Explanatory Notes About Data

Title: Data for re-evaluation of the causes of variation among mouse aggregation chimaeras

Relevant Publication:

Purpose of Study (This text is taken from the above publication.)
The composition of mouse aggregation chimaeras, comprising two genetically distinct cell populations, varies much more widely than that of X-inactivation mosaics. This is usually attributed to sampling events that occur in chimaeric aggregates but precede random X-chromosome inactivation, which begins in the epiblast lineage soon after implantation at around E5.5. An early theoretical model proposed that almost all the extra variation in chimaeras arises by spatially constrained, geometrical allocation of inner cell mass (ICM) cells to the epiblast and primitive endoderm (PrE). However, this model is inconsistent with more recent embryological evidence. The first aim of the study was to analyse the composition of chimaeric blastocysts and mid-gestation chimaeras, produced in several previously published studies, to discriminate between the original hypothesis and a spatially unconstrained two-step hypothesis. The second aim was to simulate the first two allocation events in preimplantation chimaeras, without spatial constraints, to determine how epiblast cell numbers and the extent of variation at step 1 affected variation of the epiblast in simulated 64-cell chimaeric blastocysts.

Summary of Results (This text is taken from the above publication.)
Analysis of published results for chimaeric blastocysts suggests that some variation already exists among aggregates before segregation of the ICM and TE. This would include both biological variation between aggregated embryos and experimental variation introduced by the chimaera production procedures. Variation among chimaeric ICMs was greater than among whole blastocysts, implying that significant additional variation arises when cells are allocated to the ICM or trophectoderm (TE). Moreover, in mid-gestation chimaeras, the compositions of the epiblast derivatives were not negatively correlated with the PrE derivatives, again suggesting that significant variation existed among ICMs before epiblast formation. The composition of samples from postimplantation stage chimaeras were more strongly correlated within the epiblast lineage or within the PrE lineage than between the two lineages, implying that the second allocation step, involving the segregation of the epiblast and PrE lineages, is also a significant source of variation among chimaeric epiblasts. Computer simulation results were consistent with the conclusion that the wide variation in chimaeric epiblast composition arises by two stochastic allocation steps, during the formation of the ICM and epiblast respectively. Later allocation events will cause variation among both chimaeras and X-inactivation mosaics. We also suggest that previously published U-shaped frequency distributions for chimaeric placenta composition might be explained by how TE cells are allocated to the polar TE and the subsequent movement of TE cells from the polar TE to the mural TE.

Three types of data are used for this study:

1. Data on the composition of mouse blastocyst chimaeras
These data were used for Fig. 2 and Supplementary Table S2 in the above publication and are explained below on pages 2–3.

2. Computer simulation data on the composition of simulated mouse blastocyst chimaeras
The computer simulations were used to prepare Fig. 5 and Supplementary Figs S3–S6 in the above publication and are explained below on pages 4–5.
3. Data on the composition of E12.5 postimplantation mouse chimaeras

These data were used to prepare Figs. 3 & 4, Supplementary Tables S3–S8 and Supplementary Figs S1 & S2 in the above publication. The data are archived elsewhere: Flockhart, J. H., Tang, P.-C., MacKay, G. E., Keighren, M. A., Kissenpfennig, A., Wilson, L., Pratt, T., Mason, J. O., Price, D. J. and West, J. D. (2017). Data on the composition of four balanced and four unbalanced series of E12.5 fetal mouse chimaeras, [dataset]. University of Edinburgh, Edinburgh Medical School. 
http://dx.doi.org/10.7488/ds/2056.

1. Data on the composition of mouse blastocyst chimaeras

Relevant Publications:
The data are numerical data from the following two published studies:

Nature of the Blastocyst Chimaera Data: Two Excel workbooks show numerical data on the composition of different regions in two series of mouse blastocyst chimaeras. The chimaeras were made by aggregating two intact cleavage stage embryos (originally called chimaera series CeB) or two half embryos (originally called chimaera series PCT-III or series S2n↔*S2n). In each case, one of the aggregated embryos carried the TgN(Hbb-b1)83Clo reiterated transgenic marker (abbreviated to Tg) and cells carrying this marker were identified by DNA in situ hybridisation to the reiterated transgene in serial sections of blastocysts. All cells were scored as positive or negative and the percentage of Tg-positive cells was corrected using control data from Tg/-↔Tg/- control chimaeras. The percentage of Tg-positive cells in the control Tg/-↔Tg/- chimaeras was less than the theoretical 100% because parts of the same nucleus may occur on adjacent sections but the Tg sequences may only be present in one section. The corrected percentage of Tg-positive cells in the whole trophectoderm (TE) was calculated as the weighted mean of polar trophectoderm (pTE) and mural trophectoderm (mTE) values and the corrected percentage of Tg-positive cells in the whole blastocyst was calculated as the weighted mean of the inner cell mass (ICM), pTE and mTE, using the cell numbers in ICM, pTE and mTE.

Blastocyst Chimaera Experiment 1

Experimental chimaera series CeB
Strain combination: 8-cell (C57BL x CBA)F2↔8-cell [(C57BL x CBA)F1 x TGB]
Genotype combination: wild-type (no transgene) -/- ↔ hemizygous, Tg/-
Culture period before analysis: E2.5 to E3.8

Positive control blastocyst chimaeras (used to correct % Tg+ve cells in series CeB)
Control series name: CeP2
Strain combination: 8-cell (C57BL x CBA)F2↔8-cell [(C57BL x CBA)F1 x TGB]
Genotype combination: hemizygous, Tg/- ↔ hemizygous, Tg/-
Culture period before analysis: E2.5 to E3.8

Mouse Strain Abbreviations
CBA = CBA/Ca
C57BL = C57BL/OlaHsd
(TGB = Stock carrying TgN(Hbb-b1)83Clo marker transgene)
Excel Workbook (with 2 spreadsheets):
‘Tg-positive cells in chimaeric blastocysts_Everett & West, 1996.xlsx’

Spreadsheet 1: Mean %Tg-pos in series CeB. This shows the raw data (Tg-positive and Tg-negative cells) in different regions of the blastocyst for 17 -/-→Tg/- chimaeras from series CeB. It also shows the calculated uncorrected percentage of Tg-positive and uncorrected percentage of Tg-negative cells. The percentage of Tg-positive cells was corrected, as shown, to allow for the expected frequency of false-negative nuclei, using data from the 7 Tg/-→Tg/- positive control chimaeric blastocysts in positive control series CeP2, reported in Everett & West (1996), which showed that 433/777 nuclei (55.7%) were Tg-positive.

Spreadsheet 2: Frequency distribution %Tg-pos. The corrected percentages of Tg-positive cells in the inner cell mass (ICM), whole trophectoderm (TE) and whole blastocyst were sorted separately so data could be plotted as frequency distributions.

Blastocyst Chimaera Experiment 2
Experimental chimaera series S2n↔*S2n (also called series PCT-III)
Strain combination: ½ 2-cell (C57BLxCBA)F2↔½ 2-cell [(C57BLxCBA)F1 × TGB]
Genotype combination: wild-type (no transgene), -/-↔ hemizygous, Tg/-
Culture period before analysis: E1.5 to E4.5

Positive control blastocyst chimaeras (to correct % Tg+ve cells in series S2n↔*S2n)
Control series name: *S2n↔*S2n
Strain combination: 8-cell (C57BLxCBA)F2↔8-cell [(C57BLxCBA)F1 × TGB]
Genotype combination: hemizygous, Tg/- ↔ hemizygous, Tg/-
Culture period before analysis: E1.5 to E4.5

Excel Workbook (with 5 spreadsheets):
‘Tg-positive cells in chimaeric blastocysts_Tang et al 2000.xlsx’

Spreadsheet 1: Control for Tg signal. This shows the number of Tg-positive cells and total cell numbers in the pTE (polar trophectoderm), mTE (mural trophectoderm) and ICM (inner cell mass) for 19 Tg/-→Tg/- positive control blastocyst chimaeras from series *S2n↔*S2n. It also shows the calculated percentage of Tg-positive cells in each blastocyst region, the mean of which was used as a correction factor for the experimental chimaeras (see spreadsheet 2 below).

Spreadsheet 2: Uncorrected %Tg-pos. This shows the number of Tg-positive cells and total cell numbers in the pTE, mTE and ICM for 38 -/-→Tg/- experimental blastocyst chimaeras from series S2n↔*S2n. It also shows (1) the calculated uncorrected percentage of Tg-positive cells in each of these blastocyst regions, (2) the percentages of pTE and mTE cells in the whole TE and (3) the percentages of pTE, mTE and ICM cells in the whole blastocyst.

Spreadsheet 3: Corrected %Tg-pos. This shows the uncorrected percentage of Tg-positive cells in the pTE, mTE and ICM for the experimental blastocyst chimaeras shown in spreadsheet 2, the correction factors derived for each of these blastocyst regions in spreadsheet 1 and the corrected percentage of Tg-positive cells. If the corrected percentage was greater than 100%, it was assumed that the corrected percentage of Tg-positive cells was 100%.

Spreadsheet 4: Corrected combined lineages. This shows the calculation of the corrected percentage of Tg-positive cells for the whole TE and the whole blastocyst.

Spreadsheet 5: Frequency distribution %Tg-pos. The corrected percentages of Tg-positive cells in the inner cell mass (ICM), whole trophectoderm (TE) and whole blastocyst were sorted separately so data could be plotted as frequency distributions.
2. Computer simulation data on the composition of simulated mouse blastocyst chimaeras

Explanation of the Simulation. (Most of this text is taken from the publication shown on page 1.)

The aim was to simulate the first two allocation events in preimplantation chimaeras, without spatial constraints, to determine how epiblast cell numbers and the extent of variation at step 1 affected variation of the epiblast in simulated 64-cell chimaeric blastocysts.

Two computer simulation models of chimaeric blastocysts composed of ‘black’ and ‘white’ cells were developed to model how variation in the epiblast of chimaeric blastocysts could arise in two steps without spatial constraints. Variation in the percentage of black cells in each lineage was introduced at allocation step 1 [formation of the inner cell mass (ICM) and trophoderm (TE)] and allocation step 2 [formation of the epiblast and primitive endoderm (PrE)]. In simulation model A, the first allocation step introduced little variation. In this model, all 64-cell chimaeric blastocysts contained 28 ICM cells and 36 ICM cells. Overall there were 50% black cells and 13, 14 or 15 of the ICM cells were black cells so the percentage of black cells in the ICM only varied from 46.4% to 53.6%

In simulation model B, more variation was introduced at allocation step 1. A simulated 8-cell black embryo was aggregated with a simulated 8-cell white embryo to form an aggregate of 16 cells, which all underwent two rounds of cell divisions, without cell loss, to produce a 64-cell chimaeric blastocyst. As a chimaeric aggregate has twice the normal number of cells, it will have some inner cells from the outset and the simulation assumed that all aggregates initially had 5 inner and 11 outer cells (31.3% inner cells). In most biological aggregates, both embryos will contribute to the inner cells, so at least one black and one white inner cell were included in the inner cells of the simulated 16-cell aggregates. Three other cells were allocated randomly to the inner lineage and the remainder became outer cells.

The proportion of inner cells was increased over the next two rounds of simulated cell divisions. After the first round of cell divisions, simulated 32-cell chimaeras had 12 inner and 20 outer cells (37.5% inner cells) and after the next round, simulated 64-cell chimaeras had 28 (43.8%) inner cells (ICM cells) and 36 outer cells (TE cells). At each division all inner cells divided symmetrically (conservatively) to produce two inner cells whereas some outer cells divided symmetrically to produce two outer cells but other divided asymmetrically (differentiatively) to produce one inner and one outer cell to increase the proportion of inner cells, as required. Black and white outer cells were selected randomly for asymmetrical division.

In both simulation models, additional variation was introduced at allocation step 2, when ‘black’ and ‘white’ ICM cells were allocated to the epiblast and PrE lineages. We simulated random allocation of black and white ICM cells to the epiblast (with unallocated ICM cells becoming PrE) at the 64-cell stage in chimaeras, which is equivalent to the 32-cell stage in non-chimaeras. There were 28 simulated ICM cells and the number of cells allocated to the epiblast lineage varied (10 epiblast cells in series A1 & B1, 14 in A2 & B2 and 18 A3 & B3).

Nature of Data The archived simulation data comprise:

(1) Python source code for simulation in three formats (Word document, pdf file and sublime text document)

(2) Data from 6 series of simulations in 6 separate folders (A1 – B3). Each folder contains 10 csv files, each representing a separate set of 1000 simulations plus an Excel spreadsheet with summary data for the 10 sets in a series. In the output each set is called a ‘run’ and the 10 runs are numbered 0–9. File names indicate the series, number of epiblast cells and the run (or set) number (e.g. csv
file “b2-epi-14-run7” has data for run 7 (set 7) from series B2, where 14 of the 28 ICM cells were allocated to be epiblast cells.

Output data for each run (set) represents the % black cells for the four simulated lineages (epiblast, PrE, ICM and TE) for 1000 simulations (numbered 0 – 999). The mean % black cells and variance of the % black cells for the four simulated lineages are shown at the bottom of each column. Each of the six Excel summary spreadsheets shows the means and variances for each of the 10 sets of simulations in one of the six series of simulations.