

## MODEL ANALYSIS

He et al. are advancing a long running argument in this paper by building a statistical model of a retinal progenitor cell that expresses a general view that retinas are built from a population of "equipotent RPCs", and that the realised cell total and cell fate outcomes for individual cell lineages are in some way "stochastic". Stochasticity is rarely defined clearly in the literature, but is increasingly used to describe an array of systems which are described by models with randomly varying inputs. Here the operational rules of the model are defined clearly enough to reconstruct it (with some points of ambiguity-these may be resolved by comparing model output for different assumptions and comparing to the published data) in programmatic form even in the absence of the actual MATLAB code used to produce the model results in the paper.

The primary "stochastic" property of the model described here is the probability distribution consulted by each cell upon dividing, of undergoing a symmetric proliferative (PP), asymmetric terminal (PD), or symmetric terminal (DD) division. The term "consulted" is appropriate, as the explicitly teleological language of "choice" is used repeatedly by the authors to describe the assigning of one of these mitotic modes to the cell. Mitotic mode, of course, is a property that can only be ascribed to the cell by the retrospective analysis of the scientific observer, that is, unless the fate of the offspring are causally specified by the organisation of the mitotic event itself. It does seem strange to think of RPCs in this way, given the paper's invocation of stochastic noise in gene expression as an explanation for this "choice"- it is to reconcile the image of a nucleus buffeted by randomly fluctuating bursts of transcript with the idea of agency. A secondary "stochastic" property seems to be the use of gamma probability distributions to produce a probability of mitosis for each RPC.

The model assumes that the retina starts as a "blank slate" of identical cells in a quiescent state. A wave of proliferation and differentiation sweeps the primordial optic cup, progressively tipping quiescent RPCs into a three-stage descent down the mitotic probability slope, starting with rapid, exclusive PP divisions that expand the RPC population, a second segment of increased DD and PD probability, and a terminal segment of high DD probability. The driving "independent variable" of the model is time in lineage (TiL), which is allowed to increase from zero when the cell is washed over by the "wave" experiencing its first mitosis. TiL advances as a function of chronological time. The second phase begins 8 hours TiL, the third phase 15 hours TiL. These values are said neither to be fine tuned nor to require it to produce biologically plausible results. Percentile probabilities are specified for PP, PD, and DD-type divisions in each phase. The average cell cycle time for these mitotic modes is directly measured in the paper, but the cell cycle time is modelled with the aforementioned gamma probability distribution:

**In the following, we will use the timing of the first mitosis to define the start of the development program within each lineage.** This occurs at around 23 hpf in the central nasal region, reaching the peripheral temporal region around 16 hr later. [] For simplicity, we therefore suppose that RPCs enter their active phase at a uniform rate, expecting that deviations from this will be beyond the resolution of the data.

If we assume that, over the period from 24 to 48 hpf, RPCs are limited to the proliferative phase, measurements of the average clone size over this period suggest a cell cycle time of ca. 6 hr, allowing approximately two rounds of symmetrical cell division. **(Anticipating the results of the live-imaging study, our simulations are actually performed with a shifted gamma distribution, with a refractory period of 4 hr, mean of 6 hr, and width of 1 hr.)**

The choice of the gamma distribution is not spoken to outside of the methods section. Model output

shows clones with more than 16 cells, indicating that the number of divisions is not limited to the number of phases, but the constants to produce the gamma curves are not identified. As the Monte Carlo simulation method is being used, the implication of "simulations are performed with a shifted gamma distribution" is that a gamma curve, with a cumulative probability distribution converging to 1 with as TiL increases, is sampled on some schedule to determine whether a mitosis takes place. A refractory period seems to be modelled, although it is unclear what "width" is referring to here- probably the spread of the probability distribution. 2 years earlier [?] the same statistician did this differently:

The clone size distribution and lineage composition predicted by the stochastic models discussed in the text were determined by numerical simulation using a Monte Carlo analysis based on the Gillespie algorithm (Gillespie, 1977). The balance between proliferation and differentiation, and the specification of the different differentiated cell types, were set by fixed probabilities as defined in the main text. Where necessary, the distribution of cell cycle times was taken to be log-normal with an average and variance fixed by the fit to the experimental data as defined in the main text."

It's not clear if the gamma distribution was experimentally derived from the same kind of empirical fit. The Gillespie paper points out that the Monte Carlo method is from the nonlinear dynamics toolkit, intended to use statistical mechanics to model actual physical processes with dynamics that are known to be adequately described by probability theory, like chemical reaction rates. A significant point for He et al. and for the later application of this model is that no such correspondence to an underlying physical system exists- there is no physically specifiable system which gives rise to the mitotic mode decisions in this paper.

Although the great importance and usefulness of the differential reaction-rate equations approach to chemical kinetics cannot be denied, we should not lose sight of the fact that the physical basis for this approach leaves something to be desired. This approach evidently assumes that the time evolution of a chemically reacting system is both continuous and deterministic. However, the time evolution of a chemically reacting system is not a continuous process, because molecular population levels obviously can change only by discrete integer amounts. Moreover, the time evolution is not a deterministic process either. For, even if we put aside quantum considerations and regard the molecular motions to be governed by the equations of classical mechanics, it is impossible even in principle to predict the exact molecular population levels at some future time unless we take account of the precise positions and velocities of all the molecules in the system. In other words, although the temporal behavior of a chemically reacting system of classical molecules is a deterministic process in the full position-momentum phase space of the system, it is not a deterministic process in the N-dimensional subspace of the species population numbers, as (1) implies.

In many cases of course the time evolution of a chemically reacting system can, to a very acceptable degree of accuracy, be treated as a continuous, deterministic process. However this should not always be taken for granted, especially now that the attention of chemical kineticists is increasingly being drawn to the study of ecological systems, microscopic biological systems, and nonlinear systems driven to conditions of chemical instability. In some cases like these, the inability of the reaction-rate equations to describe the fluctuations in the molecular population levels can be a serious shortcoming. Moreover, contrary to widespread belief, it is not even guaranteed that the reaction-rate equations will provide a sufficiently accurate account of the average molecular population levels; for, except for very simple linear systems, the average molecular population levels will not exactly satisfy any closed system of equations...

It is not obvious these are descriptions of the same method of modelling RPCs proliferation in any case. He et al. seems to indicate that a gamma probability distribution is sampled for some (cumulative?) probability number over time. Gomes et al.'s probability function is derived from a fit of the cell cycle distribution data, so the output specifies a particular cell cycle length from a log-normally distributed at each mitotic event, and does not poll a probability function between these events.

This is the first significant point of methodological ambiguity. The model is not further explained in the supplementary material, so we are left with Fig 4D and the descriptions quoted to determine how cell cycle is modelled. It is unclear whether the mitotic probability distribution allows cells to not mitose at all, or whether all cells are guaranteed to mitose with an average time of 6 hours. If it is the latter, it is unlikely that this "stochastic factor" actually needs to be modelled probabilistically at all- having the cells divide every 6 hours will do. If it is the former possibility, and cells simply encounter a "pulse" of mitosis probability, there is a serious problem here- 4D depicts a decreasing peak mitotic probability for each of 4 subsequent generations. Although none of these probability functions are defined, if cells may pass through this phase of TiL without experiencing a mitosis, the output of the model may be arbitrarily damped by the declining peak of the mitotic probability distribution. Doing so may have made the model fit trivial.

An interesting, more recent Harris modelling paper Boije et al. [?] proposes a model that substantially improves on He et al.'s failure to connect their choice of statistical dynamics to a physical system known to exhibit stochastic behaviour, suggesting that random fluctuations in known macromolecular species (Atoh7 and ptf1a) result in the observed fate specification distribution for zebrafish RPCs. The model used is, by contrast to He et al., rigorously specified. It states explicitly:

Specifically, we supposed that nodal TFs (Atoh7 and Ptf1a) are independently expressed such that, in each generation of the lineage, cells express Ptf1a and Atoh7 with specified probabilities,  $p_{Ptf1a}$  and  $p_{Atoh7}$ . According to their expression levels, cells fall into four groups: cells expressing Atoh7, but not Ptf1a, undergo an asymmetric (PD) division generating a GC. Cells expressing Ptf1a, but not Atoh7, undergo a terminal (DD) division resulting in two differentiated cells of type AC or HC. RPCs expressing both TFs divide asymmetrically producing either an AC or HC. Finally, RPCs lacking these TFs divide symmetrically, either undergoing a proliferative (PP) division or a DD division resulting in two late-born neurons, i.e. PR or BCs. The fraction of non-proliferative divisions creating late born neurons (given that they do not express Ptf1a or Atoh7) represents a third and final adjustable parameter in our model,  $p_{ng}$ .

...

Motivated by these investigations, we therefore considered three temporally distinct regimes of TF expression: In the first three generations of a lineage, none of the differentiation factors are expressed. Then, for two rounds of division, all of these factors can be expressed, while after that only the parameter  $p_{ng}$  remains non-zero, cf. Fig. 2B. Taking the time dependence of TF expression into account, the probabilities of division modes,  $p_{PP}$ ,  $p_{PD}$ , and  $p_{DD}$  follow closely the time dependencies proposed in recent studies (He et al., 2012).

The second highlight is interesting. 3 divisions in the first phase of He et al.'s model would be either unlikely or impossible depending on how the model is programmed. In the third phase of He's model, PP divisions are still allowed, alongside DD divisions, at least opening the possibility that one of the cells could be an "eternal stem" cell, continuing to proliferate through the life of the organism. In Boije's third phase, only DD divisions are possible.

A central claim of He et al. is:

How CNS structures, like the retina, of predictable sizes and cellular compositions arise from such variable lineages is a major unresolved question in developmental neuroscience. **The variability of clones is an intrinsic cellular feature of RPCs**

There are actually two propositions here: the variability of clonal size is intrinsic, and the variability of clonal composition is intrinsic. These are distinct attributes of the clones. What is perhaps most interesting about the progression from He et al. to Boije et al., the Harris group has simultaneously shifted emphasis to the modelling of fate distribution rather than clonal size, while intrinsically limiting clonal size by the model specifications. Given what was noted above, that He et al.'s mitotic model is either trivial (cells are guaranteed to mitose regularly until a DD event) or its effect on the model substantially specifies the output of the clonal size simulations, and that Boije et al.'s model directly specifies clonal size by limiting the total number of proliferative divisions, a suspicion arises that arbitrary selection of model parameters by the experimenters may have strongly influenced the model's fit to the experimental data.

Beyond this, since our data and simulations suggest that the He et al. model cannot explain the 2nd phase of proliferation as the fish move into their first month of life, it seems very unlikely that there is not a substantial extracellular influence on either cell cycle length or mitotic mode probability distributions. It would be interesting to compare retinal and tectal proliferative activity during the juvenile period of massive CMZ expansion in the same fish. It would be difficult to explain well-correlated CMZ tectal cell production without some supervening global cause or coordination between the two niches, either of which would disprove the notion that intrinsic clonal variability in proliferative activity is the primary determinant of retinal cell number.

Both the statistical models of He et al. and of Boije et al. are implementable in NetLogo as agent models in 2d or 3d space. Some assumptions have to be made about the way cells divide, move about in space, etc, but it should be possible to check the output of the Monte Carlo models by these methods, both against our data and against their own reported data, with some additional model constraints implied by the spatial configuration of the cells. This can even be done in 3 dimensions for the embryonic eye with 2000-10000 cells. It seems likely that it can be demonstrated that He's model can be adapted to produce a good fit to observations by reducing cell cycle behaviour to a deterministic parameter. This would clarify that the important stochastic factor is the probability distribution of observed mitotic mode.

If the mitotic mode dynamic can be reduced to a statistical simulation of fluctuations of expression in two genes, and this is the main bona fide influence on retinal development being documented in these papers, it is likely that the entire cell cycle regulatory apparatus lays outside this "circuit". It further seems unlikely that it has been severed from its regular means of correspondence with the extracellular space. If this were the case, it would be difficult to understand our result showing that PCNA+ve CMZ cells in 10dpf rys sibs are less than 25

From the perspective of the original aim of our study, this is a promising possibility, as we would ideally like to be able to impose extracellular influences on peripheral retinal RPCs to promote a second, post-embryonic phase of proliferation, with fate specification being carried out by an independent, intrinsic mechanism which will reliably reproduce the distribution of cells present in the adult retina.

With our evidence we will be able to show that (a) a model based on the intrinsic dynamics observed in the embryonic and early postembryonic life of the fish cannot explain juvenile adult neurogenesis, and that (b) this "second wave" of retinal generation produces the majority of the adult retina in zebrafish, and that therefore (c) the intrinsic stochastic RPC dynamics may be mainly related to fate specification and not to proliferative dynamics. Recently some xenopus data was published with a similarly-timed "pause" [] before a proliferative expansion, which may suggest that this peripheral RPC response may not be idiosyncratic or confined to fish.

The He et al. lineage commitment data is ambiguous. Particularly, Fig 5C's lineage trees do not indicate TiL until differentiation, instead terminating at the division node that produced differentiated cells. He et al. and Boije et al. strongly suggest that RPC lineage commitment is driven by internal fluctuations in transcription factor expression. This may well be the case, but these dynamics alone cannot explain the observed three dimensional structure of the retina. Since cells need to be integrated into the correct retinal layers, and apical division is obligatory (this claim is explicitly made in a later Harris paper about the postembryonic CMZ[]), He et al. and Boije et al.'s suggestion that RGCs are specified by early PD divisions is interesting. If He et al. are not modelling an "eternal" stem cell which may throw off progeny with TiL set to zero no matter what the model time parameter is, it is impossible to see how RGCs could be generated by long-TiL neurons. Equally interesting is that if we model such a system, PD divisions would occur most frequently adjacent to the RGC layer, soon after the new progeny emerge from the stem cell. This raises the possibility that abstracting the spatial organisation of the niche has obscured extracellular influences on fate specification or even on "mitotic mode".

## SUBSTANTIVE ANALYSIS

The He et al. paper is substantially built around the claim that intrinsic stochastic effects cleanly explain for the formation of the embryonic zebrafish retina. They have subsequently gone on to make the same claim for the postembryonic retina. Their account of the significance of this claim and its ostensible alternatives is unclear. We know that this is a Systems Biology paper because a model of a dynamical system is being simulated in MATLAB. It seems obvious this model produces a reasonably good fit to the observed data. That the stochastic model produces a good fit with data from the observed system is here construed both as evidence for the sufficiency of the modelled dynamics for the formation of the eye and as evidence for the "stochastic effects" which are purported to underlay them. Of course, for any real effect, the definition of its stochasticity is that it is well-modelled by probabilistic statistical mechanics means that one is not making an interesting *ontological* claim about something by describing it as "stochastic". One is, at best, making a novel *epistemological* claim.

He et al. gestures toward, and Boije et al. is built around the premise that interacting stochastic fluctuations in gene transcript and/or product expression control fate commitment. This is a possible explanation for the observed variability both clone size and composition. A notable feature of the statistical models of both papers is that they do not include any specifications relating to the fluctuations which are purported to underlay the behaviour. In He et al., a state machine model steps through three constant probability sets for mitotic mode. In Boije et al., cells have a constant, independent probability of expressing either transcription factor, for a brief period of time. A fluctuation is a dynamic that takes place within a particular entity over a period of time. Gene transcription is subject to wildly fluctuating bursts of activity. Indeed, there is significant evidence for pervasive transcription of the genome, with all its implications for informational, transcriptional, and translational noise "noise". If this "noise" is, indeed, an evolutionarily tunable property of RPC specification genes, this is a plausible dynamic by

which cells might be destabilised from one equilibrium state (progenitor) into another (differentiated cell), as if rolled over the lip of a proliferative cradle onto the canalised specification downslope. There remains the question of whether the fit of observed data to either of these models is good evidence for the involvement of such dynamics.

Part of the problem seems to be the desire to stake out a Grand Unified Model Of Retinal Formation by ascertaining the generator of apparently "random" outcomes. The idea proceeds from the observation that cell fate outcomes in vitro have tended to be similar to those measured in vivo- extracellular influences are not required. Sakai's paper [12] is interesting from this perspective, since his optic cups even got as far as organising into layers, but seemed to stall out at that point. In any case, the structure of the retina is clearly "self organised" in at least some way. But an account consisting of the numbers of cells produced and their fate over time cannot reasonably be said to be a full account of a formation of a CNS structure without any acknowledgement of its structure, surely.

This theory is mostly counterposed to what seems to be a strawperson Connie Cepko, who is according the ignominy of originating the foolish theory of Histogenesis By Deterministic, Linear, Unsynchronised RPCs. In this story, it was once thought (presumably by those with insufficient grounding in Nonlinear Dynamics), that the observed birth order of retinal neurons (notably, early RGC birth) was explained by neurons that progressed through a linear sequence of well-defined competency states, each of which enabled the RPC to respond to some extracellular signal that would result in it differentiating to a particular fate. There is a straightforward conceptual error here: "unsynchronized" implies that the RPCs would start out distributed along their linear trajectory through their Differentiation State Space, with some variance. One would want a statistical theory about the nature of this distribution before building a model of it. One might then declare that the first theory to commit to the idea that lineage commitment might involve stochastic phenomena was Deterministic, Linear, Unsynchronized RPCs. Again, the *ontological* significance of the claim about "stochasticity" is unclear. It has been universally accepted for at least a century that biological systems are well-described as populations by statistical methods.

Any number of physical phenomena produce outcomes that, in aggregate, are well described by statistical methods. Nonlinear dynamics techniques have been created to handle complex systems that have population values that are small enough to be heavily subject to unpredictable fluctuation within some typical variance. Nonlinear dynamics are an interesting explanatory candidate for the apparently random behaviour of stem cells, as these models often exhibit multiple, well-separated equilibri in the state space. If the lineage state of the cell depends on its internal composition of fluctuating levels of macromolecules, the dynamics describing their interactions could give rise to stable "wells" (corresponding to various proliferative and differentiated states). Unstable proliferative states may tend to eventually give way to very stable differentiated states. No such behaviour has been demonstrated for the models presented.

The shift from a general model of "mitotic mode" of He et al. to the later "transcription factor fluctuations" of Boije et al. is an admirable attempt to physically ground a statistical model of RPC function, but the dynamics of the model betray the same confusion about what kind of phenomenon is being documented. In both cases the significant model input parameters are either directly specified mitotic mode probability, or indirectly specify mitotic mode by their own interacting binary states. Mitotic mode is, notably, a feature of a cellular division which depends for its identification on events after the division. This conceptual structure implies that the organisation of the division itself in some way specifies cell fate- and this appears to be supported by the observation that RGCs are produced

mainly by PD divisions. Specification by asymmetrical segregation of cellular contents during division has been known for some time, so this is plausible.

Because "mitotic mode" is too general a trait to represent a valid target for neodarwinian synthetic molecular evolution, Harris makes an attempt to introduce transcription factor fluctuations, which have an understood physical basis. "Mitotic mode" is nevertheless smuggled back in as the output of the interaction of the transcription factors, which are either expressed or not expressed with some defined, constant probability for some period of time. As noted, the model does not have any account of the dynamics of the interacting fluctuations which are said to give rise to the PP/PD/DD "choice", so this is merely the addition of a smaller black box on the first one, which goes on to produce output with similar distributions to the observed data.

The "choice" language and selection of a behaviour (mitotic mode) rather than a bona fide fluctuating macromolecular network as the modelling subject are products of the irreducible teleology implicit in most molecular accounts of cellular function. Harris would likely maintain that this is metaphorical language which can be reduced to a legitimately ateleological physicalist description of "stochastically" interacting macromolecules. This is probably not correct. As Jesper Hoffmeyer points out:

One reason why biologists, and scientists in general, so vehemently oppose any claim to the effect that natural selection might not be a sufficient explanation for evolution on Earth, is the belief that the theory of natural selection is the very element that glues biology to materialistic science. The disquieting fact that the creatures of this world so clearly exhibit purposeful behavior has always tempted philosophically minded biologists to claim the existence of peculiar vital forces or principles pertaining to life. It is probably no longer possible to make a career in biology at the university level if you adhere to such vitalist ideas, but biologists still have to somehow cope with the obvious, though tabooed, teleological aspect of life. The prime mode in which this is done in modern biology is by the application of the concept of function. Volumes have been produced to solve the problem of how to justify the concept of function inside a non-teleological frame of understanding. And this is where natural selection comes in, for natural selection will tend to optimize the capacity of species to meet the functional challenges of their ecological niche conditions. Functionality is exactly what natural selection is supposed to produce.

The term "function" in biology is understood as the answer to a question about why some object or process has evolved in a system. In other words, what is it good for? A function thus refers forward in time from the object or process, along some chain of causation to the goal or success. This inversed arrow of time (future directedness) immediately sets functions apart from other kinds of mechanisms that always refer backward along some chain of causation explaining how the feature occurred. Darwinists, however, are not worried about the teleological character of functions because they believe that natural selection will ultimately account for them through ordinary mechanistic causation. Thus, as often noted by Darwinists, adaptive traits are not explained by the consequences they will or can have but by the consequences they already have had in ancestor populations. The consequences in other words precede the effect they explain, and selection does not therefore challenge the mechanistic paradigm of traditional biology. So, as the explanation goes, the teleology implied by the concept of function is only an "as if" teleology, i.e. a teleonomy.

But something feels wrong with this argument. There is of course no reason to believe that the flowering apple tree outside my window right now has any purpose in producing flowers. Trees do not have purposes in the human sense of the term. But clearly even apple trees exhibit some kind of agency in the sense that they have a capacity to generate end-directed behaviors. The apple flowers serve to attract bees that can pollinate them and thus assure the occurrence of sexual reproduction. But how can a purely mechanical principle



(natural selection) produce agency in a world that is not supposed to carry any trace of end-directedness? Our intuitive sense of logic seems to be strained quite a bit here.

... Life is historical in the sense that its continuation depends on an ability to learn: Strategies that have proved effective in overcoming past challenges must be "remembered" so that descendent organisms will be able to cope with those same challenges; and - just as important - those strategies that did not work well must be "forgotten". The effect of these two processes is learning, and since natural selection is precisely a mechanism for remembering the fit and forgetting the unfit natural selection is a kind of learning process. Now, learning processes are different from normal mechanical processes in that they depend on the formation of some form of a coded representation (the functional response must be incorporated in the system in some way or other, for instance as here in the form of the genetic setup). But the moment a representation becomes a constituent of a mechanism "misunderstandings" will necessarily lurk in the background. An unpredictable source of change is thereby introduced into the system and change in systems with a capacity for learning is therefore historical in nature, not mechanical.

Learning is of course a semiotic process, an interpretative agency must be a part of the process, and in natural selection this interpretative agency resides in the lineage, i.e., the population as an evolving unit. Populations are coterminous with the concrete individuals living here and now, while the lineage is a historical and transgenerational subject that possess a collective agency as such, since its destiny as a temporal integrative structure is formed through the accumulated interactions of its single units (individuals) with the ir environmentsmuch in the same way that multicellular organisms are integrative structures interacting with their surroundings via the activity of individual cells (Hoffmeyer 2008, 116). The lineage is the subject that may eventually learn to cope with changes in the ecological niche conditions or, otherwise, go extinct. According to present fashion in biology this learning process can be measured simply as changes in gene frequencies down through the generations. Each generation in this simplified optic constitutes a cohort of genotypes expressing themselves as phenotypes that enter into the competitive acts of survival and reproduction. The resulting cohort of genotypes in the next generation will thus reflect the failures and successes of the preceding cohort (i.e., what was "remembered" and what was "forgotten").

In this light, we can see both the "function-problem" of the Harris papers as well as a significant epistemological alternative. The Harris models are statistical descriptions of a population of cells "making decisions" about a "future-directed" process- the differentiative outcomes of a division. The cell has a kind of "agency" in determining the latter by the former. However, Harris cells are a kind of autistic locked in a room, occasionally consulting astrological tables to make decisions about mitosis and fate. They are slaves to simple internal probabilistic dynamics, and do not take extracellular information into account in wending their way along their lineage trajectories as they age. It remains unclear how the segregation of cell types into defined lamina occurs, or what the physical correspondent is of the model parameter that "ages" the cell along the mitotic mode probability curves (in other words, why does mitotic mode go through phases)?

Feyerabend describes scientists as epistemological anarchists who, contrary to popular belief, do not hew to any stable, well defined method, but instead use a variety of means to explore the world and generate systematic descriptions of it. That is, they are not Popperian falsification robots, but historically situated humans who must cope with the plodding incrementalism of Kuhn's "normal science" or the chaotic proliferation of new theories before one is stabilised during a Kuhnian "paradigm crisis" before one of the competing host is selected as the new dominant paradigm and normal science resumes. It seems clear the Harris and Cayouette labs are engaged in exactly this kind of activity. As Feyerabend



points out, what is science and what is not is determined only in retrospect, by people evaluating what "worked" and what did not, in retrospect, using some system of values or axioms. Their efforts represent the kind of proliferating attempts to offer new types of explanations that occur during the last portion of a paradigm's life, when everyone is aware of insurmountable evidentiary obstacles to establishing the paradigm's credibility, but as no viable systematic alternative has arrived, one must nevertheless pay obeisance to the dying sovereign. However, without a clear idea of what kind of answer is being sought and to what end, HC have produced an explanation for retinal formation which is neither intellectually satisfying (being based on dubious theoretical ground and contradicted by observation) nor technically useful (providing no tractable means to modify the system).

## COMMENTARY ON HE ET AL. TEXT AS READ

### INTRODUCTION

**Most neurons are not replaced during the lifetime of the animal. Each neural progenitor, therefore, must generate a finite clone of neurons, and all these clones together must add up to the full complement of neurons in the mature nervous system.**

- "must" actually implies: No neurons are replaced

- unclear sense of "replace". There are probably degrees of dynamism but we are probably not thinking of cavefish embryonic remnants collapsing and CMZ output "replacing" those [13]. In any case the stability of the retina depends to some extent on active genetic regulation- [10] is probably an unrelated example of this, so there are in any case obviously different modes and degrees of retinal instability.

- one reason to care about retinal stability is implicit proposition that retinal stability is required for function. cellular "patch" therapy concepts like [12] suggests rely on assumption of extensive retinal plasticity on fairly short time scales

The clonal basis of vertebrate central nervous system(CNS)development has been investigated in detail in the retina, which develops from the optic cup, an outpocketing of the forebrain. The neuro- epithelial layer of the optic cup is composed of retinal progenitor cells (RPCs) that first undergo a period of cellular proliferation, followed by a phase in which cells progressively exit the cell cycle. **Individual RPCs are multipotent, giving rise to all retinal subtypes (Cepko et al., 1996; Holt et al., 1988; Turner and Cepko, 1987; Wetts and Fraser, 1988).**

- Individual RPCs are multipotent, RPCs as a population *do* and as individuals *may* give rise to all retinal cell subtypes

- **Cepko 96[4]**: review, produces model of progression of competency through series of well-defined phases in which individual cells are receptive to environmental stimuli to produce one of an ordered set of lineages. Intended to explain temporal birth order in development through formation of embryonic eye to apparently static tissue in rodents

- **Holt 88[9]**: Early xenopus BrdU and clonal analysis: suggests "a mechanism whereby determination is arbitrarily and independently assigned to postmitotic cells" after radial extension of progeny along full retinal column "as lamination begins"

-**Turner 1987**[17]: early clonal labelling expts, concerned w/ relationship between neurons and glia. interesting numerical fate table data for mice that may not have been compared to zebrafish and xenopus [14]

-**Wetts 1988**[19]: Science paper, lineage tracing of xenopus retina w/ dextran iontophoretic injection (!) variety of clonal outcomes taken to support questions about multipotency of retinal progenitors

**In addition, clones derived from single RPCs, in a number of vertebrate species, exhibit enormous variability in both size and composition (Fekete et al., 1994; Harris, 1997; Turner and Cepko, 1987; Turner et al., 1990; Wetts and Fraser, 1988).**

-**Fekete 94** [6]: retroviral clonal lineage analysis of chicken RPCs: analysis of clone size and dispersion. Some nice early spatial analyses. *Worth comparing radial clonal distributions to later Arcos evidence*

-**Harris 97** [8]: This is probably Harris' earliest attempt to formally explicate a model of the generation of retinal order from observed variable clonal behaviour. Elaborates idea of developmental time (time in lineage) as *conceptually separate* from mitotic activity. Accepts influence of extracellular influences, neuronal coordination- more diverse view of factors contributing to generation of order than the recent work, which wants to substantially simplify these away. *worth reading closely enough to produce brief remark on the evolution of Harris' views- what evidence is prioritised and what is excluded as unimportant, anomalous, or adequately explained by other factors*

-**Turner 90** [16]: More retroviral labelling experiments, demonstrating symmetrical proliferative (PP) divisions among RPCs. Calls PD-type asymmetric divisions "stem cell division pattern". "We propose a model for the generation of retinal cell types in which the cessation of mitosis and cell type determination are independent events, controlled by environmental interactions."

**How CNS structures, like the retina, of predictable sizes and cellular compositions arise from such variable lineages is a major unresolved question in developmental neuroscience. The variability of clones is an intrinsic cellular feature of RPCs (Cayouette et al., 2003).**

-*central claim.* There are at least two kinds of variability ascribed to intrinsic factors by this claim: clonal size and composition

-**Cayouette 03** [3]: rat e16-17 explants have similar fate distribution as in vivo RPCs- suggests clonal composition does not depend on extrinsic factors

**This is known because isolated rat RPCs grown in vitro produce clones of various sizes and compositions. Yet, surprisingly, when examined as a population, these isolated clones are statistically similar both in size and composition to those induced in explants. As there are few extracellular influences on isolated RPCs, these results suggest that proliferation and cell fate choice are primarily determined by cell autonomous influences, such as transcription factors and components of the cell cycle (Agathocleous and Harris, 2009)**

-Fates are distributed the same way for individual progenitors and their progeny in explants as in situ. It follows that patterning arises from some mechanism for correctly positioning the variable outcomes of the described process without any direct connection between fate and position- this would rule out

adjacent cells signalling re: apicobasal coordinate to the differentiating cells. if this is the case no need for extracellular input when modelling differentiation, cells can be thought of as, at differentiation, being sorted to the correct locations. Can compare with current (arbitrary physical displacement) model by implementing clonal analysis in agent model.

-”transcription factors” and ”components of the cell cycle” here to be refer to differentiation and proliferation behaviour as conceptually separate mechanisms

-**Agathocleous 09 [1]**: Review describing retinal proliferation and differentiation as interrelated mechanisms governed by a series of interacting ”functional modules”. /textitGood example of attempt to introduce evodevo ”module” functions into model descriptions- and limitations of concept (arbitrary siloing of molecules into ”functional groups” predefined based on some higher-level teleological scheme (perhaps as opposed to observed physical interaction) Review questions the importance of extracellular mechanisms. Unclear separation of mitotic exit from lineage commitment, eg.:

”Apart from the decisions to divide or not at G1, the duration of the cell cycle also regulates differentiation. In general, late progenitors divide more slowly than early progenitors (Alexiades and Cepko 1996), perhaps owing to an accumulation of CKIs and a reduction of cyclin activity, which would also bias them to exit the cell cycle.”

**What remains both controversial and unresolved, however, is whether individual RPCs use these factors within a variety of stereotyped programmed lineages or whether stochastic influences govern the expression of these factors within a population of essentially equipotent RPCs.**

-”these factors”- cell autonomous influences

-I either don’t understand the formulation or disagree that this is the ”controversial and unresolved issue” being contested by the model in the paper

- ”variety of stereotyped programmed lineages” may mean something like ”progenitor proceeds through a linear series of competency states as function of some intracellular phenomenon as it develops through time” as per Cepko [4]. Cepko’s model suggests that progenitors distributed in some fashion across this series (perhaps with their position along the ”time coordinate” (roughly: time in lineage, TiL below) of the series set at ”random”- to what extent is this already a stochastic theory?) nevertheless as individuals proceed linearly through it, perhaps at some fixed rate. ”Competence” means ”competent to respond to extracellular signals” here. ”variety of stereotyped programmed lineages” may also refer to some more sophisticated conception of commitment to stereotyped ”branches” on some differentiation landscape- maybe fates are hierarchically clustered, maybe not (”transcription factor tree”-type differentiation landscape implied by derepression model in Vitorino/Harris paper [18]).

-Harris takes the notion that there could be extracellular signals as resolved by the explant data (cf. my conclusion about Sasai’s paper in introductory anecdote). The remaining part to be disputed is whether some (linear?) time-dependent sequence of fate distributions is displayed. *Need to look at early retinal formation- any reason peripheral neurons at the end of central phase of differentiation biased to one side of distribution of TiL values? ie might the apparent sequence of competence states actually reflect a bias in the distribution toward late-born neurons that develops as the central phase ends? Obviously in the post-central context, the progenitor population must still be able to produce RGCs, so TiL values for CMZ cannot exclude early fates- model would have to account for this if so*

-a question for both models is how cells are distributed across the TiL in the "starting condition" and how this relates to the "stem cell" concept. The agent model currently initialises lineage ages in all progeny of an immortal, always-peripheral stem cell with a value of 0. In any case, both models can only be sustained by some mechanism to generate "young" progenitor cells with low TiL. Seems clear that in both cases the conceptual distinction between stem and progenitor cells is that stem cells are either in some separate multipotent process-state where TiL is not a relevant input, or "locked" at the very beginning of the progenitor process-state ( $TiL = 0$ ). could be thought of as an attractor state with a large activation barrier to the canalised differentiatory downslope- if the attractor state also reliably segregates the contents of the cells in a way that always produces a "stem" cell, this state could self-promote or stabilise without precluding the loss of the stem cell from the lineage completely from internal fluctuations that push it over the barrier. Model may therefore not need built-in assumption of "stem cell immortality" to function, but an "assemblage of arcs" value would be necessary to ask whether peripheral stem cell "extinction" events are happening at a reasonable pace- presumably stem cells contribution to adjacent lineages take over the work of extinguished stem cells *Need to check Arcos evidence for clear indications that lineage extinction is happening*

-the *explicitly stochastic* part of the model developed in the paper refers to the statistical description of the frequency of PP, PD, DD -type divisions- where "postmitotic" is taken to be equivalent to "differentiation"- does *NOT* refer to the distribution of cell fates. the contrast drawn here with Cepko's model is actually with the implications of the lineage tracing data. Cepko's model doesn't care about asymmetry of proliferative behaviour. "transcription factor tree"-type model [18] doesn't suggest mitosis is the timekeeper for derepression either.

**In support of the former hypothesis, several studies have shown that RPCs exhibit cell-to-cell variability in both gene expression pattern and cell fate potential** (Alexiades and Cepko, 1997; Dyer and Cepko, 2001; Jasoni and Reh, 1996; Trimarchi et al., 2008; Zhang et al., 2003).

- "former hypothesis" is that there is some stereotyped program of lineage commitment

- **Alexiades 97 [2]**: rat in vivo and explant pulse-chase/marker study evidence indicating that distribution of marker expression on BrdU-labelled cohorts changes over time

- **Dyer 01 [5]**: p27Kip1 and p57Kip2 effects on mitotic behaviour in mouse. No big effects on fate distribution (can perturb rod proportion?), fate variability over time not considered

- **Jasoni 96 [11]**: MASH-1 expression restricted to later progenitors, expressed heterogeneously, suggested correlate of differential lineage commitment

- **Trimarchi 08 [15]**: enumeration of single cell, Affymetrix array-derived, transcript heterogeneity measurements for RPCs. Provides evidence of expression clustering of early vs late transcripts. Provides evidence for stochastic variability of transcript expression. Arbitrary clustering into transcription expression "subsets". "Comparing the gene expression profiles of E12.5 RPCs and P0 RPCs by visual inspection in Microsoft Excel, however, did reveal several candidate genes whose expression appeared mainly confined to early RPCs"- literally totally arbitrary. Arbitrary clustering into cell cycle expression groups. *May be worth investigating if clustering can be performed by unbiased search of all cells and not merely analysis that assumes the conclusion. Good example of limits of enumerative -omics projects- may provide information about sources of variability in cell fate, but cannot generate an explanatory framework for tissue-level organisation*

- **Zhang 03 [20]**: Pax6 enhancer sequences promoting expression in retinal and pancreatic lineages identified. Pax6 expression associated with horizontal cells in retina.

-evidence of progression through some temporal program of differentiation only really directly provided by Alexiades, Jasoni, Trimarchi here. Structure of the evidence presented generally just assumes prior earlier model of linear progression through temporally staged differentiation pattern- Harris has some grounds to challenge this assumption here but really the earlier type of birthdating studies are more important for establishing this than studies demonstrating variability in gene expression or cell fate potential. *Harris' confusion about what is being disputed seems evident here*

**However, a recent careful statistical analysis of a set of late progenitors from the rat retina cultured at clonal density and followed in time lapse so that every division was mapped supports the latter point of view. In this study, it was revealed that the variable clone size distribution was consistent with a simple and well-constrained stochastic model in which cells were equipotent but had certain probabilities of dividing and differentiating (Gomes et al., 2011).**

- **Gomes 11 [7]**: Rat single cell explant study. "Quantitative analysis of the reconstructed lineages showed that the mode of division of RPCs is strikingly consistent with a simple stochastic pattern of behavior in which the decision to multiply or differentiate is set by fixed probabilities. The variability seen in the composition and order of cell type genesis within clones is well described by assuming that each of the four different retinal cell types generated at this stage is chosen stochastically by differentiating neurons, with relative probabilities of each type set by their abundance in the mature retina." Follows from this that both composition and clonal size is intrinsic.

-chapter 1 should include some background on arbitrary model fits to data- that is, one can \*always\* write some polynomial function to fit some arbitrary set of data on clone sizes, where variables are the probability of the cell performing an symmetric proliferative, symmetric terminal, or asymmetric terminal/proliferative division, and variables depend on time. The model in this paper divides this function into three discontinuous regimes but could just as well be a continuous function.  $\text{probPD} + \text{probDD} + \text{probPP} == 1.0$ , all three variables are constrained to  $0 \leq \text{probXX} \leq 1$ , given in the logical structure of the relationship- each offspring of a division will either go on to experience another division or will not. function to fit these data is unconstrained  $\in [0, 1]$ , so any number of fits can work- need to ask how best to *distinguish between* fits as explanations

**In many parts of the nervous system, including the retina, there is a clear histogenesis, such that some cell types tend to be born before others (Angevine and Sidman, 1961; Livesey and Cepko, 2001; McConnell, 1989; Nawrocki, 1985; Okano and Temple, 2009; Qian et al., 2000; Rapaport et al., 2004).**

-Angevine 61 [?]:

-Livesey 01 [?]:

-McConnell 89 [?]:

-Nawrocki 1985 [?]:

-Okano 2009 [?]:

-Qian 00 [?]:

-Rapaport 04 [?]:

Such histogenesis implies that, as lineages progress, the probabilities of generating distinct cell types change as a function of time or cell division. The widely accepted competence model of retinal development (Livesey and Cepko, 2001) suggests that RPCs pass through a succession of states, possibly owing to the successive expression of a set of temporally coordinated transcription factors. Indeed, homologs of temporally expressed transcription factors that orchestrate lineage progression in *Drosophila* neuroblasts (Doe and Technau, 1993) have recently been found to have similar functions in the vertebrate retina (Elliott et al., 2008).

-the implication is inescapable as far as i can see: there definitely has to be time dependence to fate outcomes to produce staged histogenesis. Explant studies which find that clonal composition is the same in vitro and in vivo *for some time point*, and retrospective clonal analyses which find constant distributions of cell types across the lineages *after all cells in the clone are postmitotic* do not explain staged histogenesis or indicate that fate composition is not time-dependent in early embryonic development. it is not clear whether this early time dependence matters for later CMZ contributions- setup of the laminar structure may be a distinct activity which occurs using raw materials provided by intrinsic "stochastic" propensities determining clonal size and composition

-Doe 93 [?]: -Elliot 08 [?]:

A common feature of retinal histogenesis is a substantial temporal overlap in the time windows for the generation of different cell types. In the competence model, this could be explained if the clones were not fully temporally synchronized. Recent investigations, however, show that branches or sublineages of a main lineage tree give rise to distinct cellular fates at similar or overlapping times (Vitorino et al., 2009).

- hierarchical clustering of fate outcomes into a lineage tree still relies on asynchronous TiL values for RPCs. The conceptual difference here is only in how the canalisation of differentiation is conceptualised. It remains unclear how you would get temporally staged cell types in early differentiation if clonal composition is intrinsic, constant, and time-independent.

Single-cell sequencing studies show that neighboring progenitors at the same stage of development have many differences in their expression of cell determination factors (Trimarchi et al., 2008). These studies suggest an alternative to the competence model in which parallel sublineages may progress side by side and give rise to distinct subsets of neurons at the same time.

-This seems to be stretching the Trimarchi evidence, not sure Cepko would agree with this interpretation. Heterogeneity in RPC expression of cell determination factors is also implied by a linear-state-progression model with asynchronous TiLs

To gain deeper insights into these basic questions of clone size variability, stochasticity versus deterministic programming, and histogenesis at the cellular level, we developed a number of approaches to label single RPCs in zebrafish embryos and to follow these clones

over time in vivo. **Our results provide a complete quantitative description of the generation of a CNS structure in a vertebrate in vivo and show how a combination of stochastic choices and programmatic discrete steps in lineage progression transform a population of equipotent progenitors into a retina with the right number and proportions of neuronal types.** These studies also reveal a surprising insight into the mechanism of early retinal histogenesis.

-”Complete quantitative description of generation of CNS structure” lacks a mechanism that would explain the relative 3 dimensional position of the cells in the structure!! -”combination of stochastic choices and programmatic discrete steps” ie the model steps deterministically through three states with different probabilities of asymmetric/symmetric division based on TiL. Still not really clear that the supposed alternative did not include both stochastic and programmatic elements. Is stochasticity vs determinism actually the issue here? Or is it the correct assignment of cell-intrinsic vs extrinsic processes to particular roles within the formation of the eye? Even if clonal size and composition are intrinsic, one clearly still needs extrinsic coordination to achieve laminar ordering.

## DATA

### Fig 1

-establishes photoconversion labelling and tracking paradigm. MAZe heatshock activation system used to activate Kaede in progenitors- some putatively random subset of cells (function of Cre recombination inefficiency and variability in hsp70l promoter driving Cre) has Cre-lox recombination that activates Gal4-UAS expression system driving Kaede. Selective photoconversion of single Kaede-expressing cells performed by 405nm laser on spinning disc. Demonstrates lineage persistence of photoconverted protein and ability to identify differentiated progeny

### Fig 2

-establishes that labelled clones are representative of early RPCs generally. Average increase in clonal size over time is proportional to the increase in retinal population as a whole, fate outcomes are qualitatively similar

### Fig 3A

-clones labelled at 24 hr and followed out to 48 hr tend to have even numbers of progeny. those followed to 72 hours have a more even distribution of lineages with even and odd numbers of progeny. PP proliferative mode must be favoured early, with PD proliferative mode occurring somewhat later (produces odd numbers of progeny). used to justify first two stages of division choice model

### Fig 3B

-clone size broken out nasotemporally to examine effect of ”nasotemporal mitotic exit wave” seen in 4A- t-test significant result showing larger temporal clones at 72 hr is taken as evidence of ”a relative delay in the developmental program between temporal and nasal parts of the retina” (seems like clear extracellular dependency of the ”proliferative program”, minimally in terms of the ”start time” of clone proliferation?). clones photoconverted at 48, followed to 72hr are mainly 1 or 2 cells large - ostensibly points to DD mode favoured in third phase

### Fig 3C



-photoconverting single cells of clones (subclones) at 32 hr demonstrates that eventual number of progeny of subclones appears to be inversely correlated with the starting size of the clone. this is taken to mean that cells lose proliferative potential depending on TiL in a lineage-independent manner- although TiL is not implicated by these data, number of mitoses is. Note: "However, what is remarkable is that the spread of subclone sizes is large in all cases. For example, subclone sizes from two-cell parent clones are as large as 15 and as small as three." Not really clear how the "lineage-independent" nature of progressive loss of proliferative potential is demonstrated by this result

#### **Fig 4A**

-zFucci mAG-zGem RPCs go postmitotic in a naso-temporal wave over 30-48hpf, demonstration of phenomenon dissected in 3B.

#### **Fig 4B**

-quantitation of zFucci mAG-zGem cells over time in sagittal sections, broken out by "zone" and depth from most peripheral section. Largest number of labelled, proliferating cells moves nasal-;temporal as developmental time increases. Not sure why pre-24hr data is not included if the idea is that 15-24hr is a quiescent phase, followed by a wave of nasotemporal entry into the rapidly proliferating three-phase TiL division/decision mode.

#### **Fig 4C**

-schematized version of N/T entry into the stochastic TiL model

#### **Fig 4D,E**

-schematic presentation of TiL proliferation mode "decision model". Model begins when RPC has it's "first" division after some quiescent phase, three phases are passed through as a function of TiL. probability of mitosis is a function of TiL within each phase, *implied by not explicitly stated that each phase **requires** a mitosis.* probability of each of 3 division types is a function of phase

*-there is not enough information provided to immediately replicate the model's programmatic structure. Need to build single-clone model to ensure that monte carlo runs produce similar results to understand what the model's assumptions were. It is not clear that the represented "mitosis probability" curve is actually modelled or whether it matters*

#### **Fig 4F,G,H**

-monte carlo model output vs experimental observations for clone size given different labelling times

#### **Fig 5A**

-confocal time series demonstrating example of tracking of cells to differentiated lineage types

#### **Fig 5B**

-example lineage tree for data in A)

#### **Fig 5C**

-summary graphic of 60 lineages from clones converted at 32 hpf

#### **Fig 5D**

-summary of mitosis modes over time for nasal/temporal/whole retina and model predictions.

#### **Fig 5E**

-measurements of cell cycle length for different differentiation modes

#### **Fig 5F**

-division time vs division time for sister divisions- indicates near-total synchrony of divisions- reasonable to model phases as mitosis rather than TiL dependent?

**Fig 5G**

-eventual clone size vs eventual clone size for sister divisions- good correlation between eventual clone sizes- sister RPCs have similar proliferative potentials at any given time

**Fig 6A**

-temporally staged histogenesis derived from 5C lineage data

**Fig 6B, 6C, 6D**

-cell fate matrix of sister cells, incl. from progenitors that were themselves the product of PD divisions producing RGCs or ACs

**Fig 6E**

-cell fate by division type -RGCs, ACs produced from PD-type and DD-type divisions, BCs, HCs, PRs from DD divisions only

**Fig 6F**

-cell fate by asymmetry of commitment in DD divisions -RGC-producing DDs are almost always fate-asymmetric, ACs are somewhat, other types are usually fate-symmetric

**Fig 6G**

-schematic of lineage barcode method

**Fig 6H**

-this figure is a mess, what the hell

**Fig 7A, B**

-WT vs Lakritz mutant micrographs showing + in cell number in Ath5 mutants

**Fig 7C, D, E**

-Ath5 mutants have greater retinal cell number, average clone size, and more even-numbered clones, apparently in line with model predictions. *not clear how model has been modified here- text references "lineage reversion" but not clear how this applies to mitosis-type choice- possibly forcing simply converting any PD producing an RGC to a PP?*

**DISCUSSION**

In the late 1980s, newly developed methods of clonal analysis in vivo revealed that retinal progenitors were multipotent (Holt et al., 1988; Turner and Cepko, 1987; Wetts and Fraser, 1988). **This initial insight led to many questions that have still not been resolved, such as (1) why are some clones bigger than others; (2) what are the mechanisms by which clonally related cells choose different fates; and (3) is there a strict order of cell genesis within clones?**

-actually important question is how complexity (diversity of cell types) and order (stereotypic retinal units) arise from variable RPC behaviour. Multipotency of some retinal precursor was a given, just a question of when lineage commitment happens. "What are the mechanisms by which" may as well just read "How do". Note constant implicit teleology in "choice" language. I'm not sure (3) ever really mattered

To address these important questions, it is obviously useful to see full clones grow and differentiate into mature neurons in real time in the CNS in vivo. Until recent improvements in imaging and genetic labeling strategies, however, this has not been possible. Using

a variation of the MAZe strategy (Collins et al., 2010) in combination with 4D microscopy, we have been able to label single progenitors at precise stages and follow their development in time lapse until all their progeny have differentiated into specific neuronal types that we could unambiguously categorize. The variability of clone size and composition, seen here and in all previous retinal studies (Holt et al., 1988; Turner and Cepko, 1987; Turner et al., 1990; Wetts and Fraser, 1988; Wong and Rapaport, 2009), raises a key question about whether RPCs have individually fixed lineage programs, like *Drosophila* CNS neuroblasts, or whether they are a set of equipotent progenitors subject to stochastic influences.

-this is notably NOT the same question as "do RPCs proceed through differential competency states characterised by the potential to respond to extracellular factors, or is lineage specification solely a function of cell-intrinsic processes"

There is good evidence for the heterogeneity of RPCs at neurogenic stages, in particular, in respect to gene expression patterns (Alexiades and Cepko, 1997; Dyer and Cepko, 2001; Jasoni and Reh, 1996; Zhang et al., 2003), and it is possible that these differences account for the variety of lineage outcomes. No experiment can absolutely rule out that the heterogeneity of clones follows from the individual and early specification of RPCs, just as no finite sequence of numbers can be proved to be part of nonrandom series.

-heterogeneity clearly has to account for variety of lineage outcomes. the question is what kind of heterogeneity. if all of the parameters we are measuring are being assayed through macromolecular proxies, it is a given that we are going to see macromolecular heterogeneity and ascribe lineage specification heterogeneity to some aspect of that

Nevertheless, in our data set, the very large variety of clone types, in size, composition, and division pattern, and particularly the variability among subclones and sister clones, seems hard to reconcile with detailed deterministic programming.

-"detailed deterministic" here means "non-variable", which has never been an issue in contention since clone size variability was established early on. "Detailed deterministic programming" specifying exact outcomes of mitosis and differentiation is not actually what Cepko or anyone else has proposed

Most importantly, the data presented here, at least in relation to clone size, are consistent with a very simple and constrained stochastic model operating on equipotent RPCs when tested against every statistical measure.

-"tested against every statistical measure" means "we plotted monte carlo model output over our experimental data and decided it looked good". the model is definitely consistent with the data, but need to know what kind of explanation it offers and what the value of this type of explanation is

One might therefore wish to consider the possibility that many of the molecular differences seen in RPCs may not be programmed but rather are the result of cycling or stochastic fluctuations in gene expression (Elowitz et al., 2002; Hirata et al., 2002; Munsky et al., 2012).

-it's hard to tell here if the authors understand what "stochasticity" means. We are told above that heterogeneity of RPC gene expression patterns is not what explains clonal variability, here it is explicitly invoked to explain it. "Stochastic fluctuations in gene expression" are an explanation for observed heterogeneity. This is clearly implied in Trimarchi's work, no reason to do single cell transcript analysis if you don't think this is a reality

Similar models of stochastic proliferation have been very successful at predicting the lineages of progenitors in homeo- static self-renewing adult tissues in vivo (Clayton et al., 2007; Klein and Simons, 2011). A recent analysis of clones generated in vitro from late-stage rat RPCs shows that simple stochastic rules similar to those uncovered here, but with different probabilities of PP:PD:DD, are very powerful in predicting the size distribution of these clones (Gomes et al., 2011).

While this model provides an excellent fit with clone size distributions seen in the zebrafish retina in vivo, it was designed specifically for clone size rather than cell fate distributions. The data set we have is simply not sufficient to allow us to generate a useful model of cell-type distributions within clones, although in the future, with advances in imaging, this should become possible.

-this appears to be boilerplate. If the tracked clones were a sufficiently large dataset to produce the histogenesis diagram in 6A, how will adding more data improve the ability of the model to address histogenesis? *The model can't fully explain cell fate distribution because cell fate distribution is only partially dependent on mitotic staging.* RGCs and to a lesser extent ACs are favoured by early PD divisions, probably rely more on asymmetric segregation of cell contents to stabilise those differentiation states.

While the variability of clonal compositions generated by sister RPCs strongly suggests that there are likely to be stochastic elements at work in terms of fate assignment, **there are also several clear trends in the data that show cell fate determination is unlikely to be purely stochastic.** For example, the frequency of same-type pairs of PRs, HCs, BCs, and ACs is much higher than one would predict from a purely stochastic model, as is the probability that the sister of an RGC will be a P cell.

-What the hell does "purely stochastic" mean? A population variable is either stochastic or it isn't. Maybe this means "cell fate decisions aren't evenly and randomly distributed across lineages" but no one ever suggested they were- histogenesis clearly implies that they are not

- "stochastic" seems to just be functioning as a poorly understood shibboleth to connect statistical model to sexy stochastic-influences-on-gene-expression work

A pervasive feature of the development of many CNS tissues is histogenesis, the general ordering of cell type by birthdate. For example, the cerebral cortex famously shows an inside-out histogenesis, and this order of cell birth is intrinsic to progenitors, as when grown at clonal density in vitro, they give rise to clones in which there is a distinct general

order of cell-type production (Qian et al., 2000). However, it is unknown why layer VI cells exit the cell cycle before layer V cells, etc. Similarly, in the retina, RGCs are born first in a variety of vertebrate species. Why should this be so? Previous studies have provided important hints about these questions by showing that temporal identity genes, homologous to those identified in *Drosophila* neuroblasts, might also act as fate-biasing factors in RPCs to increase the probability of adopting certain fates associated to a particular temporal window (Elliott et al., 2008), but such genes have not been shown to cause early cell cycle exit. Other studies show that some cell-type determination factors may also lead to cell cycle exit and vice versa (Ohnuma et al., 1999, 2001), but their timing of expression does clearly coincide with cell birthdate. It is therefore challenging to ascertain how these factors work within the context of histogenesis, especially when stochastic mechanisms appear to influence cell cycle exit and fate choice.

-We have to know what kind of explanation we are trying to offer before assembling the data. Simple, qualitative, linear models (signal x to transcription factor y to cell exit fate decision) and models incorporating stochastic variability are incompatible. They also have different intents- linear model deals with abstract generic cell, stochastic models are dealing with population level phenomena. That a stochastic model fits a population doesn't mean that the causal structure of events governing the behaviour of individuals within that population is not "deterministic" and a function of a limited number of macromolecular constituents acting in some qualitatively describable fashion

The finding that *Ath5*, already known to be essential for RGC cell fate, is also involved in early PD divisions leading to cell cycle exit at the initiation of retinal clones thus sheds mechanistic insight into how histogenesis can be accomplished within a stochastic system.

-This seems like it requires a lot more explanation than what's offered. Histogenesis is "explained" by the model to the extent that we associate early RGC specification with PD-type events, which are front-loaded relative to DD events. *Ath5* mutants seem to have fewer PD-type events. The coupling of mitotic mode to cell fate specification is what's interesting here, so it's frustrating that these processes are alternately treated as distinct or linked depending on context

**In summary, we have shown that the generation of the zebra- fish retina can be accurately described by a combination of stochastic and programmatic decisions taken by a population of equipotent RPCs.**

-they're not actually equipotent if mitotic mode results in more cells of a particular fate and probability of that mode changes over time?? combination of stochastic and programmatic "decisions" is arbitrary-eg. why is mitotic mode treated as random variable but not cell cycle length, when both vary with some (maybe gaussian?) distribution over the population of RPCs?

As these cells move through the lineage program, the stochastic model accurately describes how these cells generate clones of variable size, and it also accurately predicts many characteristics of the actual lineage trees that we have seen in our time-lapse studies. Stochasticity also appears to be a feature of cell fate assignment. **We therefore speculate that all vertebrate retinas, though vastly different in size and the proportional composition**

of different cell types, may follow similar stochastic rules but tune their proliferative and cell fate probabilities to arrive at appropriate species- specific retinal sizes and cellular compositions.

-what is a "stochastic rule"? The implication of the earlier text is that stochastic noise in gene expression gives rise to observed variability in clonal size and fate. The implication of the speculation is that this noise is "tuned" to arrive at species-specific retinal compositions. Why does this "tuning" have to happen at the genomic level though? Ie. it seems clear that "tuning" produces another round of proliferation differentiation into the first few months of zebrafish life- is this temporal pattern really "tuned" at this level? Seems far more likely to be subject to extracellular signalling?

# Bibliography

- [1] Michalis Agathocleous and William A. Harris. From progenitors to differentiated cells in the vertebrate retina. *Annual Review of Cell and Developmental Biology*, 25:45–69, 2009.
- [2] M. R. Alexiades and C. L. Cepko. Subsets of retinal progenitors display temporally regulated and distinct biases in the fates of their progeny. *Development*, 124(6):1119–1131, March 1997.
- [3] Michel Cayouette, Ben A. Barres, and Martin Raff. Importance of intrinsic mechanisms in cell fate decisions in the developing rat retina. *Neuron*, 40(5):897–904, December 2003.
- [4] Constance L. Cepko, Christopher P. Austin, Xianjie Yang, Macrene Alexiades, and Diala Ezzeddine. Cell fate determination in the vertebrate retina. *Proceedings of the National Academy of Sciences*, 93(2):589–595, 1996.
- [5] Michael A. Dyer and Constance L. Cepko. Regulating proliferation during retinal development. *Nature Reviews Neuroscience*, 2(5):333–342, 2001.
- [6] D. M. Fekete, J. Perez-Miguelsanz, E. F. Ryder, and C. L. Cepko. Clonal analysis in the chicken retina reveals tangential dispersion of clonally related cells. *Developmental Biology*, 166(2):666–682, December 1994.
- [7] Francisco L. A. F. Gomes, Gen Zhang, Felix Carbonell, José A. Correa, William A. Harris, Benjamin D. Simons, and Michel Cayouette. Reconstruction of rat retinal progenitor cell lineages in vitro reveals a surprising degree of stochasticity in cell fate decisions. *Development (Cambridge, England)*, 138(2):227–235, January 2011.
- [8] W. A. Harris. Cellular diversification in the vertebrate retina. *Current Opinion in Genetics & Development*, 7(5):651–658, October 1997.
- [9] Christine E. Holt, Thomas W. Bertsch, Hillary M. Ellis, and William A. Harris. Cellular Determination in the Xenopus Retina is Independent of Lineage and Birth Date. *Neuron*, 1(1):15–26, March 1988.
- [10] D. J. Horsford. Chx10 repression of Mitf is required for the maintenance of mammalian neuroretinal identity. *Development*, 132(1):177–187, December 2004.
- [11] Christine L. Jasoni and Thomas A. Reh. Temporal and spatial pattern of MASH-1 expression in the developing rat retina demonstrates progenitor cell heterogeneity. *The Journal of Comparative Neurology*, 369(2):319–327, May 1996.



- [12] Tokushige Nakano, Satoshi Ando, Nozomu Takata, Masako Kawada, Keiko Muguruma, Kiyotoshi Sekiguchi, Koichi Saito, Shigenobu Yonemura, Mototsugu Eiraku, and Yoshiki Sasai. Self-Formation of Optic Cups and Storable Stratified Neural Retina from Human ESCs. *Cell Stem Cell*, 10(6):771–785, June 2012.
- [13] Allen G. Strickler, Kuburat Famuditimi, and William R. Jeffery. Retinal homeobox genes and the role of cell proliferation in cavefish eye degeneration. *International Journal of Developmental Biology*, 46(3):285–294, 2002.
- [14] Raphaël Thuret, Hélène Auger, and Nancy Papalopulu. Analysis of neural progenitors from embryogenesis to juvenile adult in *Xenopus laevis* reveals biphasic neurogenesis and continuous lengthening of the cell cycle. *Biology Open*, 4(12):1772–1781, November 2015.
- [15] Jeffrey M. Trimarchi, Michael B. Stadler, and Constance L. Cepko. Individual Retinal Progenitor Cells Display Extensive Heterogeneity of Gene Expression. *PLOS ONE*, 3(2):e1588, February 2008.
- [16] D. L. Turner, E. Y. Snyder, and C. L. Cepko. Lineage-independent determination of cell type in the embryonic mouse retina. *Neuron*, 4(6):833–845, June 1990.
- [17] David L. Turner and Constance L. Cepko. Turner1987.pdf. *Nature*, 328(9), July 1987.
- [18] Marta Vitorino, Patricia R Jusuf, Daniel Maurus, Yukiko Kimura, Shin-ichi Higashijima, and William A Harris. *Vsx2* in the zebrafish retina: Restricted lineages through derepression. *Neural Development*, 4(1):14, 2009.
- [19] R. Wetts and S. E. Fraser. Multipotent precursors can give rise to all major cell types of the frog retina. *Science*, 239(4844):1142–1145, March 1988.
- [20] Xin Zhang, Shaun Heaney, and Richard L. Maas. Cre-loxp fate-mapping of Pax6 enhancer active retinal and pancreatic progenitors. *genesis*, 35(1):22–30, January 2003.