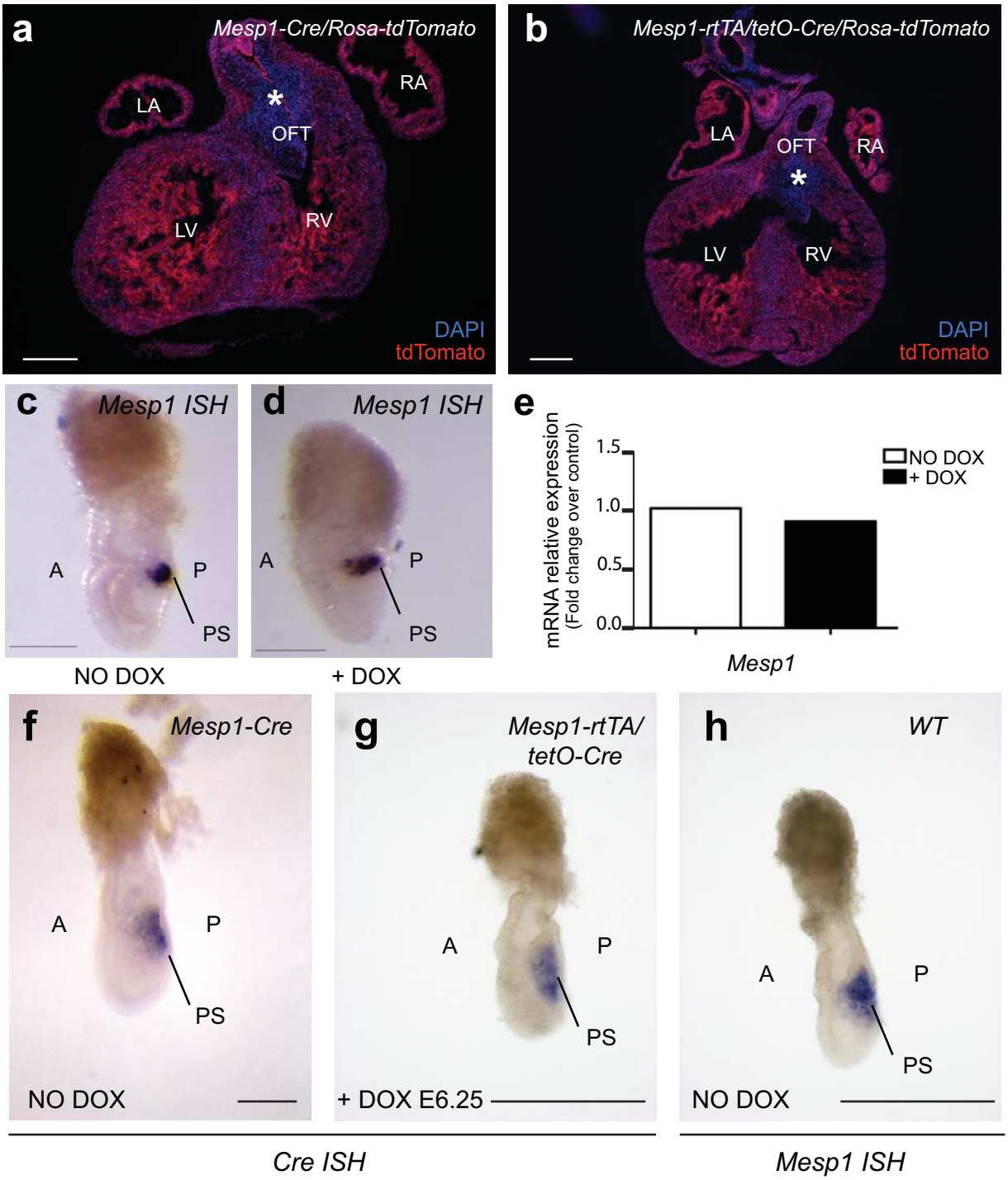
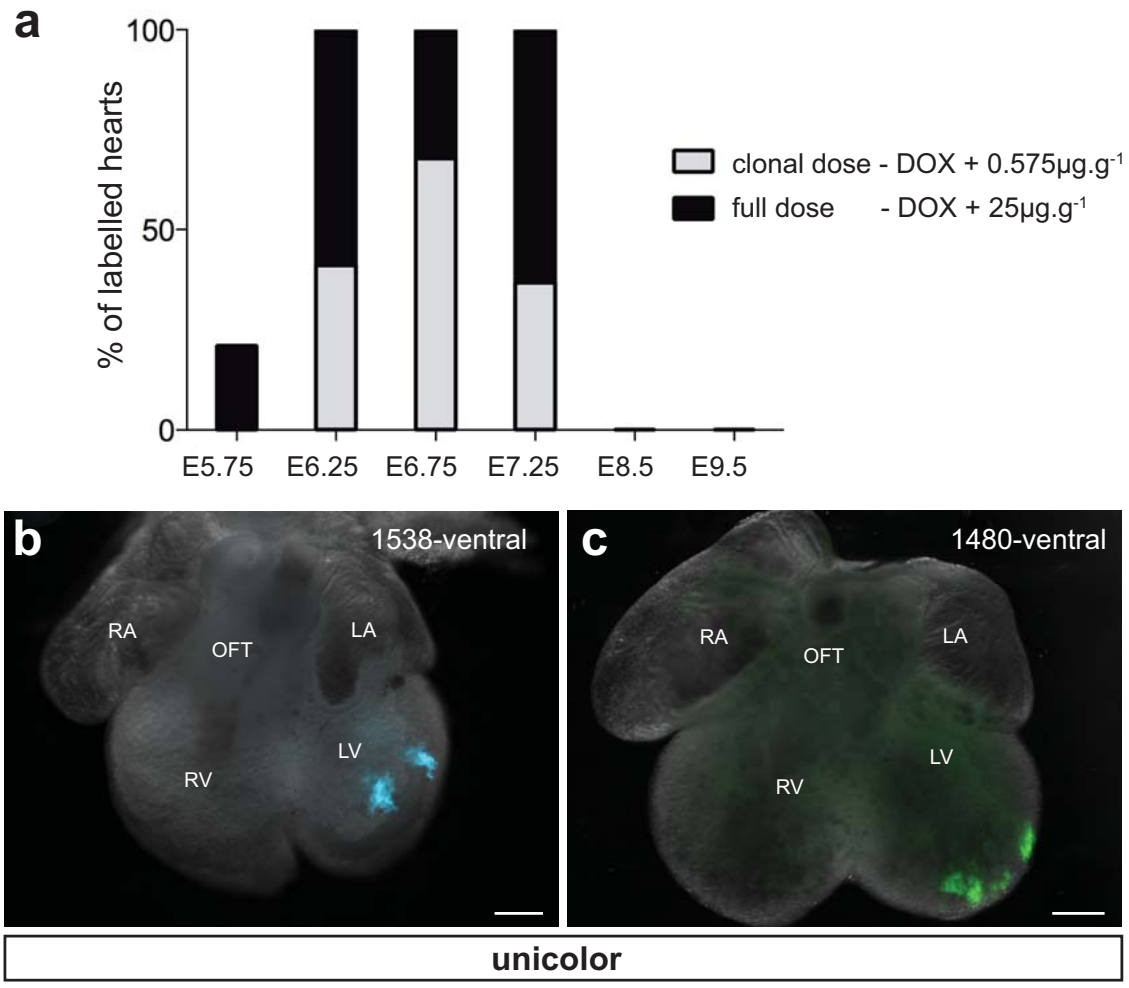
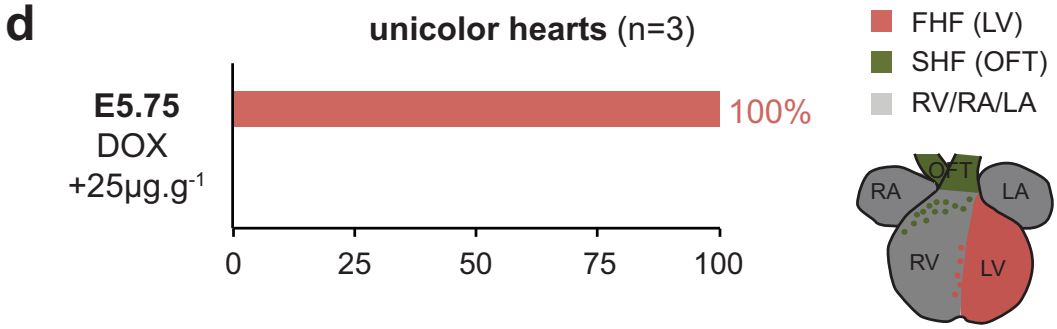
DOI: 10.1038/ncb3024

nature cell biology

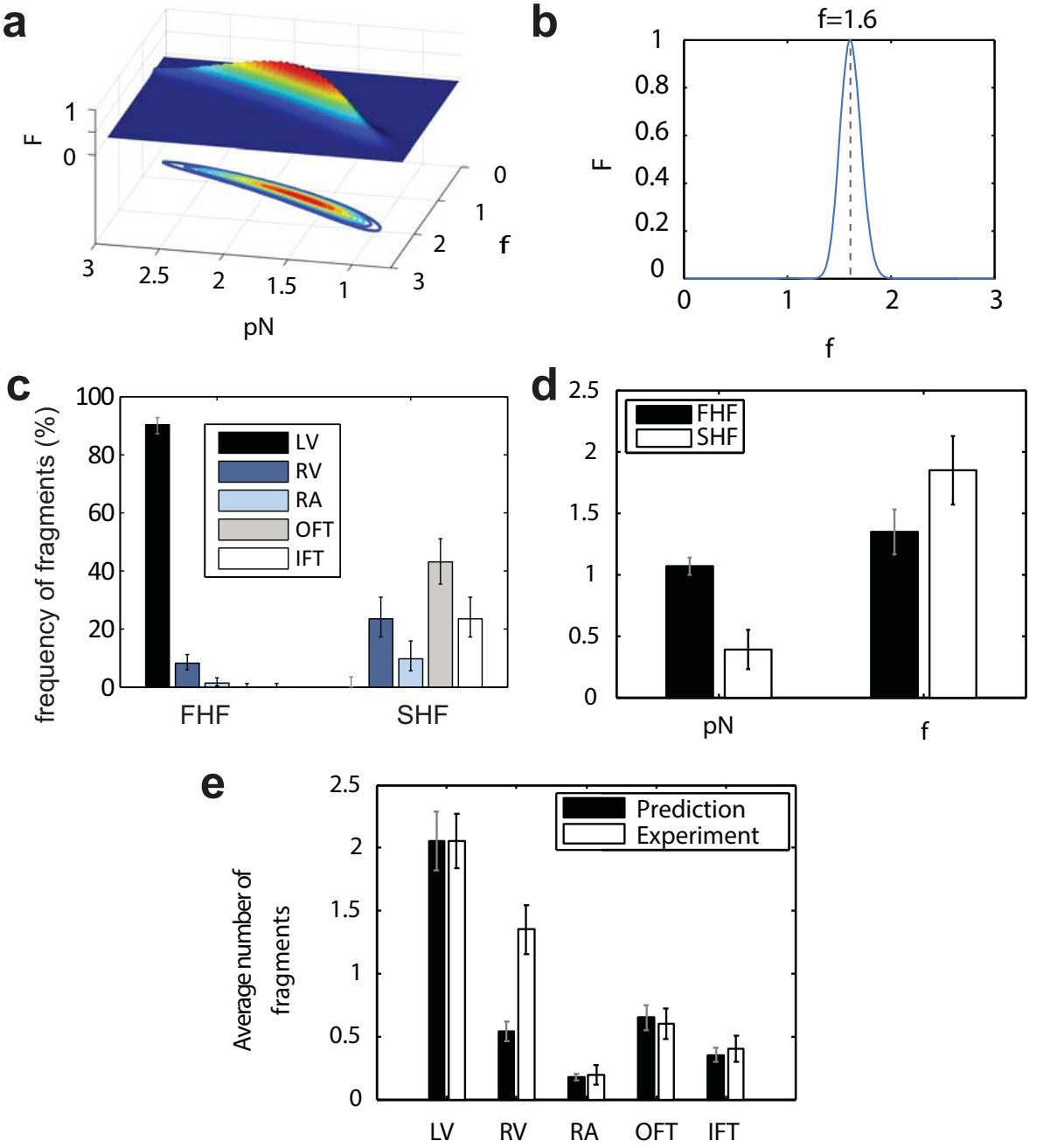


Supplementary Figure 1 *Mesp1-rtTA/tetO-Cre* transgenic mice induced Cre expression similarly to *Mesp1-Cre* Knock-in mice. **a-b.** Sections of E12.5 *Mesp1-Cre/Rosa-tdTomato* (a) and *Mesp1-rtTA/TetO-Cre/Rosa-tdTomato* hearts (induced by Dox administration between E6.25 and E7.5) (b) and co-stained with DAPI. Both transgenic hearts have a similar expression of the tdTomato with a negative region in the OFT that derive from Mesp1 negative neural crest cells (asterisks). **c-e.** Doxycycline injection has no effect on *Mesp1* expression during early mouse embryonic development. **c-d.** In situ hybridization for *Mesp1* expression in early embryo at E6.5. The detection of *Mesp1* mRNA in the primitive streak (PS) and the nascent lateral mesoderm is similar in embryo that did not receive DOX (c) and in embryos injected with DOX (+ DOX) (d). A, anterior; P, posterior. **e.** Expression of *Mesp1* analyzed by RT-qPCR in early embryos (E6.75) without (n= 9) or after doxycycline injections (n= 6). These data show no difference in *Mesp1* expression after DOX injection. **f-g.** In situ hybridization for *Cre* expression in early embryos at E6.75. The detection of *Cre* mRNA in the primitive streak (PS) and the nascent lateral mesoderm is similar in *Mesp1-Cre* knock-in (f) and in *Mesp1-rtTA/TetO-Cre* transgenic embryos injected with DOX at E6.25 (+ DOX) (g). *Cre* expression is similar to the endogenous *Mesp1* expression in wild type embryos (h). A, anterior; P, posterior. Scale bar: 500μm.

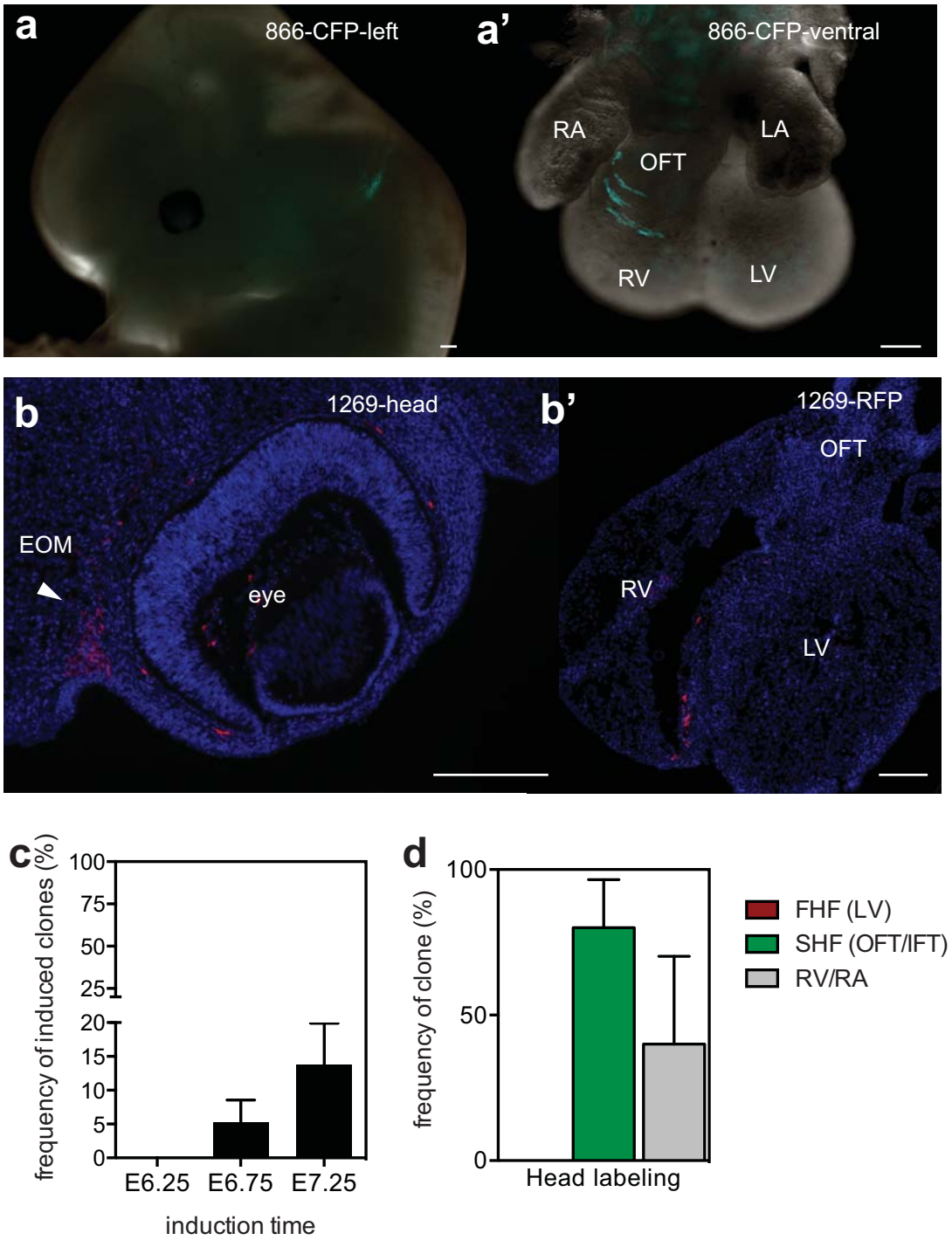




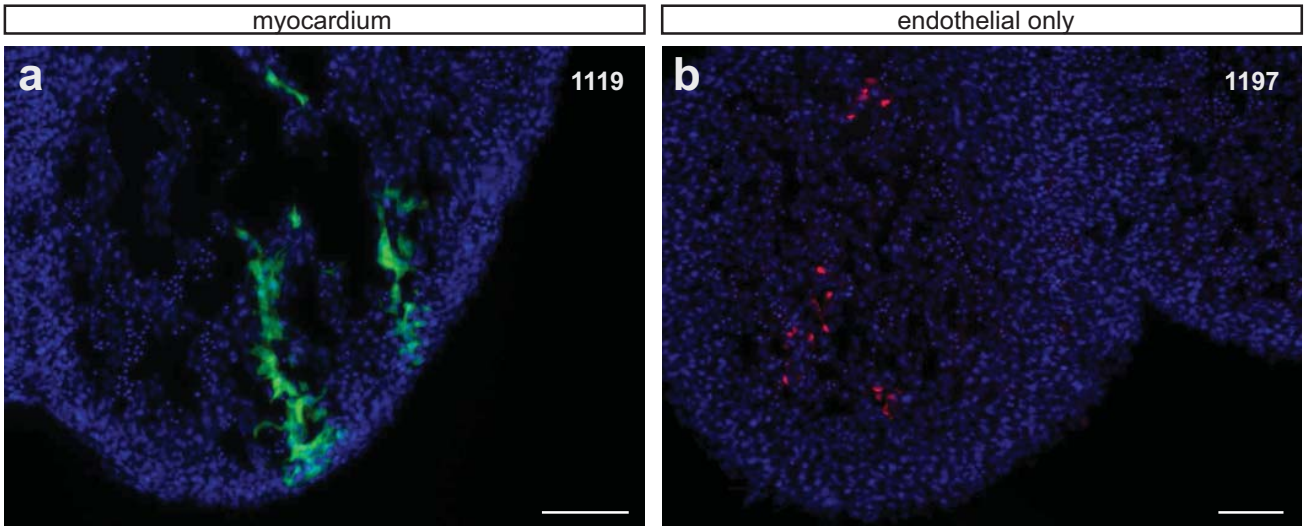
Supplementary Figure 2 Temporal Dox administration in Mes1-rtTA/TetO-Cre/Rosa-Confetti embryos. **a.** While clonal dose of DOX (0.575µg.g-1) induces labelling in *Mesp1-rtTA/TetO-Cre/Rosa-Confetti* embryos at E6.25 (n=53), at E6.75 (n=118) or at E7.25 (n=65), this dose was not sufficient to induce labelling at E5.75 (n=13). A much higher dose of Dox (25µg.g-1) was required to produce labelling at a clonal density at E5.75 (n=90). This 40 fold increase of DOX is likely to persist at a concentration sufficient to activate the Cre at the time of endogenous *Mesp1* expression. This high dose of DOX never labelled any heart after administration at E8.5 or E9.5 (n=24) supporting the absence of transgene expression after the end of endogenous *Mesp1* expression. **b,c.** Examples of *Mesp1-rtTA/TetO-Cre/Rosa-Confetti* unicolour labelled hearts at E12.5 induced at E5.75 after administration of high dose of Doxycycline (25µg.g-1). Note that each cluster is localized within the LV, FHF derivative and no labelling was detected other compartments. OFT, outflow tract; RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; IFT, inflow tract. Scale bars: 200 µm. The number on the upper right in each panel refers to the ID of the labelled heart. **d.** Quantification of the regional (FHF and SHF) contribution of patches of Mesp1 labelled cells in unicolour hearts induced at E5.75 with the high dose of Doxycycline (25µg.g-1), shows the exclusive labelling of the FHF (red) similarly to was found at E6.25 (Fig. 2m).



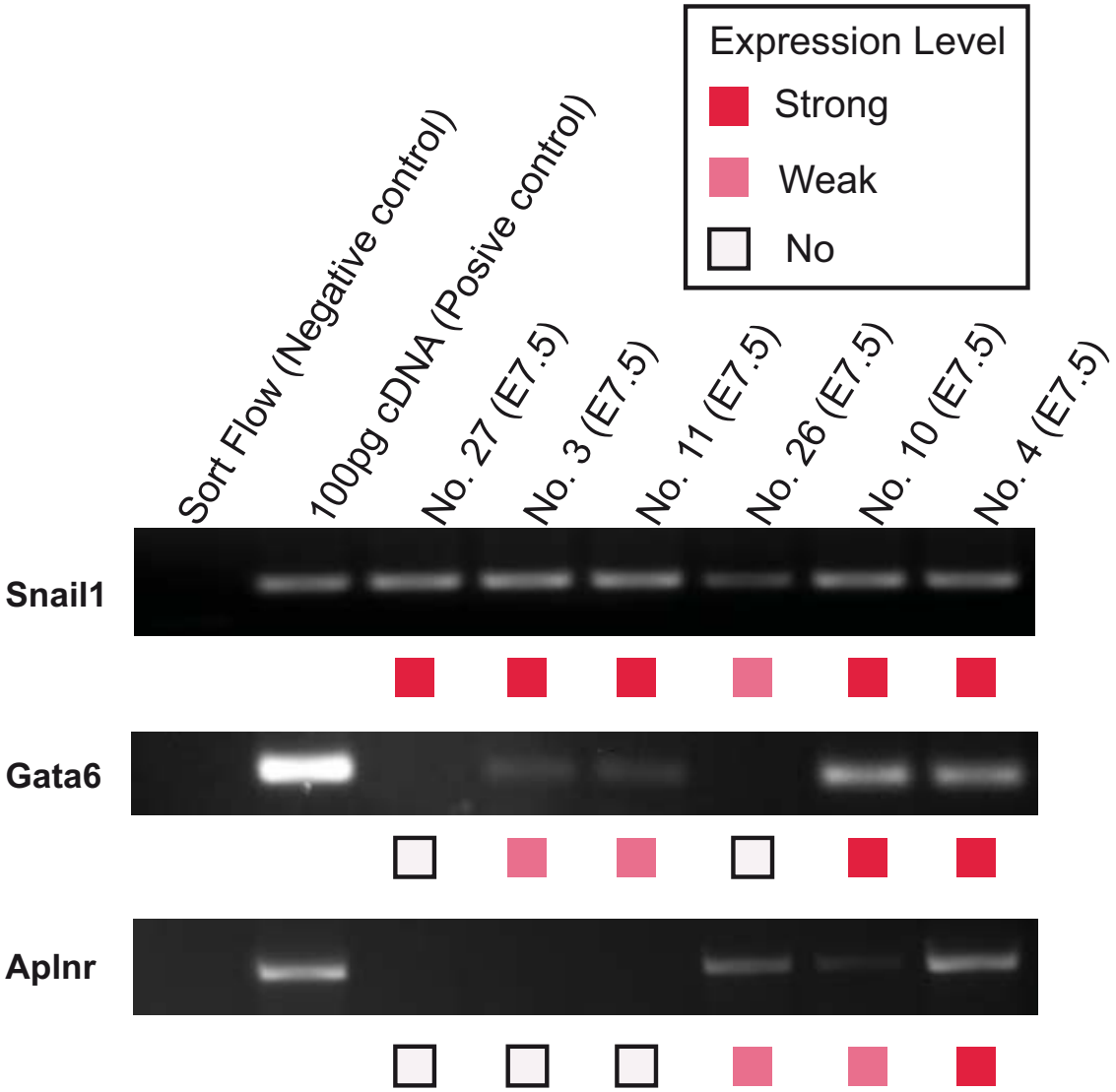
Supplementary Figure 3 Biostatistical modeling of the clonal fate data. **a.** The likelihood function F gives the probability of the experimental data for different values of the induction frequency pN and the fragmentation rate f. The numeric values have been rescaled such that the maximum of the likelihood function corresponds to 1. Color denotes the value of F, such that red signifies a large value and blue signifies small value. Lines of equal values are indicated on the bottom of the figure. One sees that the maximum value of F is relatively featureless along a curve in the pN-f-plane. To infer the values of pN and f we must therefore refer to an independent measurement of one of the two parameters. **b.** The multicolour labelling strategy allows us to independently infer the induction frequency pN=1.3 by evaluating the abundances of hearts with a given number of colours. With this, we are left with a slice through the pN-f-plane and the fragmentation rate can be determined with a higher accuracy. **c.** Monoclonal datasets (n=89) identify two subpopulations in *Mesp1* expressing cells: FHF progenitors, which contribute to the LV and SHF progenitors, which contribute to OFT and IFT. The plot shows the probabilities of monoclonal fragments in the different heart compartments. **d.** Values for the induction frequency, pN, and the fragmentation rate, f, for the two FHF (n=188) and SHF (n=102) precursors. While the overall induction frequency is higher for FHF precursors, which we attribute to highest expression of *Mesp1* at the early time points, the fragmentation rate is higher for SHF precursors. **e.** We may use the distribution of monoclonal fragments (c) to predict the distribution of fragments in all hearts (n=263). We find an excellent agreement with the notable exception of the RV, which might suggest the existence of an independent pool of progenitors contributing to RV morphogenesis. Error bars indicate one sigma (c and e) or 95% (d) confidence intervals.



Supplementary Figure 4 Late *Mesp1* progenitors also contribute to the head. **a-a'.** Example of a *Mesp1-rtTA/TetO-Cre/Rosa-Confetti* embryo with co-labelled head (a) and heart (a') at E12.5. Scale bars: 200 µm. **b-b'.** Sections of a *Mesp1-rtTA/TetO-Cre/Rosa-Confetti* labelled embryo showing labelling in the head in an extraocular muscle (EOM) (b) as well as in the right ventricle RV (b'). Scale bars: 200 µm. **c.** Temporal appearance of head muscle labelling inferred from all datasets (n=105 independent embryos translating to n=181 embryos by colour). Plotted is the fraction of head muscle labelling induced at each induction time for a given colour. Head muscles are preferentially labelled at the late time points. **d.** Regional contribution of head progenitors in monoclonal datasets (n=5), showing the co-labelling of the head with the heart and preferentially SHF derivatives or the right ventricle (RV). OFT, outflow tract; RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium. Errors bars indicate one sigma confidence intervals.



Supplementary Figure 5 Cohesive versus dispersive mode of growth of the myocardium and the endocardium. **a-b.** Sections of E12.5 *Mesp1-rtTA/tetO-Cre/Rosa-Confetti* unicolour hearts. **a.** Example of a compact myocardial YFP-labelled clone showing cohesive growth of the myocardium. **b.** Example of a dispersed endocardial RFP-unicolour clone showing dispersive mode of growth of the endocardium. Scale bars: 200 µm. The number on the upper right in each panel refers to the ID of the labelled heart.



Supplementary Figure 6 Semi-quantification of single cell RT-PCR analysis. Examples showing strong, weak and no gene expression in single cells.

nature cell biology

DOI: 10.1038/ncb3024

Supplementary Table legends

Supplementary Table 1 Table summarizing the clonal fate data according to their regional contribution and their probability of being monoclonal. Description of the labelling in *Mesp1-rtTA/tetO-Cre/Rosa-Confetti* induced heart. For each region, the number of labelled clusters is indicated. OFT, outflow tract, RV, right ventricle, LV, left ventricle, RA, right atrium, LA, left atrium, IFT, inflow tract.

Supplementary Table 2 Up-regulated genes in *Mesp1* GFP+ cells *in vivo.* Description of genes displaying a change in expression of >2 fold between *Mesp1*-GFP+ and *Mesp1*-GFP- cells at E6.5 and 7.5. (Fold change over GFP- cells at E6.5 ; Fold change over GFP- cells at E7.5) in 2 independent biological replicates. A gene ontology analysis was used to classify the up-regulated genes in the following categories : Transcription Factors/Chromatin Remodelling, Signaling pathways, Migration/Polarity/Guidance and Others (all biological function related to early embryo development that we can not put in any previous classes). **In bold** (overexpressed in Mesp1 GOF ESC) Underlined (overexpressed in Mesp1-GFP ESC).

Supplementary Table 3 Gene up-regulated in both *in vivo* and *in vitro* arrays.

Supplementary Table S4 primer sequences.

Supplementary Note

The interpretation of clonal fate data in the developing heart is complicated by the potential of clones to "fragment" into disconnected "clusters" as the tissue expands and cells rearrange, making the assignment of clonal progeny potentially ambiguous. However, by implementing a mathematical framework to analyze the statistics of the resulting clonal fragments, it is possible to faithfully recover information on the lineage potential and timing of the marked Mesp1 expressing cells. In the following supplementary note section, we detail the analytical program, presenting only the summary of the method in the main text.

We begin by introducing a simple stochastic framework to model clone fragmentation. Making use of the multicolor labeling strategy, we then use the observed clonal fate data to infer the induction frequency and the fragmentation probability of Mesp1 expressing cells. This enables us to assign clonal fragments to single-cell induction events. As a result, we show that Mesp1 expressing cells consist of two discrete subpopulations, one committed to the first heart field (FHF) derivatives and the other committed to the second heart field (SHF) derivatives. Further, with this assignment, we then show that these subpopulations are temporally distinct: while FHF precursors are mostly induced during the earliest two induction time points, SHF precursors are mainly induced during the latest two induction time points.

We note that the present scheme provides a general framework, which can be used to decipher the fate behavior and potency of progenitors using inducible genetic labeling methods in other developing tissues.

Induction frequency and clone fragmentation

The fragmentation of clones into separate clusters complicates the interpretation of clonal fate data. As both genetic labeling of cells and clone fragmentation happen in a stochastic manner, one finds a broad distribution of fragment numbers in labeled hearts. The number of precursors associated with such fragments is therefore not straightforwardly inferable from the data, as neither the induction frequency nor the degree of fragmentation is known (Fig. 3c). Fortunately, by addressing a statistical ensemble of labeled hearts, we can make use of statistical inference to assign with known confidence the provenance of the observed fragments. To this end, we first identify the induction frequency and degree of fragmentation for the heart as a whole by pooling data from all of the labeled hearts. With this result, we can then identify which of the collection of monochromatic patches in a heart are derived from a single induced cell. The basic strategy is illustrated in Figs. 3d,f. Restricting our analysis to monoclonal fragments, we then address the question of lineage potential. In addition, by analyzing the data by induction time using all labeled hearts, we also reveal the timing of lineage specification in the first and second heart field.

Induction frequency of Mesp1 progenitors following inducible labeling

To analyze the clonal fate data, let us begin by defining the probability, p, that following Dox administration, an early Mesp1 expressing progenitor cell becomes induced. Of course, this probability may vary according to the specific color of the fluorescent reporter gene. However, for now, let us consider just one of the colors and later generalize to multiple color

combinations. Then, if there are a total of N Mesp1 expressing progenitor cells at the time of induction, if the induction probability of each cell is considered statistically uncorrelated with its neighbors, the probability distribution for the number of induced cells for a given color is given by the binomial distribution,

$$P(m) = \binom{N}{m} p^m (1-p)^{N-m},$$

where the binomial coefficient is defined by $\binom{n}{k} = n!/[k!\,(n-k)!]$. Then, if the induction probability is clonal (i.e. p is of the order of $1/N$), we can make a Poisson approximation,

$$P(m) \approx \frac{N^m}{m!} e^{-pN} p^m = \frac{(pN)^m}{m!} e^{-pN}.$$

In particular, the probability that the tissue remains completely unlabeled is given by e^{-pN} and, as expected, the mean number of induced cells is $\langle m \rangle = pN$. Let us now consider the potential for Mesp1 cell-derived clones to undergo fragmentation.

Clone fragmentation

Once a precursor cell has been induced, in the course of its clonal expansion through cell proliferation, cells may disperse and the clone may fragment into multiple subclones. To account for this process of fragmentation, we may once again model these events as a statistically uncorrelated Poisson random process, so that the probability that an individual clone ends up in k fragments (i.e. it undergoes $k-1$ fragmentations) is given by

$$R(k) \approx \frac{f^{k-1}}{(k-1)!} e^{-f},$$

where f denotes the degree of fragmentation, defined as the average number of fragmentations experienced by a single cell-derived clone over the time course from induction to analysis. The degree of fragmentation represents the time-integral of the underlying fragmentation rate, which may itself vary over time. Of course, the degree of fragmentation may depend on the total size of the clone, i.e. large clones may fragment more than small clones. To investigate this, we calculated the surface area (SA) of clones in unicolor hearts, i.e. the percentage of the heart's surface clones cover ($n = 18$). We indeed found that clones vary significantly in SA at each induction time. However, comparing the size of these clones to the number of fragments did not show any significant correlation (Spearman's rank correlation, $\rho = 0.19$, $p = 0.45$). Therefore, since we will later see that most of these hearts are monoclonal, there is no evidence in the data that the degree of fragmentation depends on the size of clones.

With this definition, what then is the probability distribution of finding a total of k labeled fragments if m cells of a common color have been induced? In this case the number of fragmentation events is given by the total number of fragments minus the number of induced cells, $k-m$. Then, taking the fragmentation and induction events to be statistically independent, the branching probability, $S(k|m)$, is described by a Poisson process with an effective rate $m \cdot f$, and

$$S(k|m) = \frac{(mf)^{k-m}}{(k-m)!} e^{-mf},$$

where $k \geq m$. Therefore, with this result, we can infer the joint probability distribution for finding a heart with *m* induced cells giving rise to a total of k fragments of a given color,

$$J(m,k) = S(k|m)P(m) = \frac{(mf)^{k-m}}{(k-m)!}\frac{(pN)^m}{m!}e^{-mf-pN},$$

with $k \geq m$. (Note that, as defined here, the number of fragments must obviously be bound by the number of induced cells.)

In practice, in any given experiment, only the total number of labeled fragments (of a given color) is accessible – the underlying number of induced cells (clones) cannot be recovered for any given cluster of fragments. Moreover, we only have access to the frequency of clone fragments when at least one cell has been induced. The frequency of non-induced hearts is not recorded. Therefore, we should exclude the contribution of $m = 0$ from the statistical ensemble. In this case, the joint size distribution of "labeled" clones is therefore given by

$$J^{labeld}(m,k) = \frac{1}{(1-e^{-pN})}J(m,k)$$

where $m > 0$. For these persisting clones, since we measure only clonal fragments, we should combine all possible induction outcomes, from which we obtain the persisting fragment distribution,

$$F(k) = \sum_{m=1}^{k} J^{labeled}(m,k) = \frac{e^{-pN}}{1-e^{-pN}}\left[\sum_{m=1}^{k}\frac{(mf)^{k-m}}{(k-m)!}\frac{(pN)^m}{m!}e^{-mf}\right].$$

From this expression, we find that the average number of labeled fragments is given by

$$\langle k\rangle^{labeled} = \frac{pN(1+f)}{1-e^{-pN}}.$$

Hence, the average number of fragments in labeled hearts increases linearly with the degree of fragmentation f and, for moderately large values of pN, linearly with the induction frequency.

Fitting the data

Already at this point we may try to give an estimate on the values of pN and f for the heart as a whole. For our analysis, we do not take into accounts hearts, which are labeled in a specific color in the epicardium. The reason for this choice is that the outer unicellular layer is formed by cell migration quite late compared to the induction time (at E9.5), leading to very dispersed cells across the epicardium. This makes it difficult to distinguish labeled cells in the epicardium from those in the IFT and OFT. Making use of the formula for F and explicitly denoting its dependence on the parameters, we calculate the probability that the observed fragment numbers are found for any given degree of fragmentation, f, and induction frequency, pN. As the observations $k_1, k_2, \ldots$ are statistically independent this probability is given by:

$$F(\{k_1, k_2, \ldots\}|pN, f) = F(\{k_1|pN, f) \cdot F(k_2|pN, f) \cdot \ldots .$$

Treating F as a function of f and pN (which is then generally called the *likelihood*), we may now ask for the maximum of this likelihood function: the values of pN and f that yield the experimental data with highest probability. We consider these values to be the best estimate for the degree of fragmentation and induction frequency. From this analysis, we find that $pN = 1.7 \pm 0.8$ and $f = 1.3 \pm 1.0$ (95% confidence intervals). The large confidence intervals reflect the fact that the maximum of the likelihood function cannot be precisely determined along a curve in the pN-f-plane (Supplementary Fig. S3a). In other words, all of these parameters fit the experimental data equally well within the limits of statistical significance. In the following we therefore develop an independent approach to further constrain the two fitting parameters.

The multicolor labeling assay provides a means to independently infer the induction frequency, pN. To understand how, consider first the probability that a heart remains unmarked in any one of the three colors following Dox administration (we do not consider the GFP+ contributions as the induction frequency of these cells is found to be negligible – only one heart was found to contain any GFP+ cells. If the relative induction frequency of the three colors (YFP, RFP, and CFP) is equal, then this probability is given by $J(0,0)^3 = e^{-3pN}$. Therefore, the probability that an induced tissue involves all three colors (regardless of the number of fragments) is given by the probability that the tissue is clonally labeled in all three colors divided by the probability that the tissue is labeled at all:

$$C_{tricolor}^{labeled} = \frac{[1 - J(0,0)]^3}{1 - J(0,0)^3}.$$

Similarly, the chance that an induced tissue involves two out of three colors is given by

$$C_{bicolor}^{labeled} = \frac{3J(0,0)[1 - J(0,0)]^2}{1 - J(0,0)^3},$$

while those that involve only one color is set by,

$$C_{unicolor}^{labeled} = \frac{3J(0,0)^2[1 - J(0,0)]}{1 - J(0,0)^3}.$$

Since these probabilities are independent of the fragmentation probability, f, they can be used to provide an independent estimate of the induction frequency, pN.

To estimate the induction frequency of cells, we could immediately apply the results above to investigate the relative frequency of unicolor, bicolor and tricolor clones. However, in this case, we have to exercise some caution: Analysis of the unicolor clones shows that the induction frequency of the CFP is significantly smaller than the RFP and YFP with only 3 CFP+ clones out of a total of 23. By contrast, both the bicolor and tricolor hearts have a roughly equal representation of the three colors: In hearts which are labeled in any of the two heart fields, the multiplicity of RFP:YFP:CFP is 23:22:13 for bicolor and 55:57:66 for tricolor. While the multiplicity of colors is far from perfectly equal in the bicolor case, the comparison of the fitted distribution with the experimental data will further validate our approach. Then, since

$$\frac{C_{bicolor}^{labeled}}{C_{tricolor}^{labeled}} = \frac{3\,J(0,0)}{1 - J(0,0)},$$

we can use the ratio of the number of bicolor to tricolor clones to infer the induction probability, pN. With a measured ratio of 1.10, we find that $pN = 1.31 \pm 0.05$. We performed the same calculations by explicitly taking the lower induction frequency of CFP in bicolor hearts into account. With this approach we obtain an average induction frequency, pN, of roughly 1.4 for all fluorescent markers in tricolor hearts and RFP and YFP in bicolor hearts, i.e. we find only a minor deviation for most observed hearts. The induction frequency of CFP in bicolor hearts involving CFP would correspondingly be about 0.7. Therefore, the pN value only changes significantly for CFP in bicolor hearts. As we will see below, this only marginally influences the outcome of the statistical analysis.

Therefore, on average, pooling all of the data from the three induction times, we expect that approximately 1.3 cells are induced per color in each heart. However, this estimate includes hearts where there are no marked clones at all. To obtain the induction frequency for clones that contain at least one marked cell we have to divide $P(m)$ by the probability that a heart is unlabeled in a given color, $1 - P(0)$. Therefore, when restricting attention to labeled hearts, the probability of m induction events is given by

$$P^{labeled}(m) = \frac{P(m)}{1 - e^{pN}}.$$

Consequently, we obtain for the mean number of induced cells in labeled hearts

$$\langle m \rangle^{labeled} = \frac{pN}{1 - e^{-pN}},$$

which, in the present case, is approximately equal to 1.8.

With this estimate for pN, we may now turn to consider the probability distribution of fragment numbers, $F(k)$, and the estimate of the degree of fragmentation, f. By fixing the value of pN we can restrict the possible parameters to a slice through the pN-f-plane (Supplementary Fig. S3a and b). Making use of the formula for $F(k)$, analysis of the maximum likelihood shows that $f = 1.6 \pm 0.2$. Both, the induction frequency and the degree of fragmentation are in agreement with the values obtained above.

As a further consistency check, we may note that, with these values of $pN$and f, the average number of fragments labeled hearts, $\langle k \rangle^{labeled}$, is given by 4.6, which compares excellently with the experimentally measured value of 4.7. Indeed, with the fitted values, the predicted fragment number distribution compares favorably with the measured distribution, as indicated in Fig. 3e.

Note that, to estimate pN, we made use of the fact that the induction frequency is roughly equal for all fluorescent markers for bicolor and tricolor hearts. However, for unicolor hearts, the frequencies of the different colors are manifestly different. If the statistical weights of the different colors in unicolor hearts were representative for all hearts, this might lead one to conclude that induction frequencies are also different for multicolor hearts. Since the degree of fragmentation should not depend on the color of fluorescent label, we may analyze the total fragment numbers, $\langle k \rangle^{labeled}$, to test whether or not the overall induction frequency does indeed depend on the color of the fluorescent marker. If unicolor hearts were representative for all hearts, one would expect that $\langle k \rangle^{labeled}$ depends sensitively on the induced color. Taking all labeled hearts (uni-, bi- and tricolor), we find that

the average number of fragments is 4.6 ± 0.3, 4.1 ± 0.4 and 5.4 ± 0.3 for YFP (n=87), CFP (n=83) and RFP (n=92), respectively, which suggests that the induction frequency is only weakly dependent on the color of the fluorescent marker. We attribute the apparently non-representative induction of unicolor hearts to a thresholding effect in which the sensitivity of different colors to induction is amplified when the level of Cre expression is low.

Mesp1 positive cells are restricted to either the first or the second heart field

With the value of the induction frequency and degree of fragmentation fixed, we may now make an informed decision on which hearts are monoclonal. To begin, we note that the probability that k patches derive from m clones is given by

$$L(m|k) = \frac{J(m,k)}{F(k)}.$$

Therefore, the probability that k fragments are of clonal origin is given by $L(m = 1|k) = J(m = 1|k)/F(k)$. Similarly, the probability that k patches derive from more than one clone is obtained by summing over all induction outcomes larger than one,

$$L(m > 1|k) = \sum_{m=2}^{\infty} \frac{J(m,k)}{F(k)} = 1 - L(m = 1|k).$$

To make a decision on the maximum number of fragments we consider to be of clonal origin, we compare these two: we consider k fragments to be monoclonal if the probability that that they stem from a single cell, $L(m = 1|k)$, is larger than the probability that they stem from more than one cell, $L(m > 1|k)$. Specifically, in the spirit of the theory of Bayesian inference, we compute the logarithm of these two probabilities and multiply by -2,

$$D = -2\ln\left[\frac{L(m > 1|k)}{L(m = 1|k)}\right].$$

With this definition, fragments are considered monoclonal if $D > 0$. Taking the values for pN and f obtained in the previous section, we find that 3 or less fragments of a single color are likely monoclonal, cf. Fig. 3 g and h. Indeed, with this classification, we expect that some 12% of hearts designated as monoclonal would in fact be polyclonal.

How does the approximation of equal induction frequencies in bicolor hearts affect this threshold value? If the type of fluorescent protein does not influence the degree of fragmentation, this would mean that the likelihood that a given number of patches is monoclonal is higher for CFP than for the other fluorescent markers in bicolor hearts. In other words, treating the induction frequency separately would allow us to treat some bicolor hearts as monoclonal, which have slightly more than 3 patches. Hence, treating the induction frequencies of different colors separately only marginally increases the sample size of monoclonal hearts.

With these results we may now restrict our analysis to hearts, in which a single clone has been labeled per color. Remarkably, we find that, of the 89 cases of hearts that are deemed to have marked fragments of clonal origin in either the FHF (LV) or the SHF (OFT and IFT), all are restricted to one or the other heart field. None of these clones contribute to both heart

fields. (We note that this apparently perfect segregation of clones is further assisted by the histogenesis which, as we will see below, leads to the temporal separation in the specification of progenitors of the two heart fields.) By contrast, of the 69 clones that have fragments in the FHF, 15% also have fragments in the other heart compartments (i.e. the RV and the RA). Similarly, of the 20 clones that have fragments in the SHF, 55% have fragments in other heart compartments. Figure. 3i shows the percentage of clones that contribute to the different heart compartments given that they contribute to the FHF (left) or SHF (right), respectively. We conclude that, by the time of induction, Mesp1+ cells are already lineage restricted, contributing to either the first or second heart field, but not both. However, both Mesp1+ subpopulations are able to contribute to cells in the remaining heart compartments.

To further scrutinize the properties of Mesp1+ cells we calculated the non conditional probabilities with which these cells contribute fragments to the different heart compartments (Supplementary Fig. 5c). For example, we find that about 10% of the fragments of FHF precursors end up in other compartments. This means, as 85% of FHF precursors exclusively contribute to the FHF, that from the remaining 15% approximately two out three fragments end up in other heart compartments.

These clones have an average number of fragments of $k_F = 2.10 \pm 0.01$ and $k_S = 2.60 \pm 0.02$ for FHF and SHF precursors, respectively. Taking into account the fact that, by introducing a threshold of $k = 3$, we neglect clones with a large number of fragments (i.e. those lying in the tail of $L(m = 1|k)$), this result agrees well with the predicted value for the overall population, viz. $f + 1 = 2.6$. This also tells us that fragmentation of SHF precursors is slightly higher than fragmentation of FHF precursors, which raises the question of whether the former might migrate more.

As a consistency check we may estimate the induction frequency and the degree of fragmentation of the two types of precursors independently by following the steps from the previous section. Since most of the contributions of these cells go into the FHF (LV) and the SHF (OFT and IFT), respectively, we restrict our analysis to fragments in these compartments. For the FHF precursors we find that $pN = 1.07 \pm 0.07$ and $f = 0.78 \pm 0.17$ for FHF precursors and $pN = 0.39 \pm 0.16$ and $f = 1.00 \pm 0.22$ for SHF precursors (the values for pN are shown in Supplementary Fig. S3d). On the one hand, these results are in agreement with a higher degree of fragmentation of SHF precursors. On the other hand, this tells us that the induction frequency is significantly higher for FHF precursors. Moreover, noting that, as the overall induction frequency is the sum of the individual induction frequencies, the individual values for pN are in good agreement with the values obtained for the whole population of Mesp1 expressing cells. The fragmentation rates are, expectedly, lower, as we neglected fragments located in heart compartments other than the LV, OFT and IFT.

With the probability of single clones to contribute to the different heart compartments defined, we may now predict the overall distribution of fragments in all hearts. To this end, we may account for the neglect of large monoclonal clusters by calculating the effective degree of fragmentation of the two subpopulations as follows: $\bar{k}_{F,S} = 2k_{F,S}/(k_F + k_S) \cdot (f + 1)$. In other words, we use the monoclonal data to infer the relative deviation of each subpopulation from the average number of fragments of a single clone, $f + 1$, in the overall population. From this we obtain an estimate for the fragmentation rates of each subpopulation, viz. $f_F = 1.4 \pm 0.2$ and $f_S = 1.9 \pm 0.3$ (Supplementary Fig. S3d). Since the 95% confidence interval of the difference between these two values does not contain 0, this

difference in fragmentation rates is statistically significant. With this result, we are then able to predict the experimentally observed distribution of fragments in all hearts with remarkable accuracy, cf. Supplementary Fig. S3e. One notable exception is the number of fragments in the RV, which is twice as large as that expected. We attribute this apparent discrepancy to the fact that there are clones that exclusively contribute to the RV. These are not included in the analysis of the monoclonal hearts, but they do, of course, contribute to the overall distribution of fragments.

Temporal induction of the FHF and SHF progenitors

To investigate the temporal order of fate specification we now take into account the time point of Dox administration. First, we address the proportion of FHF and SHF precursors that are labeled at each induction time. From the previous results, we know that FHF and SHF precursors are mutually exclusive with respect to their contributions to the LV (FHF) on the one hand and the OFT and IFT (SHF) on the other. As the labeling of clones can be considered statistically independent, the average number of induced FHF and SHF precursors is proportional to the average number of fragments in these compartments $\langle m_{F,S}\rangle(t) = p_{F,S}(t)N_{F,S}(t) \propto \langle k_{F,S}\rangle(t) = K_{F,S}(t)/H$, irrespective of whether or not the hearts are monoclonal. Here, $p_{F,S}(t)$ denotes the induction probability of a single FHF or SHF derivatives, respectively, $N_{F,S}(t)$ are the total numbers of FHF or SHF derivatives in the early tissue, and $\langle k_{F,S}\rangle(t)$ signifies the average number of fragments in the corresponding heart compartments. The total number of heterozygotic mice, H, is, by the design of the experiment, independent of the induction time. With this, we can calculate the proportion of induced cells at each time point t of Dox administration, $r_{F,S}(t)$. To this end, we divide the average number of clones that were induced at time t, $\langle m_{F,S}\rangle(t)$, by the total number of induced clones, to obtain

$$r_{F,S}(t) = \frac{K_{F,S}(t)}{\sum_t K_{F,S}(t)} = \frac{p_{F,S}(t)N_{F,S}(t)}{\sum_t p_{F,S}(t)N_{F,S}(t)}.$$

Then, since the probability $p_{F,S}(t)$ of a single Mesp1 expressing cell to be induced should not depend on the particular time point of induction, we can make the simplification

$$r_{F,S}(t) = \frac{N_{F,S}(t)}{\sum_t N_{F,S}(t)}.$$

Therefore, the ratio represents the proportion of FHF and SHF derivatives that are induced at time t. Importantly, this proportion can be estimated by analyzing the total numbers of fragments in *all* hearts. From this analysis, we find that most FHF derivatives are induced at induction times E6.25 and E6.75 (89%) while most SHF derivatives are labeled at induction times E6.75 and E7.25 (95%), cf. Fig. 3k.

Finally, one may also use the analysis of data from the monoclonal fragments alone as a consistency check. Here, the number of induced clones, $m_1(t)$ and $m_2(t)$ are directly accessible. In agreement with the results incorporating all hearts we find that FHF precursors are mostly induced early (E6.25 and E6.75) and SHF progenitors are mostly induced late (E7.25). However, in contrast to the results obtained from the full data set, none or very few of the SHF precursors are induced at E6.75. We attribute this to the fact that, at this time point, Mesp1 is only expressed at low levels in SHF precursors. As a result, these cells will

only be induced at induction frequencies higher than the relatively low doses that define the monoclonal data points.

We may also infer the total numbers of induced clones of each subpopulation for each time point by dividing the total number of fragments, $K_{F,S}(t)$, by the average number of fragments that single clones contribute to the FHF or SHF, respectively, $\bar{k}_{F,S} \cdot \pi_{F,S}$: $m_{F,S}(t) = K_{(F,S)}(t)/(\bar{k}_{F,S} \cdot \pi_{F,S})$. Here, $\bar{k}_{F,S}$ is the corrected overall number of fragments and $\pi_{F,S}$ denotes the probability that a fragment of such a single clone ends up in the FHF (LV) or the SHF (OFT and IFT), respectively, cf. Supplementary Fig. S3c. With $\langle m_{F,S} \rangle(t) = m_{F,S}(t)/H$, of course, as neither $\bar{k}_{F,S}$ nor $\pi_{F,S}$ do not significantly depend on time, this exactly reproduces $r_{F,S}(t)$.

From this analysis, we find that, overall, 254 ± 22 FHF and 138 ± 17 SHF precursors have been induced. Hence, FHF precursors have roughly twice the induction frequency, pN, of SHF precursors, which compares favorably with the estimated induction frequencies for the two subpopulations that we obtained by comparing the numbers of tricolor and bicolor hearts. Given that clone induction is statistically independent, the numbers of FHF and SHF precursors follows a binomial distribution. In this case we may employ *Fisher's exact test* to calculate the probability that differences in the number of induced clones between two induction times are the result of pure chance (Fisher, 1922). We find that this probability is small when comparing any two induction times. The differences in the number of lineage specified cells are statistically highly significant ($p < 0.0001$). Hence, we find that Mesp1 positive cells consist of two temporally distinct subpopulations. While FHF derivatives are largely specified early, most SHF derivatives are induced at the latest time points.

This completes the quantitative statistical analysis of the clonal fate data. In summary, making use of a multicolor labeling strategy, we employed statistical inference to estimate the induction frequency and the degree of fragmentation in a pooled dataset. This allowed us, for a given color, to identify the ensemble of monoclonal hearts. Restricting the analysis to these hearts we showed that Mesp1 expressing cells are already committed to either contributing to the first heart field or the second heart field. We calculated the contribution of these two types of precursors to the different heart compartments and showed that the precursors to the two heart fields are induced in two distinct temporal regimes.

Bibliography

Fisher, R. A. (1922). On the interpretation of χ2 from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society , 85* (1), 87-94.