail chick chimeras have been used to follow the fate of numerous embryonic territories such as the neural crest, the somites, various regions of the neural plate and central nervous system [19]. Tetraparental mice have shown that most of the structures of the organism have a polyclonal origin and therefore that cell fate determination concerns groups of cells rather than single cells [20].

As already mentioned, these approaches are global and do not easily and unambiguously reveal the behavior of individual cells. This is made possible by clonal analysis methods based on the labeling of single cells. There are two main clonal analysis approaches: prospective and retrospective.

2.2. Prospective clonal analysis

In prospective methods, the cell to be labeled is chosen by the investigator and therefore its position is known. Its fate is determined by analyzing its clonal descendants. The cell is labeled by intracellular injections of an enzyme (horseradish peroxidase), a dye (usually fluorescent e.g. dextran linked to rhodamine, fluorescein or biotin; fluorescent dextran conjugates) or DNAs encoding a reporter gene (GFP, LacZ). These labeling methods are cell autonomous because the high molecular weight dyes cannot diffuse from one cell to another [21].

Prospective clonal analysis has been used to establish fate maps of many embryos and

embryonic territories of invertebrates and vertebrates such as the mouse, amphibians, and zebrafish. The complete lineages of several invertebrates such as ascidians (*Ciona intestinalis*) [22–24], sea urchin (*Strongylocentrotus purpuratus*) [25, 26], nematodes (*Caenorhabditis elegans*) [27] and leeches (*Helobdella triserialis*) [28] have been established using prospective clonal analysis. These methods are extensively used to study defined populations of cells in vertebrates.

Labeling can also take the form of a genetic modification induced at a given time during development. Such labels are permanent and are transmitted to the descendants. The induction of mitotic clones based on homologous recombination is extensively used in *Drosophila*. Recombination can be induced by irradiation [29] or by a heat shock using the FLP/FRT system [30]. This technique has also been used in zebrafish [31].

Prospective clonal analysis has several limitations:

(1) The dilution of dyes following injection limits the analysis to four to six cell divisions.

(2) Unavoidable embryo manipulation can disturb development. This applies especially to mammals, as their embryos must be cultured ex utero for such experiments.

(3) This approach is not totally objective as the investigator chooses the cell to inject. Therefore, it is difficult to be sure that all cells at the origin of a structure have been targeted. To overcome this problem, it is more appropriate to use methods for the systematic screening of all possible labels. The retrospective methods satisfy this requirement in principle.

2.3. Retrospective clonal analysis

Retrospective clonal analysis consists of labeling a single cell at random and then observing the derived clone. Information concerning the fate of the labeled cell and

the behaviors of its descendants are retrospectively deduced. This method is based on genetic labeling techniques.

Recombinant retroviruses (RRV) can be used to introduce a reporter gene into embryonic cells [32, 33]. Following infection, a single DNA copy of the virus is integrated into the host genome as a provirus. As a consequence, the labeling is stably transmitted to descendants. RRV are defective, that is unable to form new viruses. Secondary infection is therefore impossible and the labeling is strictly cell autonomous. The RRV encodes a reporter gene such as LacZ, alkaline phosphatase, GFPs or luciferase. In principle, this method allows the labeling of any cell of the embryo if, as is usually the case, the RRV used does not exhibit tropism restriction. It is therefore possible to screen all possible sorts of labeling at a given embryonic stage.

This method has been, and still is, largely used in the chick and the mouse. It has been extraordinarily useful to study lineages in the CNS. However, this method has severe limitations: (1) the intra-uterine injection of the RRV in mice can damage the embryo; (2) the expression of the provirus depends on its site of integration in the host genome; (3) the provirus can be silenced by a variety of epigenetic repression mechanisms; (4) the accessibility and infectivity of all cells of the embryo are not uniform. For instance, murine RRVs do not infect post-mitotic or resting cells. This limitation has recently been overcome by using lentiviruses (HIV1-based) [34].

Cells can also be labeled using the LaacZ system, which is described in detail below.

2.4. Retrospective clonal analysis using the LaacZ labeling system

The LaacZ system is based on a particular reporter gene that was obtained by modifying the bacterial LacZ gene [13]. The duplication of 289 or 1000 bp in the coding sequence of the gene generates many in-frame stop codons. As a conse-

quence, the LaacZ gene encodes a truncated, non-functional β -galactosidase.

A functional LacZ gene can be restored subsequent to intragenic homologous recombination initiated by a double strand break in the duplicated region. Repair mechanisms, in particular double strand break repair (DSBR), result in a homologous recombination event that re-establishes the LacZ coding frame generating a functional β -galactosidase (Fig. 1A). The recombined LacZ is then transmitted to all descendants of the modified cell. The resulting clone is detectable by β -galactosidase histochemical staining. This spontaneous event is rare (1 per 10^6 cells per generation in culture).

The LaacZ gene is combined with a promoter that targets its expression in a chosen tissue or structure. This technique requires the generation of LaacZ transgenic lines. This limits its use to organisms for which transgenesis is possible. The establishment of these transgenic lines requires time and effort and the characteristics of the chosen animal must be taken into account. Several mouse lines with different promoters already exist. The α AChR-LaacZ line (in which the LaacZ reporter is under the control of the promoter of the α sub-unit of acetylcholine receptor) is useful to study the myotome [13, 35]; the α actin-cardiac LaacZ line is useful to study the cardiac and skeletal muscular systems [36]; the NSE LaacZ line (in which the reporter is under the control of the promoter of the neuron-specific enolase) is useful to study the central nervous system [37]; and the Dct-LaacZ line (in which the reporter is under the control of the promoter of the dopachrome tautomerase gene) is useful to study neural crest derivatives such as melanocytes [38].

In a LaacZ embryo, intragenic homologous recombination occurs at random, in any cell and at any time during development. This leads to random labeling of cells without any need to manipulate the embryo. As homologous recombination is an extremely rare event, it generally concerns only one

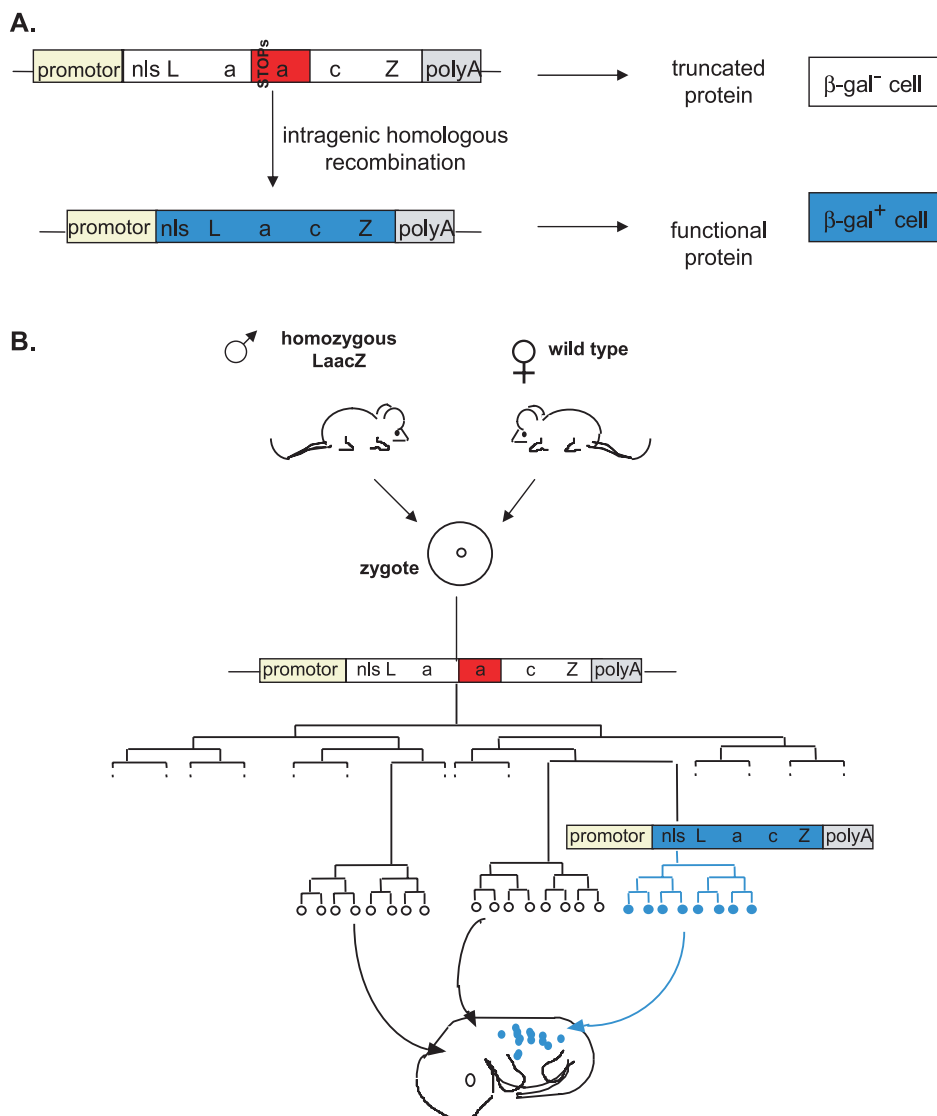


Figure 1. The LaacZ labeling system. **(A)** The LaacZ reporter gene. LaacZ was derived from the bacterial LacZ gene by duplicating a portion of its coding sequence. This generates stop codons leading to the translation of a truncated inactive β -galactosidase. The LacZ gene can be restored by intragenic homologous recombination between the duplicated sequences. This re-establishes the coding frame meaning that functional β -galactosidase can be produced. **(B)** Production of clones. LaacZ embryos are produced by crossing homozygous LaacZ transgenic males with wild-type females. In these embryos, a spontaneous homologous recombination event can occur randomly in any cell and at any time during development. This rare event initiates the labeling. The recombined cell carries a functional LacZ. β -galactosidase-positive ($\beta\text{-gal}^+$) clones can be visualized by X-gal histochemistry.

cell in the embryo (Fig. 1B). Labeled clones can be visualized in the structure or tissue in which the reporter gene is expressed. If a tissue-specific promoter is used, it is possible to visualize cells derived from very early labelings at a time when perhaps the promoter was inactive. Libraries (collections) of clones can be produced. When large enough, these libraries reflect the labeling of all possible precursors of a given structure.

This method presents several advantageous characteristics for the analysis of successive stages of clonal development:

- (1) As the LaacZ gene can recombine even when not expressed, all cells of the embryo can potentially be labeled.
- (2) The labeling does not necessitate any manipulation of the embryo.
- (3) The labeling is genetic, therefore permanent, allowing long-term analysis.

A different system of genetic cell labeling for retrospective clonal analysis has been engineered in *Antirrhinum majus* (snapdragon). It is based on the random deletion of a transposon that leads to a red pigmentation in a white background [39]. It has been used to study clonal behaviors in the petals [40].

3. TERMINOLOGY AND CONCEPTS OF RETROSPECTIVE CLONAL ANALYSIS

The LaacZ system makes it possible to initiate clones at any time during the formation of a structure. To facilitate the analysis of the libraries of clones obtained, it may be useful to bear in mind a few theoretical situations of the formation of prototypical structures [41].

3.1. Definitions

3.1.1. Progenitors and founder cells

The cells that form the structure called *A* (Fig. 2) are those present at the time of the

analysis (observation). These cells can correspond to several cell types if the structure is multitissular and already, at least partly, mature.

In the lineage of the structure *A*, the progenitor cells are those cells whose descendants all participate in the final structure and the founder cells are the earliest cells whose descendants all participate in the structure (Fig. 2). The progenitor cells derive from the division of founder cells. The descendants of a founder cell in structure *A* are defined as a cohort. When a founder cell or a progenitor cell is labeled, all its descendants are in structure *A* and are detected.

3.1.2. The ancestral cells

If only a fraction of its descendants contribute to structure *A*, a cell is an ancestral cell of structure *A* (Fig. 2). The other descendants will contribute to other structures of the embryo. For instance, before gastrulation, a cell of the presumptive mesoderm can be an ancestral cell of the cardiac muscle but can also contribute to other mesodermal derivatives such as the skeletal muscle system. If an ancestral cell is labeled, the cells detected in structure *A* do not represent all the clonal descendants of this cell.

The period that precedes the founder cells of a structure (from the ancestral cells to the founder cells) is called the pre-historic period of the structure because its lineage is more or less indeterminate during this period. The period from the determination of founder cells is called the historic period of the structure because its lineage is determinate.

3.1.3. Horizontal and vertical clonal growth

The division of an ancestral cell gives rise to a novel ancestral cell for *A* and to another cell whose descendants will participate in other structures (Fig. 2A). This type of growth is called “non-clonal” as only one of the two daughter cells will have descendants

in *A*. It is also known as “vertical growth” as the number of cells contributing to structure *A* remains constant during subsequent generations.

However the pool of ancestral cells can increase in size if an ancestral cell of structure *A* gives rise to two daughter cells both with ancestral characteristics (Fig. 2B). The growth is known as “clonal” as both daughter cells contribute to structure *A*. It is also called “horizontal” growth as the number of cells contributing to structure *A* increases.

From the founder cell stage onwards, growth is always clonal and horizontal except if the structure exhibits apoptosis. If apoptosis occurs, the number of progenitor cells of structure *A* can remain constant from one cell generation to the next.

3.1.4. Monoclonal or polyclonal origin of a structure

A structure has a monoclonal origin when it is formed from the descendants of only one founder cell. Its origin is polyclonal when it is formed from the descendants of several founder cells. When the structure has a monoclonal origin, only two periods characterize the formation of the structure: a non-clonal period of vertical growth for the ancestral cells and then a clonal period of horizontal growth for the founder and progenitor cells (Fig. 2A). This situation is nevertheless exceptional. In vertebrate embryos, it has been repeatedly found that most organs, including very tiny ones such as the hair follicle in mammals, have a polyclonal origin.

When the structure has a polyclonal origin, several cases are possible. Two examples are given in Figures 2B and 2C. In the first case (Fig. 2B), an intermediate pool of ancestral cells is formed at a certain phase during development with a period of clonal and horizontal growth, before the determination of the founder cells. Therefore, there are successive periods of clonal and non-clonal growth. In the second case (Fig. 2C), structure *A* is formed by several independent lineages. This can result from the

recruitment of novel cells at a certain time during development. Growth can follow either scheme in each lineage.

3.1.5. Families of clones and saturation of the libraries

The random labeling of cells with the LaacZ system generates different types of clones that can be grouped into families on the basis of common characteristics: size, spatial distribution, cell type, etc.

The labeling reaches saturation when the library of clones obtained has more than one example of any possible labeling. At this stage, additional clones do not provide any new information.

3.2. Concepts

3.2.1. Modes of growth

Several modes of growth are possible: a proliferative mode, when each division produces two equivalent daughter cells; a stem cell mode, when the division produces a cell similar to the parental one and another cell with a different fate; and a diversification mode, when the division produces two different daughter cells that are also different from the parental cell. Clonal analysis makes it possible to distinguish between these three modes.

As the number of cells doubles at each generation in the proliferative mode, the probability of reaching a cell of a given generation also doubles with each generation. Therefore the probability of obtaining a clone of a given size is inversely proportional to this size. The most frequent clones have only one cell and the rarest are the largest (Fig. 3A).

In the stem cell mode of growth, each cell divides asymmetrically giving one cell that self renews the pool of stem cells and another cell with a different fate. As in the proliferative mode, the most frequent clones are those with only one cell. Whatever the size of the clone (> 1 cell), the probability of labeling remains constant. If the clones are

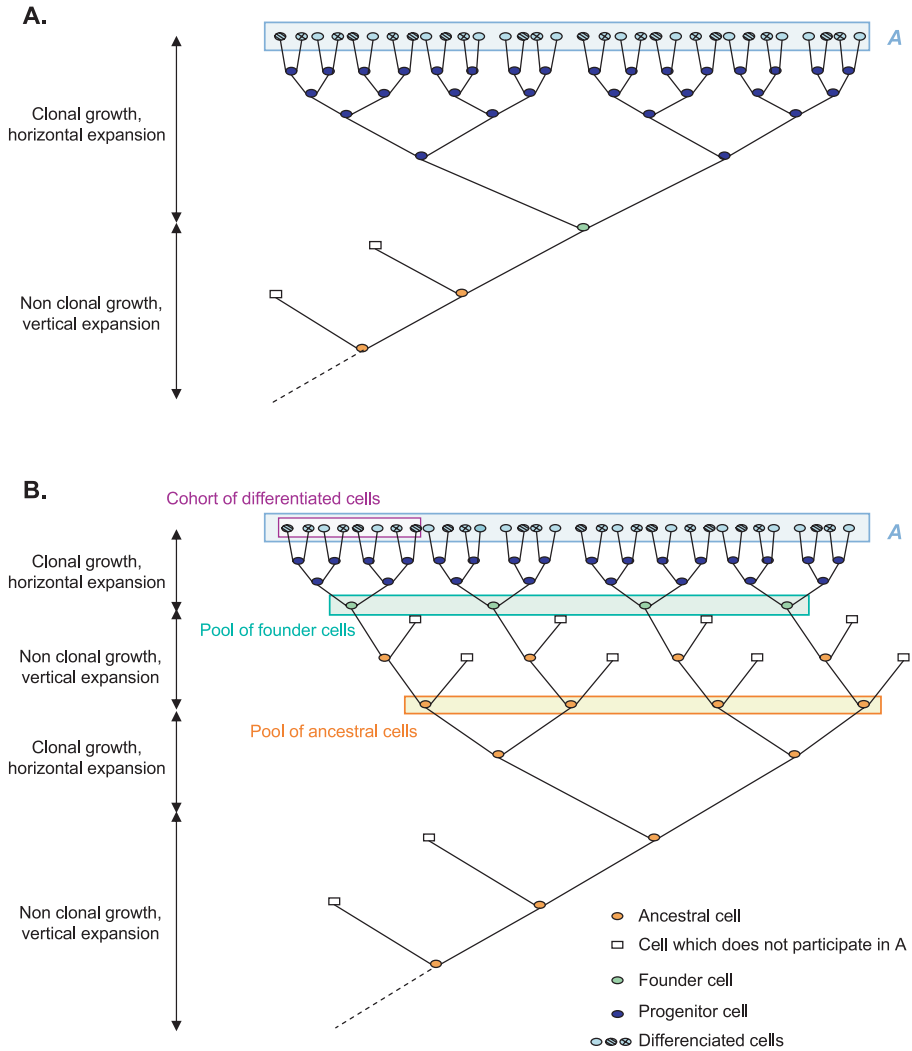


Figure 2. Genealogical trees for simple model structures. **(A)** Monoclonal origin. The various differentiated cells (pale blue ovals) that constitute *A* (the blue rectangle) are derived from a single founder cell (green oval). Two successive periods of growth occur during the formation of *A*: a period of non-clonal growth during the vertical propagation of the ancestral cells (orange ovals); a period of clonal growth during the horizontal expansion of the pool of progenitor cells (dark blue ovals) from the single founder cell (green oval). The progenitor cells produce the differentiated cells that constitute *A*. **(B)** Polyclonal origin. The various differentiated cells (pale blue ovals) that constitute *A* (the blue rectangle) are derived from a pool of several founder cells (green ovals and the green rectangle). A period of clonal growth leads to the expansion of the pool of ancestral cells (orange ovals and the orange rectangle). A cohort of differentiated cells is represented in purple. **(C)** Dual clonal origin. The cells that constitute *A* (the blue rectangle) are derived from two distinct lineages. The first lineage has a monoclonal origin (one founder cell, in green, as in **2.A**). The second lineage has a polyclonal origin (several founder cells, green ovals and green rectangle, as in **2.B**) and an intermediate pool of ancestral cells (orange ovals and orange rectangle). The cells from the two lineages intermingle in the final structure. For a colour version of this figure see www.edpsciences.org/rnd.

C.

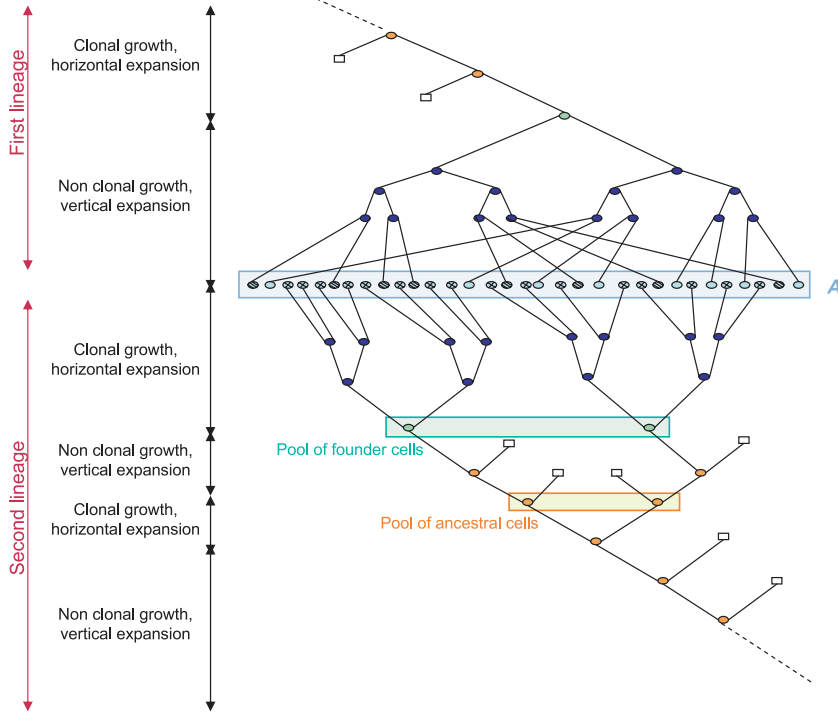


Figure 2. Continued.

grouped in classes following an exponential progression (to make comparison with the proliferative mode of growth possible), it becomes clear that the frequency of production of labeled clones of the different classes increases exponentially with their size (Fig. 3B). This is the opposite of the proliferative mode of growth (compare Figs. 3B and 3A).

Obviously the formation of a structure may involve both of these modes. For instance, a period of proliferation can follow a period of stem cell growth. In such cases, the graph of the frequencies of the clones in relation with their sizes will exhibit an exponential decrease characteristic and indicative of a period of exponential growth and then an exponential increase

indicative of a period of stem cell growth (Fig. 3C). Careful examination of the size of the clones in a library can therefore reveal periods of exponential (proliferative) and linear (stem cell) growth.

3.2.2. Crossing index, compartments and regionalized structures

Embryonic structures are sometimes regionalized during their formation. After regionalization, clones are no longer randomly distributed in the structure: they are preferentially located in a given region. This contrasts with the clones initiated before regionalization, which can contribute to several regions of the structure. To detect a process of regionalization with a

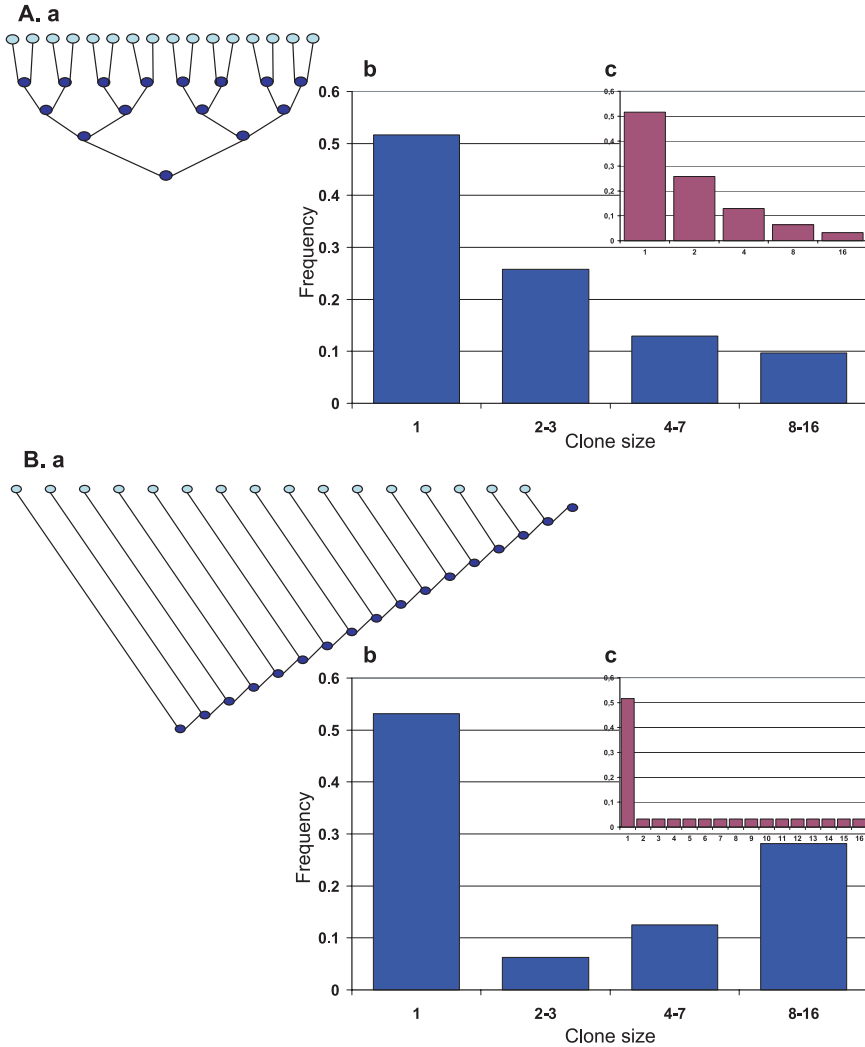


Figure 3. (A) Proliferative mode of growth. Each cell (blue ovals) divides into two equivalent daughter cells (Aa). The frequency of clones (y axis) decreases exponentially according to their size (x axis) (Ab and Ac). The graphs are the same whether we consider clone size (Ac) or groups of clones in classes of increasing exponential size (Ab). (B) Stem cell mode of growth. Each stem cell (dark blue ovals) divides unequally to give rise to a new stem cell and to another cell with a different fate (light blue ovals) (Ba). The most frequent clones have one cell and the frequency of clones is the same when the clones are composed of more than one cell (Bc). When clones are grouped in classes of increasing exponential size, the distribution of clone size increases exponentially (Bb). (C) Combined mode: stem cell mode followed by proliferative mode. In this case, the division of a stem cell produces a new stem cell and another cell that shows a proliferative mode of growth. This transient amplifying cell (blue rectangle) produces a cohort of four cells in the final structure (purple rectangle). When clones are grouped in classes of increasing exponential size, the distribution of clone size first decreases exponentially (typical of a proliferative mode of growth) and then increases exponentially (typical of a stem cell mode of growth) (Cb). For a colour version of this figure see www.edpsciences.org/rnd.

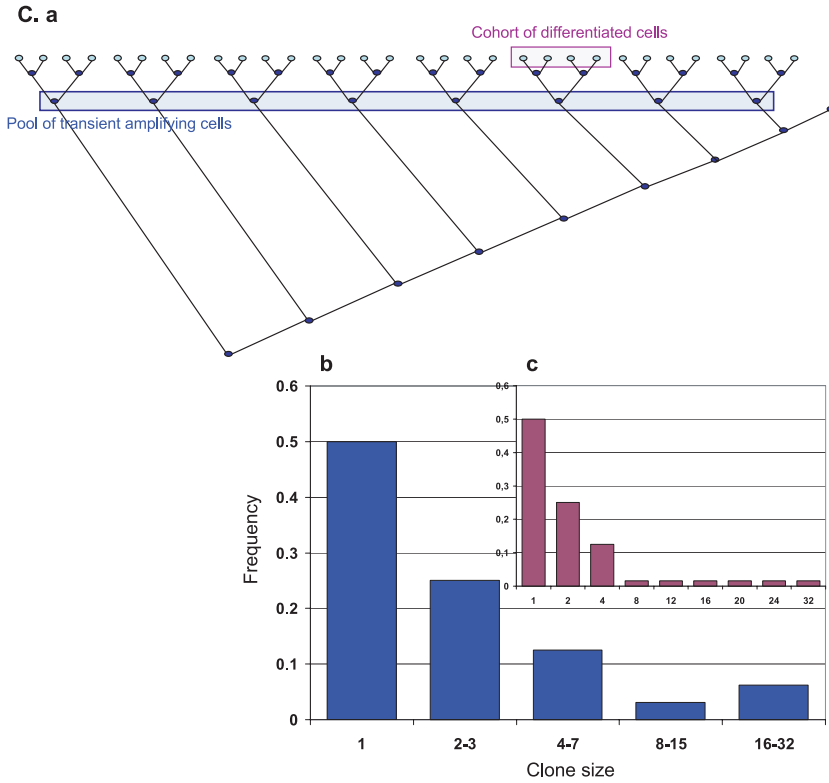


Figure 3. Continued.

random labeling technique such as the LaacZ method, the clone library must be saturated. It is next best to use a function called “crossing index” to detect regionalization (Fig. 4). Practically, the structure is divided into units either arbitrarily or on the basis of morphological marks. For each limit between the units, the value of the crossing index is equal to the number of clones that crosses it. In non-regionalized structures, the value of the crossing index remains constant (Fig. 4A). In regionalized structures, the value of the crossing index presents one or several minima at the level of the limits between the regions (Fig. 4B). To detect regions, the crossing index can be calculated for each family of clones in the library. If for a family the minimum of the crossing index is equal to 0, it means that

the cells of the clones do not cross the limits between regions (Fig. 4B, c and d). The regions may thus correspond to developmental compartments, although regions are frequently not compartments. The study of a family of clones with minimum crossing indexes of 0 makes it possible to date the appearance of regions or compartments.

3.2.3. Clonal complexity and regions of preferential proliferation

Using saturated libraries, it is also possible to detect regions of preferential proliferation. The use of clonal complexity values helps in this search (Fig. 5). The clonal complexity of a region is equal to the number of clones that contribute to it. For instance, in Figure 5B, region 3 is labeled

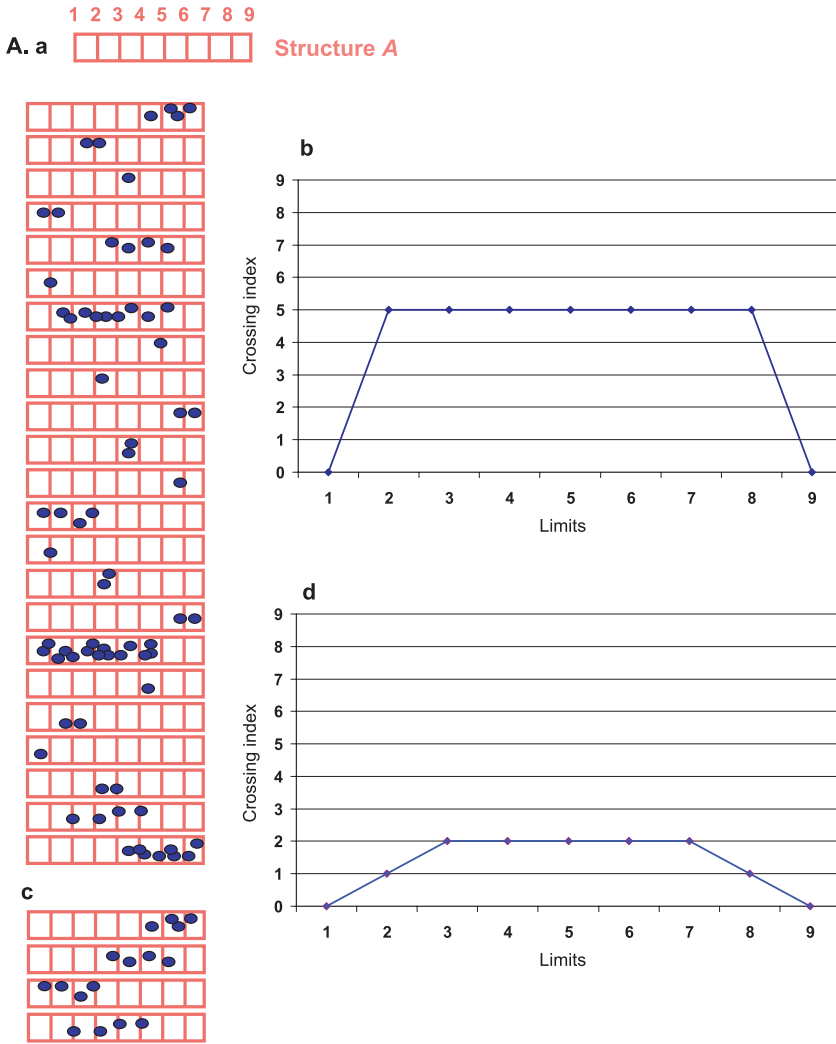


Figure 4. Crossing index. A structure is arbitrarily divided into eight equal units. The limits between the units are numbered 1 to 9 (the red rectangle). Considering the limit between two units, the value of the crossing index is equal to the number of clones that crosses it. In structure *A* and structure *B*, a collection of 23 clones is represented in blue, with one clone per line (**Aa** and **Ba**). The class of clones of four cells is represented separately (**Ac** and **Bc**). The graphs (**Ab** and **Ad**, **Bb** and **Bd**) represent the values of the crossing index (y axis) of the limits (x axis). (**A**) Crossing index of a non-regionalized structure. Graph **Ab** corresponds to the 23 clones represented in **Aa**. The crossing index is constant and equal to 5. When considering the four-cells clones (**Ac**), the crossing index remains constant and equals 2 (graph **Ad**). (**B**) Crossing index of a regionalized structure. Graph **Bb** corresponds to the 23 clones represented in **Ba**. Graph **Bd** corresponds to the four-cells clones represented in **Bc**. In graph **Bb**, the crossing index has two minima, for limits 3 and 6, demonstrating that the structure is regionalized. In graph **Bd**, when considering only the four-cells clones, the crossing index equals 0 for limits 3 and 6, suggesting that the regions become compartments when the four-cells clones are formed and, therefore, limits 3 and 6 are frontiers. For a colour version of this figure see www.edpsciences.org/rnd.

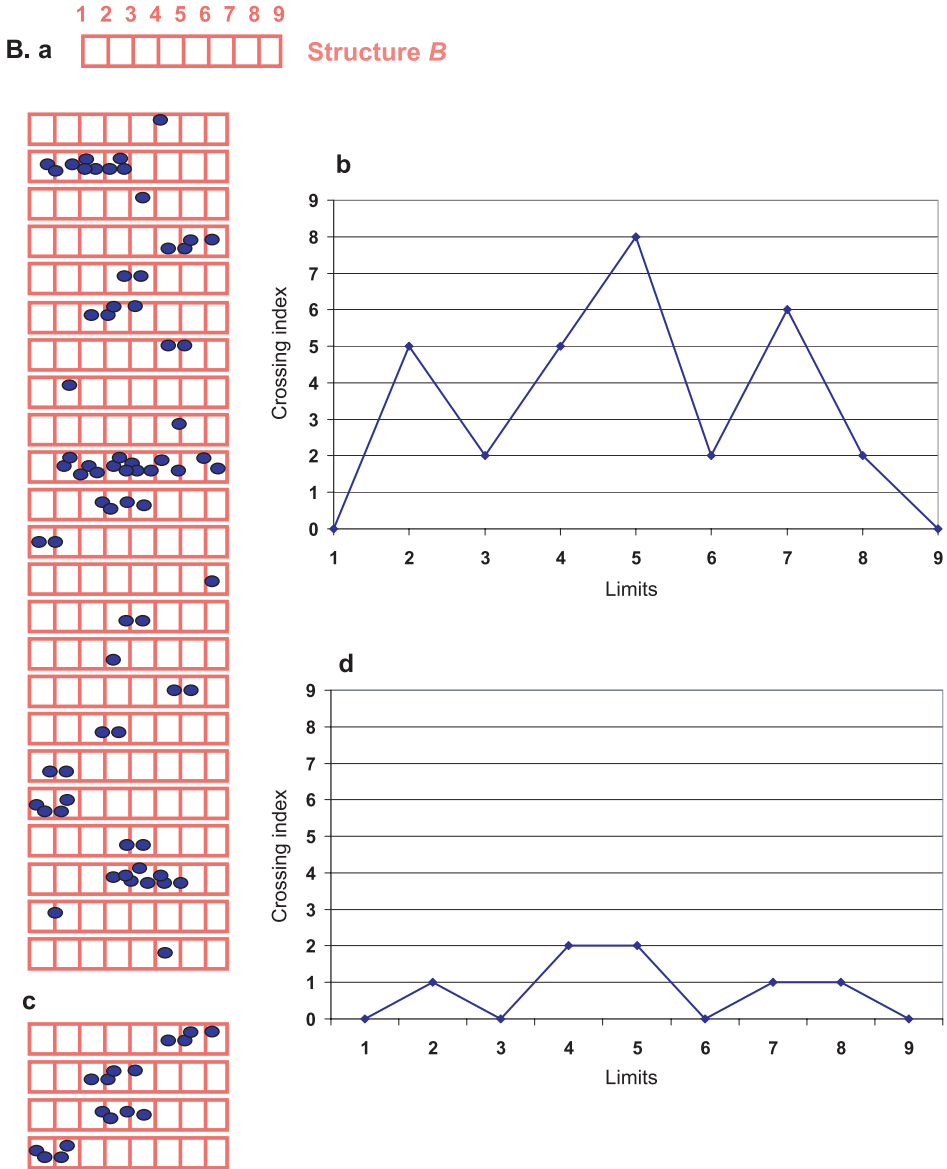


Figure 4. Continued.

15 times (15 clones): some of these clones contribute only to this region but others also contribute to region 2 or to regions 1 and 2. In contrast, region 1 is labeled only three

times, but the clones extend into regions 2 and 3. The clonal complexity of region 3 is higher than that of region 1. This reveals that the zone of preferential proliferation

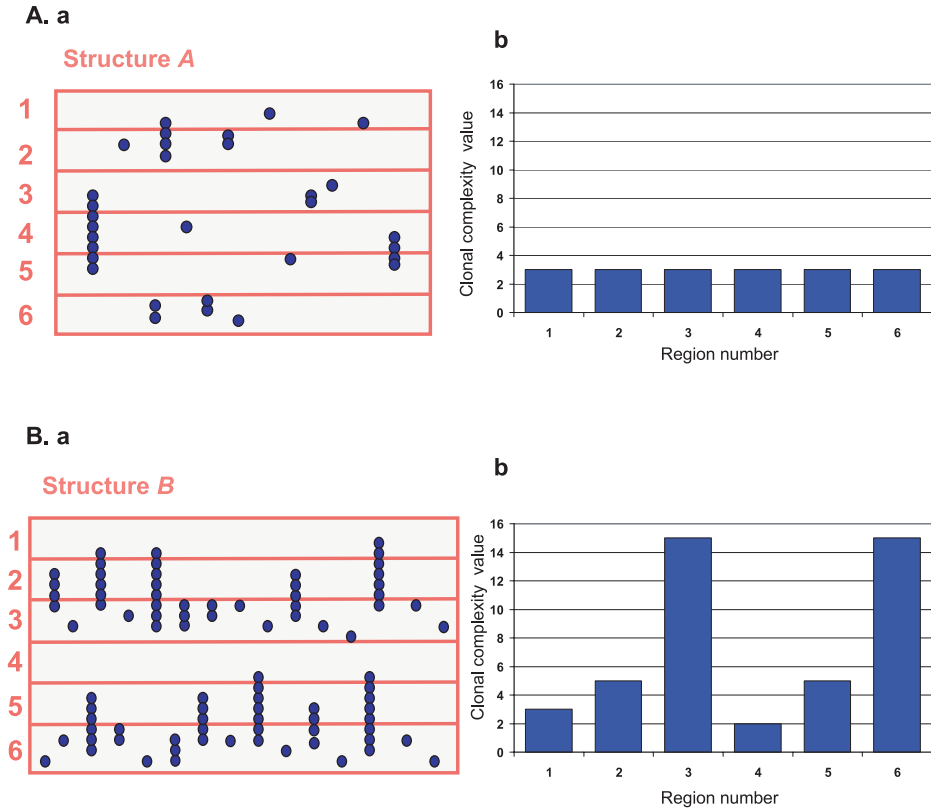


Figure 5. Clonal complexity. Structure *A* and structure *B* (in red) are divided into six regions, numbered 1 to 6. Fourteen clones (in blue) are represented along vertical lines in structure *A* (**Aa**) and 30 clones in structure *B* (**Ba**). Clonal complexity is the number of clones that participate to a region. The graphs show the clonal complexity (y axis) for each region (x axis). (**A**) A structure with regions of equivalent growth potential. Clonal complexity is constant (graph **Ab**). (Note that structure *A* follows a proliferative mode of growth.). (**B**) A structure with two regions of preferential growth. The graph (in **Bb**) shows two maxima for regions 3 and 6, demonstrating the presence of two regions of preferential growth. (Note that structure *B* follows a stem cell mode of growth.) For a colour version of this figure see www.edpsciences.org/rnd.

corresponds to region 3. The preferential proliferation may be due to stem cells.

If the cells of all regions of a structure have equivalent growth potential, the clonal complexity is equivalent for all regions (Fig. 5A).

The measurement of the clonal complexity identifies potential zones of preferential proliferation.

3.3. Information that can be deduced from the description of the clones

In general, different families of clones can be identified in a library. It is again important that the library be saturated to have a fair representation of all possible labeling. As the clones are generated at random and at any time during embryogenesis, the most recent were generated in the structure

itself and the others were generated at the level of the progenitors, founders or ancestral precursors. However the cellular generation in which the labeling was initiated provides no clear answers about the size and the topography of the clones.

3.3.1. Establishment of the lineage of the structures

When observing clones, the most direct information possibly relates to the type of differentiated cell from which they are composed. When a structure is multitissular, if a clone contributes to only one tissue, it means that the labeled cell was unipotent. However, if the clone contributes to several tissues, then the labeled cell was pluripotent. This information can be used to build a tree of the genealogical relations between the different cell types in the structure. For instance, the cerebellum is formed of several cell types: the deep nuclei neurons, the Purkinje cells, the granular neurons, the Golgi cells, the stellate and the basket neurons. Using a large library of clones obtained using NSE LaacZ transgenic line, it has been demonstrated that all cell types except the granular neurons are produced sequentially (first the neurons of the deep nuclei, then the Purkinje cells, then the Golgi cells and finally the stellate and basket neurons) from a unique common precursor [37]. The granular neurons have a distinct clonal origin. This indicates that the cerebellum has a dual clonal origin.

3.3.2. Families of clones and regionalization

To date the birth of a clone, it is necessary to compare it with the diverse families of clones in the library. It is interesting to attempt to identify the families corresponding to the ancestral, founder and progenitor cells. However, this is not a trivial problem because several families of clones can correspond to one category of precursor cells if these precursors are not equivalent and/or can adopt different behaviors. Further-

more, the comparison of the different families of clones may reveal that one family (e.g. A) has the properties of potential sub clones of another family (e.g. B). This may indicate that the B family clones were generated later than the A family clones. This sort of temporal classification is also very valuable for other reasons.

The probability of labeling the cells at the origin of a given family of clones is directly proportional to the size of its pool (total number of cells at the origin of this family). This number depends on the number of cells present at a given cellular generation in the embryo but also on the number of generations during which the cells at the origin of the family have existed. For instance, the ancestral cells of a structure are constituted of a relatively small number of cells at a given cellular generation but may exist for many generations. Therefore, the pool can be very large. In contrast, the founder cells of a structure are only present during one cellular generation and, therefore, they are always more rarely labeled than ancestral cells.

Finally, by quantifying the contribution of an ancestral or founder cell to a structure, it is possible to estimate the number of founder cells in this structure. This approach showed that the pools of founder cells at the origin of the myotome [12], the Purkinje cell layer of the cerebellum [37] and cardiac tissue [36] contain 100–150, 80 and 140 cells respectively.

3.3.3. Topological properties of the clones

The size and the spatial distribution of the cells of a clone can give information about many aspects of the characteristics of a structure such as whether it is or not regionalized. To extract the important features of the topology of the clones, it is necessary to find a method to describe it. It is often practical to divide the structure into arbitrary units, in relation or not with morphological criteria (such as segments in the

myotome or in the rhombencephalon), with the axes of the structure (dorso-ventral, antero-posterior, medio-lateral, proximo-distal) or with morphological boundaries. The clonal contribution to these units can then be described. Some clones contribute to several units whereas others are restricted to only one unit. It is possible that this analysis reassesses these arbitrary units, in other words that this analysis reveals limits between real regions that cannot be detected using morphological markers. The validity of the regionalization can be tested using the crossing index method described in Section 3.2.2.

It is then possible to describe the geometry of the clones within each defined region. This description includes the orientation of the clone and quantitative data such as its position and the value of the angle in relation with the axes. For instance, in the myotome, which has been studied with the α AchRnls LaacZ line, the clones were first described in relation with the somites and with their most anterior position along the longitudinal axis [12] and then within each somite [35].

The clonal complexity of each region may give information about the presence of regions of preferential proliferation (see Sect. 3.2.3.). This approach ultimately leads to the delimitation of regions and boundaries within the structure in relation with the distribution of the clones. The contribution of the clones to these different regions and their properties can next be studied.

3.3.4. The topology of the clones, cell movements and oriented growth

The spatial characteristics of clones can provide information about the dispersal of the precursor cells of a structure. Indeed, a clone can correspond to one, several or many clusters of cells. A clone formed of several clusters is suggestive of a phase of cell dispersal during development. The dispersion can be oriented or isotropic

(Figs. 6C and 6D). In contrast, a clone formed of a unique coherent group indicates that the growth has (probably) always been coherent. In this case, growth can be either oriented or isotropic (Figs. 6A and 6B). A period of coherent growth can follow a period of dispersion. The final arrangement of cells in a clone results from this sequence of behaviors (Fig. 6). It is important to note that if a period of dispersion follows a period of coherent growth, the period of coherent growth cannot be detected by looking at the final aspect of the clone. Similarly, a period of coherent oriented growth will mask a previous period of coherent isotropic growth.

3.3.5. The size of the clones and the modes of growth

In the case of regional and proliferative growth from strictly equivalent cells in the absence of apoptosis, the size of a clone gives a reliable indication of its date of birth. However, when one of these conditions is not fulfilled, the size of the clones and their date of birth are not directly correlated. For instance, if the cells have different generation times, the labeling of a cell with a poor division potential at the beginning of the lineage will produce a small clone (possibly dispersed in different regions of the structure) and the labeling of a cell with a high division potential closer to the end of the lineage will produce a large clone (possibly restricted in only one region of the structure). Similarly, an early clone may be smaller than a late clone if a fraction of its descendants undergo apoptosis.

Finally, it is interesting to consider the situation of a structure produced following a temporal mode from a pool of stem cells such as the myotome. In such cases, a small clone that contributes to the first segments produced may have been initiated very early in the embryo and a large clone contributing to several caudal segments may have been initiated later during the development of the structure.

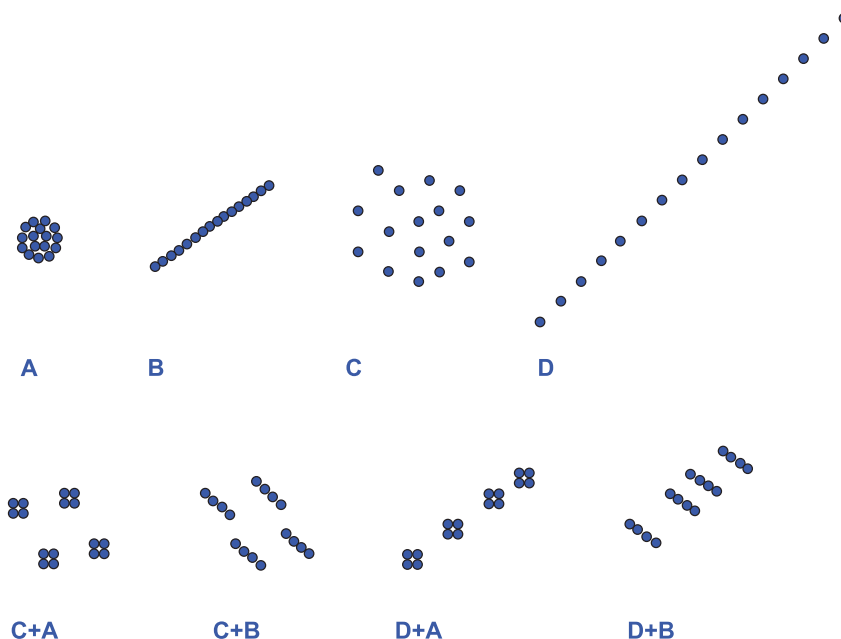


Figure 6. Clonal modes. The growth of a structure generally combines several elementary growth behaviors, called clonal modes. First line: Elementary clonal modes. (A) Isotropic-coherent growth. (B) Oriented-coherent growth. (C) Isotropic-dispersive growth. (D) Oriented-dispersive growth. Second line: a few examples of the growth when clonal modes are combined.

The size of clones can also give very valuable information concerning the mode of growth of the structure (proliferative or stem cell mode). Clones must first be classified by size. The study of the probability of labeling of these different clones then makes it possible to distinguish between these different modes and their sequences and therefore to determine whether stem cells are involved in the formation of the structure (see Sect. 3.2.1.) [10].

4. CONCLUSION

Clonal analysis is one of the most powerful ways to elucidate the history of a structure. Retrospective clonal analysis using the LaacZ method makes it possible to detect with certainty the successive steps of the formation of a structure because any ancestral or precursor cell can be marked.

The lineage, as well as certain aspects of cell behavior, can be deduced by analyzing libraries of clones. This method also provides more complex information such as the level of regionalization of a structure, the presence of developmental compartments or the mode of growth (proliferative, stem cell-based or diversificative). This kind of approach can be performed for any structure if an appropriate promoter can target the expression of LaacZ in it. Another very interesting problem is the relationship between the different embryonic structures. It is fair to say that too little is known concerning the lineages and the cell behaviors involved. Therefore, it would be interesting to control LaacZ by a promoter conferring ubiquitous expression such as the promoter of Rosa26 locus. This would make it possible to detect clones whatever the complexity of their tissue composition and at any

time. This work, which is now in progress, will open many novel possibilities.

One difficulty with the LaacZ method is that the date of birth of the clones is not known. The establishment of a system allowing the temporal induction of clones would therefore complement the LaacZ system. Unfortunately, all attempts have been unsuccessful because the existing induction systems such as the Cre/Lox system do not fulfill all necessary requirements. Firstly, the background activation of clones must be extremely low. Secondly, the induction of only one clone per embryo must be possible. Thirdly, induction must be possible at any stage during development. In the future, it may be possible to derive a temporal induction system from the LaacZ method. This would provide the opportunity to study the cellular history of the relations between embryonic structures.

ACKNOWLEDGMENTS

We thank several members of "Cells into organs" for their comments on the manuscript. This work was supported by grants from Cells into organs (European Network of excellence), Eurostem (European consortium for stem cell research), GPH Stem cells (Institut Pasteur), ARC (Association pour la Recherche sur le Cancer) and AFM (Association Française contre les Myopathies). J.F. Nicolas is from the Institut National de la Recherche Médicale (INSERM).

REFERENCES

- [1] Gilbert SF, Raunio AM (Eds). *Embryology: constructing the organism*. Sunderland, USA, Sinauer Ass, 1997.
- [2] Raff A. *The shape of life: genes, development and the evolution of animal form*. Chicago, University of Chicago Press, 1996.
- [3] Stern CD. *Gastrulation: from cells to embryo*. Cold Spring Harbor Laboratory Press, 2004.
- [4] Keller R. Shaping the vertebrate body plan by polarized embryonic cell movements. *Science* 2002, 298: 1950–1954.
- [5] Psychoyos D, Stern CD. Fates and migratory routes of primitive streak cells in the chick embryo. *Development* 1996, 122: 1523–1534.
- [6] Myers DC, Sepich DS, Solnica-Krrezel L. Convergence and extension in vertebrate gastrulae: cell movements according to or in search of identity? *Trends Genet* 2002, 18: 447–455.
- [7] Glickman NS, Kimmel CB, Jones MA, Adams RJ. Shaping the zebrafish notochord. *Development* 2003, 130: 873–887.
- [8] Gardner RL, Lawrence PA. Single cell marking and cell lineage in animal development. *Phil Trans R Soc Lon B* 1985, 312: 1–187.
- [9] Simpson P. Parameters of cell competition in the compartments of the wing disc of *Drosophila*. *Dev Biol* 1979, 69: 182–193.
- [10] Mathis L, Nicolas JF. Different clonal dispersion in the rostral and caudal mouse central nervous system. *Development* 2000, 127: 1277–1290.
- [11] Mathis L, Sieur J, Voiculescu O, Charnay P, Nicolas JF. Successive patterns of clonal cell dispersion in relation to neuromeric subdivision in the mouse neuroepithelium. *Development* 1999, 126: 4095–4106.
- [12] Nicolas JF, Mathis L, Bonnerot C. Evidence in the mouse for self-renewing stem cells in the formation of a segmented longitudinal structure, the myotome. *Development* 1996, 122: 2933–2946.
- [13] Bonnerot C, Nicolas JF. Clonal analysis in the intact mouse embryo by intragenic homologous recombination. *C R Acad Sci USA* 1993, 316: 1207–1217.
- [14] Mathis L, Nicolas JF. Cellular patterning of the vertebrate embryo. *Trends Genet* 2002, 18: 627–635.
- [15] Clarke JD, Tickle C. Fate maps old and new. *Nat Cell Biol* 1999, 1: E103–E109.
- [16] LeDouarin NM. Particularités du noyau interfase chez la caille japonaise (*Coturnix coturnix japonica*). Utilisation de ces particularités comme marqueur biologique dans les recherches sur les interactions tissulaires et les migrations cellulaires au cours de l'ontogenèse. *Bull Biol Fr Belg* 1969, 103: 435–452.
- [17] Mintz B. Genetic mosaicism in adult mice of quadruparental lineage. *Science* 1965, 148: 1232–1233.
- [18] Pedersen RA. Potency, lineage and allocation in preimplantation mouse embryos. In: Pedersen RA, *Experimental approaches to mammalian embryonic development* New York, Press Syndicate of the University of Cambridge, 1986, p 3–33.

- [19] LeDouarin N. The neural crest. Cambridge, Cambridge University Press, 1999.
- [20] Mintz B. Gene control of mammalian differentiation. *Annu Rev Genet* 1974, 8: 411–470.
- [21] Beddington R, Lawson KA. Clonal analysis of cell lineages. In: Cockcroft DL (Ed), *Postimplantation Mammalian Embryos – A practical Approach*, Oxford, Oxford University Press, 1990, p 267–291.
- [22] Nishida H, Satoh N. Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. I. Up to the eight-cell stage. *Dev Biol* 1983, 99: 382–394.
- [23] Nishida H, Satoh N. Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. II. The 16- and 32-cell stages. *Dev Biol* 1985, 110: 440–453.
- [24] Nishida H. Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. III. Up to the tissue restricted stage. *Dev Biol* 1987, 121: 440–453.
- [25] Hörstadius S. The mechanisms of sea urchin development studied by operative methods. *Biol Rev Cambridge Phil Soc* 1939, 14: 132–179.
- [26] Cameron RA, Davidson EH. Cell type specification during sea urchin development. *Trends Genet* 1991, 7: 212–218.
- [27] Sulston JE, Schierenberg E, White JG, Thomson JN, Von Ehrenstein G. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev Biol* 1983, 100: 64–119.
- [28] Weisblat DA, Shankland M. Cell lineage and segmentation in the leech. *Phil Trans R Soc Lond B* 1985, 312: 39–56.
- [29] Bryant PJ. Cell lineage relationship in the imaginal wing disc of *Drosophila melanogaster*. *Dev Biol* 1970, 22: 389–411.
- [30] Theodosiou NA, Xu T. Use of FLP/FRT system to study *Drosophila* development. *Methods* 1998, 14: 355–365.
- [31] Streisinger G, Coale F, Taggart C, Walker C, Grunwald DJ. Clonal origins of cells in the pigmented retina of the zebrafish eye. *Dev Biol* 1989, 131: 60–69.
- [32] Cepko CL, Ryder E, Austin C, Golden J, Fields-Berry S, Lin J. Lineage analysis using retroviral vectors. *Methods* 1998, 14: 393–406.
- [33] Sanes J, Rubinstein J, Nicolas JF. Use of a recombinant retrovirus to study post-implantation cell lineage in mouse embryos. *EMBO J* 1986, 5: 133–142.
- [34] Naldini L, Blomer U, Gallay P, Ory D, Mulligan R, Gage FH, Verma IM, Trono D. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector (see comments). *Science* 1996, 272: 263–267.
- [35] Eloy-Trinquet S, Nicolas JF. Clonal separation and regionalisation during formation of the medial and lateral myotomes in the mouse embryo. *Development* 2002, 129: 111–122.
- [36] Meilhac S, Esner M, Kelly RG, Nicolas JF, Buckingham M. The clonal origin of myocardial cells in different regions of the embryonic mouse heart. *Dev Cell* 2004, 6: 685–698.
- [37] Mathis L, Bonnerot C, Puelles L, Nicolas JF. Retrospective clonal analysis of the cerebellum using genetic lacZ/lacZ mouse mosaics. *Development* 1997, 124: 4089–4104.
- [38] Wilkie AL, Jordan SA, Jackson IJ. Neural crest progenitors of the melanocyte lineage: coat colour patterns revisited. *Development* 2002, 129: 3349–3357.
- [39] Coen ES, Carpenter R, Martin C. Transposable elements generate novel spatial patterns of gene expression in *Antirrhinum majus*. *Cell* 1986, 47: 285–296.
- [40] Rolland-Lagan AG, Bangham AJ, Coen E. Growth dynamics underlying petal shape and asymmetry. *Nature* 2003, 422: 161–163.
- [41] Eloy-Trinquet S, Mathis L, Nicolas JF. Retrospective tracing of the developmental lineage of the mouse myotome. *Curr Top Dev Biol* 2000, 47: 33–80.