nature neuroscience

The development of cortical circuits for motion discrimination

Gordon B Smith[1,8], Audrey Sederberg[2,7,8], Yishai M Elyada[1], Stephen D Van Hooser[3], Matthias Kaschube[4–6,9] & David Fitzpatrick[1,9]

Stimulus discrimination depends on the selectivity and variability of neural responses, as well as the size and correlation structure of the responsive population. For direction discrimination in visual cortex, only the selectivity of neurons has been well characterized across development. Here we show in ferrets that at eye opening, the cortical response to visual stimulation exhibits several immaturities, including a high density of active neurons that display prominent wave-like activity, a high degree of variability and strong noise correlations. Over the next three weeks, the population response becomes increasingly sparse, wave-like activity disappears, and variability and noise correlations are markedly reduced. Similar changes were observed in identified neuronal populations imaged repeatedly over days. Furthermore, experience with a moving stimulus was capable of driving a reduction in noise correlations over a matter of hours. These changes in variability and correlation contribute significantly to a marked improvement in direction discriminability over development.

Accurate visual discrimination depends critically on the selective responses of neurons in visual cortex for features of the visual scene such as the orientation of edges and their direction of motion. Other aspects of cortical responses, especially those that influence the spatial and temporal patterns of neuronal activity, are also important in visual discrimination. These include response variability[1–3], the number of responsive neurons[4–6] and the degree of correlation in neuronal response, all of which affect the performance of population coding in the mature visual cortex[7–10]. How these four features of the population response emerge and reach their mature state during the development of the visual cortex remains unclear.

Most is known about the development of stimulus selectivity, and studies in the ferret indicate that the time course of emergence and the role of experience differ according to the type of selectivity. For example, orientation selectivity is present and organized in a columnar fashion around the time of eye opening[11], whereas tuning for direction selectivity emerges shortly after eye opening in a process that requires visual experience[12]. Much less is known about the development of the temporal properties of the cortical population response, beyond the characterization of single units as "sluggish" and unreliable before and around the time of eye opening, becoming more crisp and reliable with continued experience[13,14]. Moreover, how these changes in single unit properties are related to the number of responsive neurons and the correlation structure of evoked responses remains unclear. However, two recent reports in rodents suggest that both of these properties may undergo significant postnatal maturation[15,16].

In this study we used two-photon *in vivo* calcium imaging to characterize the spatial and temporal response properties of large numbers of single neurons in ferret visual cortex to assess how these factors change during postnatal development. We found that cortical responses at eye opening were characterized by a high density of active neurons that displayed prominent wave-like activity, a high degree of variability and strong noise correlations. Over the next three weeks, the population response became increasingly sparse, wave-like activity disappeared, and variability and noise correlations were markedly reduced. The decrease in variability and noise correlations both contribute significantly to improvements in the ability of cortical neuronal activity to discriminate motion direction, and both the decrease in noise correlations and improvement in direction discriminability appear highly sensitive to visual experience. Taken together with previous observations in the ferret[12,17], these results indicate that the period following eye opening is distinguished by rapid changes in a number of neuronal response properties that are critical for motion discrimination.

RESULTS

Ferrets were imaged in three age groups (naive: postnatal day (P) 29–32, immature: P33–36, and mature: P48–50, with 0–1, 4–6 and >15 d visual experience, respectively) following intracortical injections of AAV expressing the fluorescent protein GCaMP3 (**Fig. 1a**). In animals imaged at eye opening, we observed dense and vigorous responses with strong orientation selectivity but weak direction selectivity, whereas in older animals responses were considerably

[1]Department of Functional Architecture and Development of Cerebral Cortex, Max Planck Florida Institute for Neuroscience, Jupiter, Florida, USA. [2]Department of Physics, Princeton University, Princeton, New Jersey, USA. [3]Department of Biology, Brandeis University, Waltham, Massachusetts, USA. [4]Frankfurt Institute for Advanced Studies, Frankfurt, Germany. [5]Faculty of Computer Science and Mathematics, Goethe University, Frankfurt, Germany. [6]Bernstein Focus: Neurotechnology Frankfurt, Frankfurt, Germany. [7]Present address: Department of Organismal Biology and Anatomy and Department of Neurobiology, University of Chicago, Chicago, Illinois, USA. [8]These authors contributed equally to this work. [9]These authors jointly directed this work. Correspondence should be addressed to D.F. (david.fitzpatrick@mpfi.org) or M.K. (kaschube@fias.uni-frankfurt.de).

Received 5 November 2014; accepted 10 December 2014; published online 19 January 2015; doi:10.1038/nn.3921

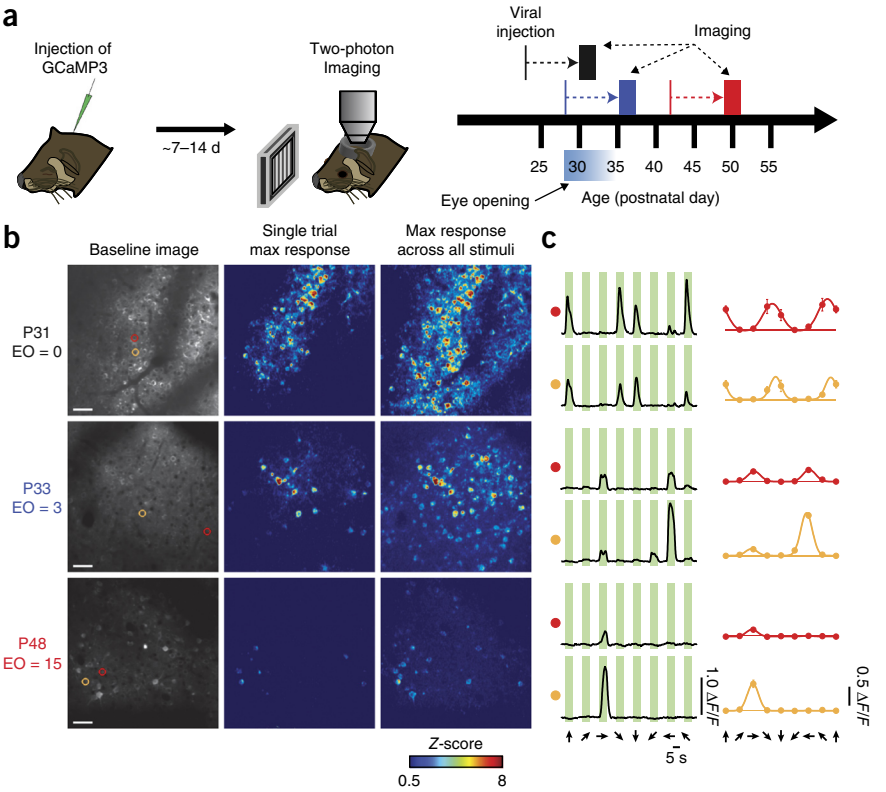

Figure 1 Response properties change dramatically following eye opening. (**a**) Experimental timeline. GCaMP3 expressing AAV was delivered intracortically via microinjection and two-photon imaging was performed 7–14 d later. Animals were imaged at either P29–32 (naive), P32–36 (immature) or P48–50 (mature). (**b**) Representative responses from animals in each age group. Left, baseline image. Scale bars, 50 µm. Middle, single-trial response to preferred stimulus (maximum projection across stimulus duration). Right, maximum response across all stimuli and all trials. EO, days after eye opening. (**c**) Responses for individual neurons highlighted in **b**. Left, response to eight directional stimuli, averaged across trials. Right, tuning curves fit with a two-peaked Gaussian. Horizontal line indicates the mean response to a blank stimulus.

sparser and direction selectivity was greatly increased (**Fig. 1b**,**c**). Pooling across animals, we observed similar results to those seen in previous work[12], with strong selectivity for orientation and weak selectivity for direction in naive animals, both of which increased significantly over the following weeks (orientation, **Fig. 2a**; Kruskal-Wallis test (KW): $\chi^2(2) = 309.45$, $P < 0.001$; pairwise Mann-Whitney U test (MW): naive: $Z(1,811) = -12.23$, $P < 0.001$; immature: $Z(1,447) = -15.30$, $P < 0.001$; mature: $Z(992) = -6.99$, $P < 0.001$; direction, **Fig. 2b**; KW: $\chi^2(2) = 473.87$, $P < 0.001$; pairwise MW: naive: $Z(1,555) = -16.51$, $P < 0.001$; immature: $Z(1,174) = -18.06$, $P < 0.001$; mature: $Z(883) = -7.56$, $P < 0.001$). Beyond these expected changes, we found that the trial-to-trial response variability decreased significantly in immature and mature animals as compared to naive animals (**Fig. 2c**; KW: $\chi^2(2) = 190.24$, $P < 0.001$; MW: naive versus immature: $Z(1,555) = -16.51$, $P < 0.001$; naive versus mature: $Z(1,174) = 18.06$, $P < 0.001$). Variability rebounded slightly but significantly from immature to mature animals (MW: $Z(883) = -7.56$, $P = 0.001$). Over this same period, the amplitude of the response evoked by the preferred stimulus did not change (**Fig. 2d**; $\Delta F/F$ mean ± s.e.m.: 0.148 ± 0.004, 0.147 ± 0.006 and 0.138 ± 0.007 for naive, immature and mature respectively; KW: $\chi^2(2) = 2.99$, $P = 0.22$).

To quantify the population sparseness[4], we examined the fraction of identified neurons within a field of view (FOV) that exhibited a response to at least one stimulus on a given trial, making this measurement resistant to changes in direction selectivity. The fraction of responsive cells did not change from naive to immature animals (**Fig. 2e**; mean ± s.e.m. across all trials for all stimuli: 70.2 ± 1.7% versus 68.7 ± 2.1%; KW across groups: $\chi^2(2) = 103.2$, $P < 0.001$; MW: naive versus immature: $Z(226) = 0.35$, $P = 0.72$). However, responses in mature animals were considerably sparser (44.4 ± 1.2%; MW: naive versus mature: $Z(238) = 9.38$, $P < 0.001$; immature versus mature: $Z(202) = 8.13$, $P < 0.001$). These results cannot be explained through enhanced direction selectivity, decreased responsivity to preferred stimuli (**Fig. 2d**), a change in the density of labeled neurons (KW: $\chi^2(2) = 2.69$, $P = 0.26$) or toxicity related to GCaMP expression (correlations between expression time and response density were nonsignificant in all groups: naive: $r(7) = 0.48$, $P = 0.19$; immature: $r(3) = -0.65$, $P = 0.24$; mature: $r(3) = -0.27$, $P = 0.65$), but rather may result from developmental changes in receptive field structure[13,18]. Notably, the response density in naive animals does not reflect a global nonspecific hyperexcitability, but rather was highly specific for both stimulus orientation and direction (**Fig. 2f**, within-group KW across stimuli: naive: $\chi^2(7) = 227.03$, $P < 0.001$; immature: $\chi^2(7) = 107.99$, $P < 0.001$; mature: $\chi^2(7) = 139.40$, $P < 0.001$; dominant versus orthogonal: naive: $Z(526) = 14.12$, $P < 0.001$; immature: $Z(382) = 8.36$, $P < 0.001$; mature: $Z(430) = 9.53$, $P < 0.001$; dominant versus opposite: naive: $Z(262) = 4.09$, $P < 0.001$; immature: $Z(190) = 3.31$, $P < 0.001$; mature: $Z(214) = 4.78$, $P < 0.001$).

Wave-like propagating responses in naive animals

In addition to the high density of the responsive neurons, a prominent feature of the visual response in young animals was a wave-like propagation of activity during the stimulus period (**Fig. 3a** and **Supplementary Movie 1**). The spatiotemporal pattern of activity was largely consistent both within and across stimuli (**Fig. 3b**,**c**) and frequently resembled a linear traveling wave (**Fig. 3d**). Both the fraction of trials eliciting a linear wave and the wave velocity exhibited age-dependent declines (wave incidence, **Fig. 3e**, mean ± s.e.m.: naive: 30.7 ± 5.0%, $n = 3$ FOV from 3 animals; immature: 10.2 ± 4.5%, $n = 6$ FOV from 3 animals; mature: 1.6 ± 0.5%, $n = 4$ FOV from 2 animals; KW: $\chi^2(2) = 8.62$, $P = 0.013$; wave velocity, **Fig. 3f**; naive: 256 ± 24 µm/s, $n = 115$; immature: 91 ± 11 µm/s, $n = 60$; mature: 44 ± 7 µm/s, $n = 6$; KW: $\chi^2(2) = 65.07$, $P < 0.001$).

The stimulus-evoked wave-like pattern of activity demonstrated here does not propagate continuously over the surface of the cortex but is limited to domains that are tuned to the orientation of the stimulus. Consistent with the significant orientation tuning that is present at eye opening, the likelihood of eliciting a linear wave within the field of view was highly dependent on the orientation of the stimulus (**Fig. 3g**, Friedman's test: naive: $\chi^2(7) = 19.05$, $P = 0.008$; immature: $\chi^2(7) = 16.00$, $P = 0.025$; mature: $\chi^2(7) = 21.00$, $P = 0.004$). There was also a bias toward a directional preference in the likelihood of eliciting a linear wave in all age groups (compare preferred to null stimulus in **Fig. 3g**).

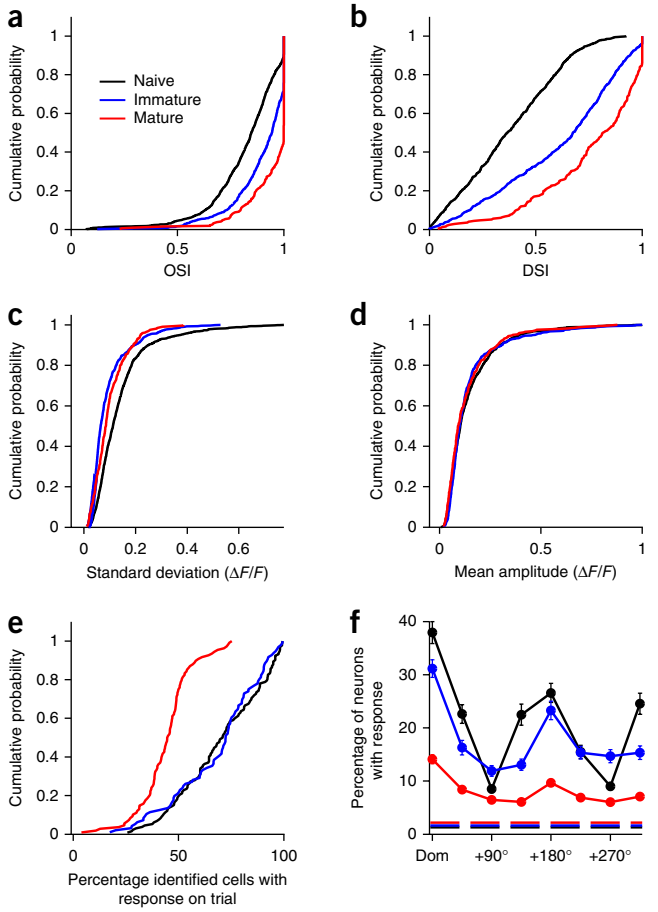

Figure 2 Stimulus selectivity increases and population response density decreases with age and experience. Naive, P29–32 (n = 9 animals); immature, P33–36 (n = 5 animals); mature, P48–50 (n = 5 animals). (**a**) Orientation selectivity (OSI) increased significantly with age (n = 1,134, 679 and 315 neurons for naive, immature and mature, respectively). (**b**) Direction selectivity (DSI) increases significantly with age (n = 924, 633 and 252 for naive, immature and mature, respectively). (**c**) Response variability (shown as s.d. across trials) to the preferred stimulus decreases with age. (**d**) Response amplitude to the preferred stimulus does not change across age (mean response across trials). (**e**) Response density (fraction of active neurons on a given trial out of all identified neuronal ROIs with at least one response) declines significantly from the naive and immature groups to the mature group. (**f**) Response density is stimulus specific. In all age groups, the fraction of neurons active on a given trial is significantly greater for the dominant (Dom) stimulus (the stimulus producing activity in the largest fraction of neurons, aligned across animals) than for a stimulus with opposite direction of motion (null, +180°) or an orthogonal orientation. Dashed line indicates mean fraction of neurons active during blank stimuli. Error bars are mean ± s.e.m. across animals.

The waves had a strong tendency to travel across the cortex in a consistent direction that was distinct for each animal (**Fig. 3h**; for FOVs with minimum 10 traveling waves (5 of 13, 3 naive and 2 immature); Hodges-Ajne test for nonuniformity: $P < 0.05$ in 4 of 5 cases with ≥30 waves each; fifth case $P = 0.49$ with 13 waves; statistics in **Supplementary Table 1**). The propagation direction varied significantly across animals (circular nonparametric multisample test for equal medians (CM)[19]: $M(4) = 35.075$, $P < 0.001$) and did not vary as a function of the direction of stimulus motion (CM within animal: $P > 0.14$ in all cases; statistics in **Supplementary Table 2**).

Noise correlations decrease with age

The presence of highly variable but highly dense responses in naive animals suggests a large degree of correlated variability at this age. As shared trial-to-trial variability has strong implications for neural coding[8,10], we examined pairwise noise correlations across development. Across all neuronal pairs, noise correlations exhibited a significant decline with age (mean ± s.e.m. across animals after averaging pairs within each FOV: 0.125 ± 0.023, 0.020 ± 0.003, 0.010 ± 0.002 for naive, immature and mature, respectively; KW: $\chi^2(2) = 19.94$, $P < 0.001$; MW: naive versus immature: $U(15) = 72$, $P < 0.001$; naive versus mature: $U(16) = 81$, $P < 0.001$; immature versus mature: $U(15) = 63$, $P = 0.008$).

In mature animals, noise correlations tend to be higher between nearby neurons with similar tuning properties[20–22]. We examined how this difference emerged during development. Noise correlations declined with increasing spatial distance for all groups (**Fig. 4a**, Friedman's test within group: naive: $\chi^2(13) = 106.55$, $P < 0.001$, $r(98{,}643) = -0.49$, $P < 0.001$; immature: $\chi^2(13) = 84.23$, $P < 0.001$, $r(34{,}222) = -0.19$, $P < 0.001$; mature: $\chi^2(12) = 35.79$, $P < 0.001$, $r(7{,}617) = -0.14$, $P < 0.001$), and the decrease in correlations with age was evident at all distances examined (Mack-Skillings test for age while controlling for distance: $\chi^2(2) = 253.76$, $P < 0.001$). Noise correlations also decreased with age for all intra-pair tuning differences (**Fig. 4b**, Mack-Skillings test for age while controlling for angular difference: $\chi^2(2) = 130.44$, $P < 0.001$). We further observed a strong relationship across ages between the mean noise correlation and the mean direction selectivity of each neuronal population we imaged (**Fig. 4c**, $r(24) = -0.59$, $P = 0.001$), indicating that these changes occur over similar timescales during development. Notably, neuron pairs with similar tuning and small spatial separation exhibited the highest noise correlations in naive animals (**Fig. 4d**) and also showed the largest decreases with age (**Fig. 4e**,**f**).

To determine whether the strong wave-like activity present in naive animals contributes to high noise correlations, we selected stimuli capable of eliciting linear waves and calculated noise correlations using only trials with and without waves. Trials with wave-like activity showed significantly higher noise correlations than those without (mean ± s.e.m. across 3 animals: wave present, 0.266 ± 0.019; absent, 0.125 ± 0.021; Wilcoxon signed-rank test (WSR): $Z(2{,}195) = 39.54$, $P < 0.001$, n = 2,197 pairs from 3 animals), indicating that wave-like responses are a major, albeit not the only, source of correlated variability in the naive cortex.

Chronic two-photon imaging of emergence of direction selectivity

To track the developmental changes that occur in individual neurons, we performed longitudinal imaging of a defined neuronal population over several days following eye opening. We successfully imaged four ferrets over two or three imaging sessions starting around eye opening (**Fig. 5a**,**b**). In these experiments, we only considered neurons that could be conclusively identified across all imaging sessions (73.8 ± 7.8 of cells identified in first imaging session, mean ± s.e.m. across 4 animals). In the example shown in **Figure 5c**,**d**, we imaged 126 visually responsive neurons over 3 imaging sessions from P32 to P37. In this population, we found an increase in direction selectivity from P32 to P35, which did not change further by P37 (**Fig. 5d**, Friedman's test: $\chi^2(2) = 14.02$, $P < 0.001$; *post hoc* WSR: P32 versus P35: $Z(124) = -3.84$, $P < 0.001$; P32 versus P37: $Z(124) = -3.33$, $P < 0.001$; P35 versus P37: $Z(124) = 0.31$, $P = 0.754$). To examine the changes underlying this increase in selectivity, we compared orientation and direction preferences for identified neurons across imaging sessions. We found that orientation preference was largely stable over

this period, while orientation selectivity increased significantly as would be expected from previous work[11] (**Supplementary Figs. 1** and **2a**,**c**,**e**; WSR: $Z(317) = 12.66$, $P < 0.001$). Interestingly, direction preference exhibited approximately 180° reversals in a subset of neurons (**Fig. 5e**,**g**). Among neurons with a stable preferred direction, selectivity increased significantly over imaging sessions (**Fig. 5f**; WSR: $Z(239) = 5.61$, $P < 0.001$).

Increased direction selectivity could occur through either potentiation of responses to the preferred direction, suppression of null responses or both. In neurons that maintained a stable direction preference (see Online Methods and **Fig. 5g**, top), the rise in selectivity was due to a potentiation of the response to the preferred direction (**Fig. 5h**, top; mean $\Delta F/F \pm$ s.e.m., with WSR, initial versus final: preferred: 0.33 ± 0.02 versus 0.43 ± 0.02, $Z(113) = 5.051$, $P < 0.001$; null: 0.17 ± 0.01 versus 0.18 ± 0.01, $Z(113) = 1.066$, $P = 0.286$; orthogonal: 0.04 ± 0.01 versus -0.01 ± 0.00, $Z(113) = -7.725$, $P < 0.001$; $n = 115$), consistent with the initial changes reported using pooled acute single-unit recordings[23]. Notably, neurons that exhibited reversals in preference (for example, **Fig. 5g**, middle) showed both a potentiation of the final preferred response and a depression of the

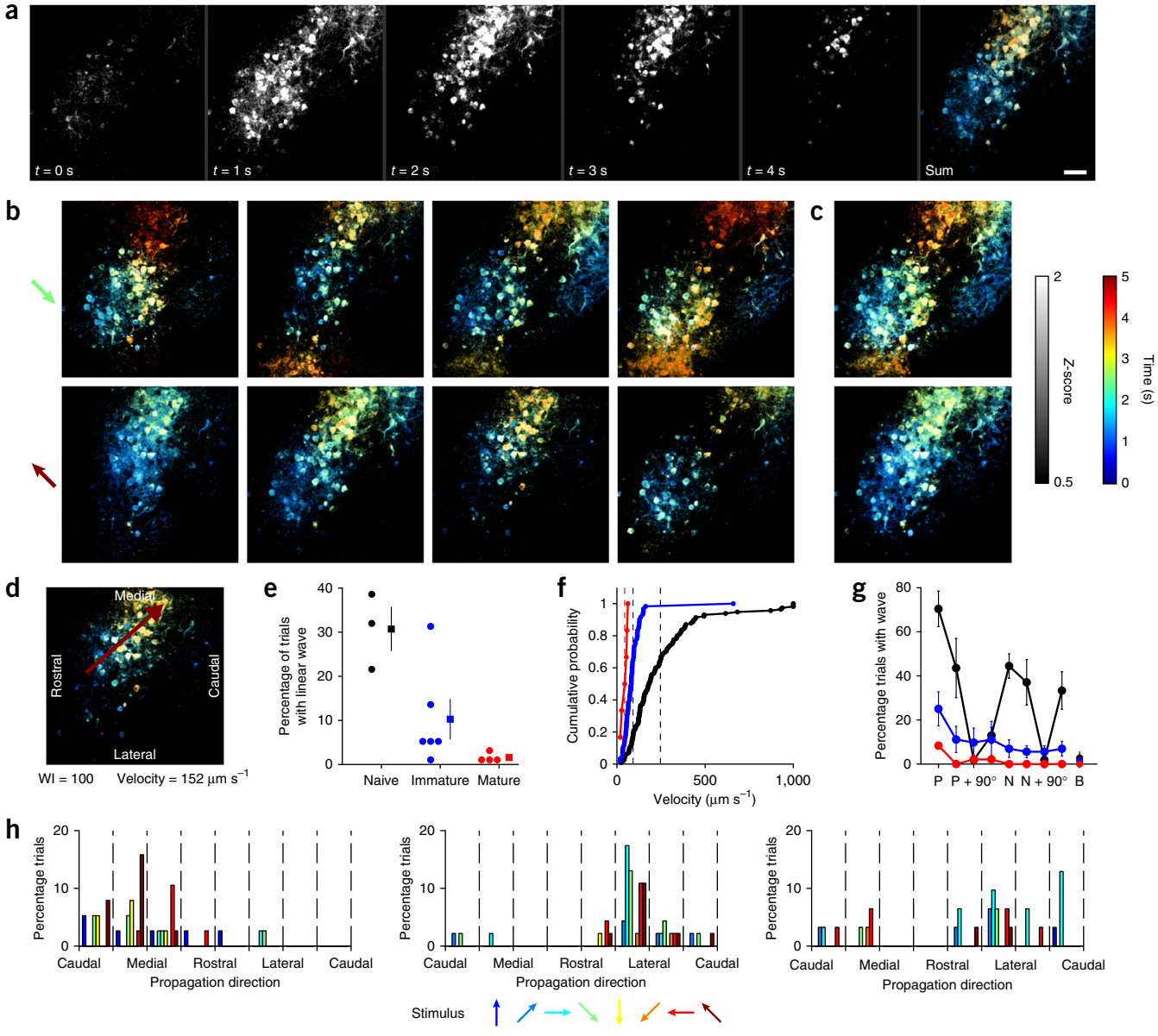

Figure 3 Wave-like responses to visual stimulation in young animals. (**a**) Pseudocolored time course of response to single stimulus. Propagating activity appears as a gradient from blue to orange. Scale bar, 50 µm; applies to **a**–**d**. (**b**) Single-trial example responses. Top row, responses to grating drifting down and to the left; bottom row, responses to stimulus of same orientation but opposite direction of motion. (**c**) Average response across all trials with strong wave-like activity. (**d**) Example response well fit by a linear traveling wave (wave index (WI) = 100). (**e**) Frequency of linear waves declines significantly with age ($n = 3$ FOV from 3 animals, 6 FOV from 3 animals and 4 FOV from 2 animals for naive, immature and mature, respectively). Circles indicate individual FOVs; squares indicate mean ± s.e.m. (**f**) Velocity of linear waves declines significantly with age (mean ± s.e.m.: naive: 256.8 ± 23.7 µm/s, $n = 115$ waves; immature: 91.3 ± 10.6 µm/s, $n = 60$ waves; mature: 44.2 ± 7.1 µm/s $n = 6$ waves). Dashed lines indicate geometric means. (**g**) Occurrence of linear waves is stimulus specific. P, preferred stimulus; N, null stimulus with opposite direction of motion from preferred; B, blank stimulus. Error bars are mean ± s.e.m. across animals. (**h**) Wave direction is largely consistent across all stimuli within an animal but differs across animals. Each histogram shows propagation directions for one animal in the naive group. Bar color indicates stimulus identity.

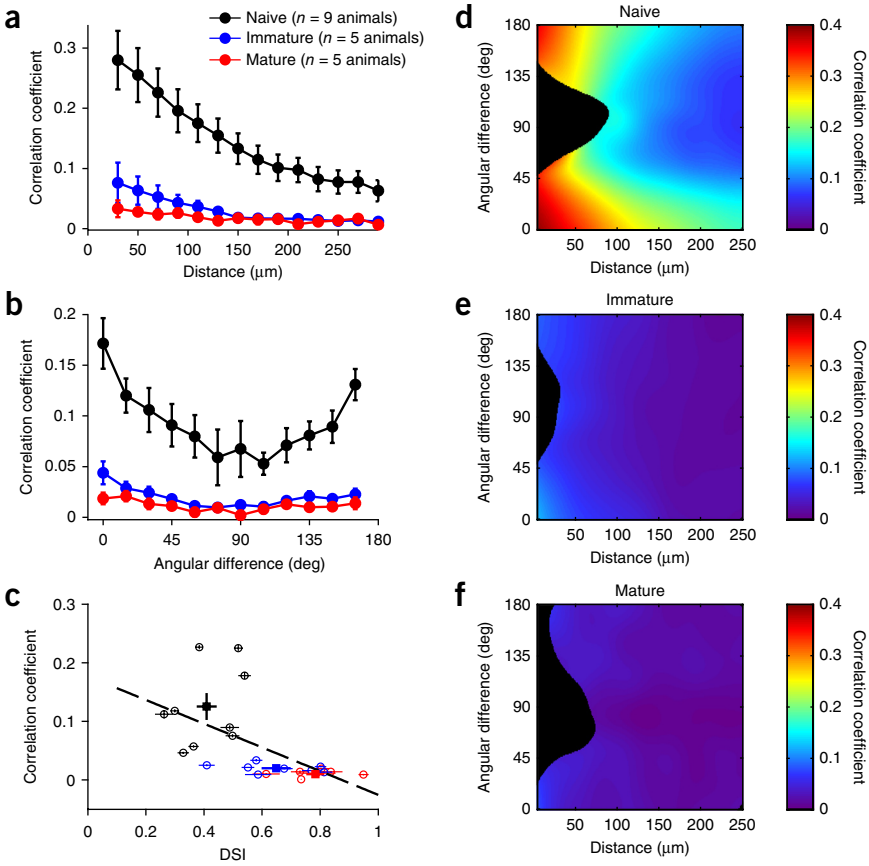

Figure 4 Noise correlations decline with age and experience. (**a**) Pairwise noise correlation as a function of intra-pair spatial distance. There is a significant decrease in noise correlation across age groups, as well as a significant decrease in noise correlation as a function of distance within each group. (**b**) Pairwise noise correlation as a function of the intra-pair difference in preferred direction. For all age groups, noise correlations were significantly higher for pairs with similar orientation preferences than for those with orthogonal preferences. Likewise, all age groups exhibited significantly higher correlations for pairs with similar as opposed to opposite direction preferences. Data in **a**,**b** are represented as the mean ± s.e.m. across animals after taking the mean across all pairs within an animal. (**c**) Noise correlations decrease and DSI increases with age. Circles, mean ± s.e.m. for each FOV; squares, mean ± s.e.m. across all animals in each age group. There is a significant correlation ($R = -0.59$; $P < 0.01$) between noise correlation and DSI across all experiments. (**d**–**f**) Maps of pairwise noise correlation as a function of intra-pair spatial distance and difference in preferred direction (local average based on Gaussian kernel, $\sigma_x = 15$ μm, $\sigma_y = 15$ deg). Early, noise correlation is large for small distances, in particular in pairs with similar preferred direction. In the mature cortex, this bias is reduced and noise correlations are small irrespective of distance and tuning difference. Black indicates regions with insufficient data (less than 20 pairs per 30-μm × 22.5-deg region).

initial preferred (final null) (**Fig. 5h**, middle; mean $\Delta F/F \pm$ s.e.m., with WSR: initial versus final: preferred: 0.24 ± 0.03 versus 0.53 ± 0.06, $Z(22) = 4.286$, $P < 0.001$; null: 0.34 ± 0.04 versus 0.27 ± 0.03, $Z(22) = -2.171$, $P = 0.030$; orthogonal: 0.04 ± 0.01 versus −0.04 ± 0.01, $Z(22) = -3.829$, $P < 0.001$; $n = 24$), as did cells with an initially uncertain preference that developed over time (**Fig. 5g**,**h**, bottom; mean $\Delta F/F \pm$ s.e.m., with WSR, initial versus final: preferred: 0.26 ± 0.02 versus 0.37 ± 0.02, $Z(96) = 5.732$, $P < 0.001$; null: 0.21 ± 0.02 versus 0.17 ± 0.01, $Z(96) = -2.748$, $P = 0.006$; orthogonal: 0.07 ± 0.01 versus 0.00 ± 0.01, $Z(96) = -6.844$, $P < 0.001$; $n = 98$).

These changes in single-cell response were accompanied by a significant decline in pairwise noise correlation among the longitudinally imaged populations (**Fig. 6a** and **Supplementary Fig. 3**; WSR: $Z(18{,}946) = -98.10$, $P < 0.001$, $n = 18{,}948$ pairs). We next assessed the relationship between the change in selectivity of individual neurons and the structure of the population response during the period following eye opening (**Fig. 6b**). In three of four longitudinal imaging experiments, direction selectivity increased from the initial to the final imaging session (Friedman's test: $P < 0.05$ for each experiment with 33, 147 and 126 neurons per experiment, respectively; statistics in **Supplementary Table 3**), while pairwise noise correlations among this same population decreased (Friedman's test: $P < 0.001$ for each experiment, with 489, 10,618 and 7,770 pairs per experiment, respectively; statistics in **Supplementary Table 3**). In the fourth experiment, only 13 neurons were both identifiable and visually responsive across days and neither changes in direction selectivity nor in noise correlation were significant (Friedman's test: direction selectivity: $P = 0.58$, $n = 13$ cells, correlation: $P = 0.49$, $n = 69$ pairs). We also observed a strong decrease in trial-to trial variability over this same period (initial versus final imaging session, WSR: $Z(317) = 15.48$, $P < 0.001$). These results clearly show that individual neurons in defined populations become both more selective for direction of motion, less variable and less correlated following eye opening.

Given that noise correlations can reflect the influence of common inputs and recurrent connectivity[24,25], it is possible that high noise correlations between pairs of neurons early in development may predict shared tuning properties later on. To examine this, we compared noise correlations during the initial imaging session for neurons with similar preferred directions (within 45°) that maintained that similarity (S–S pairs) versus those that adopted opposite preferences by the final imaging session (S–O pairs) (**Fig. 6c** and **Supplementary Fig. 4**). Initial noise correlations were significantly higher for S–S pairs than S–O pairs (**Fig. 6d**; MW: $Z(4{,}857) = 3.01$, $P < 0.001$, $n = 3{,}098$ and 1,761 for S–S and S–O, respectively). Likewise, initial noise correlations were higher between pairs with opposite preferences that would ultimately adopt similar tuning (O–S pairs) than for pairs that would maintain opposing preferred directions (O–O pairs) (**Fig. 6e**; MW: $Z(2{,}448) = 8.74$, $P < 0.001$, $n = 921$ and 1,529 for O–S and O–O, respectively). Taken together, these results indicate that initial pairwise noise correlations may predict future tuning similarity, suggesting the presence of functionally interconnected subpopulations destined to adopt similar tuning.

Decline in variance improves direction discriminability

Traditional measures of direction selectivity are based on the average response of neurons over multiple stimulus trials in order to 'average out' the noise that can be present on a single trial. However, reliable behavioral discriminations of motion direction are based on neuronal responses to a single stimulus presentation. Thus, in addition to selectivity, the variability in single neuron response and correlation in

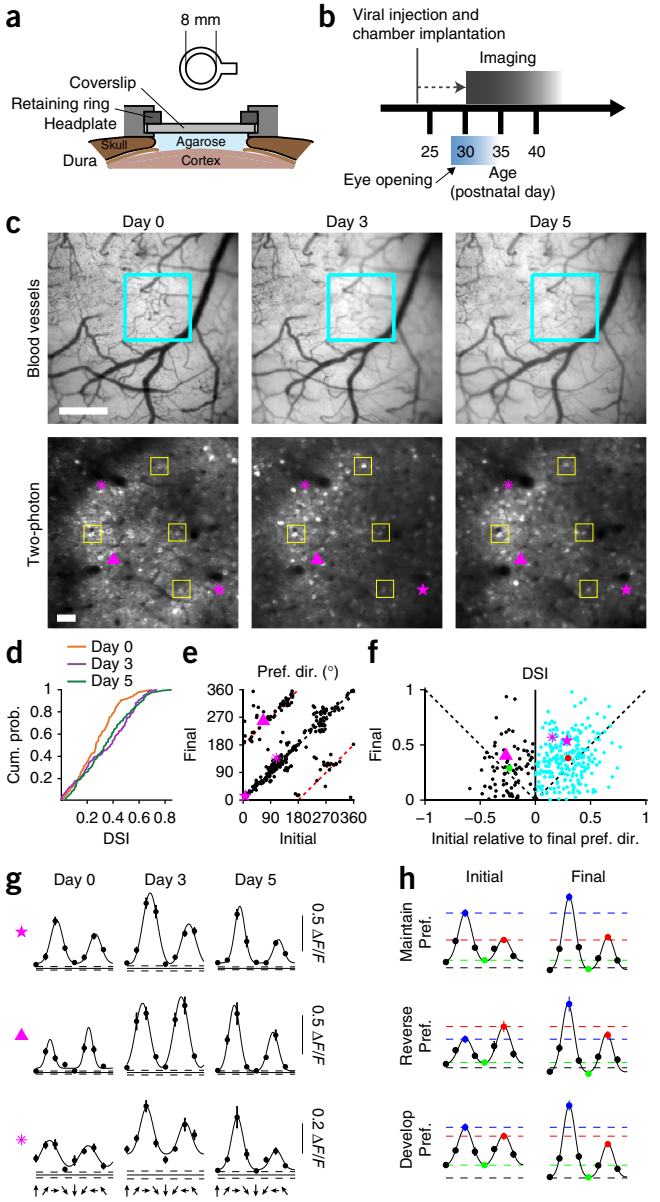

Figure 5 Longitudinal two-photon imaging reveals emergence of direction selectivity in identified neurons. (**a**) Schematic of chronically implanted headplate. (**b**) Time course of imaging experiments. (**c**,**d**) Longitudinal imaging across 5 d following eye opening. (**c**) Top, cortical blood vessel pattern from day 0 to day 5. Blue box indicates area for functional imaging. Bottom, imaging field over days. Yellow boxes outline examples of corresponding neurons. 126 neurons (63% of day 0) remained visible and visually responsive across all imaging sessions. Images were aligned across days with an affine transform. Purple symbols indicate example neurons shown in **g**. Scale bars: top, 500 µm; bottom, 50 µm. (**d**) Direction selectivity (DSI) increases in a population of neurons imaged from a single animal over 5 d ($n = 126$ neurons). Cum. prob., cumulative probability. (**e**) Direction preference is stable in the majority of neurons, whereas a subset of cells exhibit 180-degree reversals. Red dashed lines indicate 180-degree shift from day 0. Pref. dir., preferred direction. (**f**) DSI increases in neurons with stable preferences (blue symbols; red circle indicates mean, with s.e.m. smaller than marker size) from the initial to final imaging session. Neurons that reverse preference (black; green indicates mean, with s.e.m. smaller than marker size) do not show an increase in selectivity. (**g**,**h**) Example (**g**) and average (**h**) tuning curves showing neurons that maintained a preferred direction (top), reversed preferred direction (middle) and developed a clear preferred direction (bottom). Purple symbols in **c**,**e**–**g** indicate corresponding neurons. Error bars in **g** are s.e.m. across trials ($n = 12$). Error bars in **h** are s.e.m. across neurons ($n = 115$, 24, 98 for maintained, reversed and developed a preferred direction, respectively). For ease of comparison, tuning curves were shifted to align preferred stimuli on day 5.

response of the active population may affect the capacity of network activity to discriminate different directions of motion.

As would be expected, the direction discriminability (**Supplementary Fig. 5a**) of single neurons improved considerably from the naive to immature groups (**Fig. 7a**; MW: $Z(1,920) = -18.3$, $P < 0.001$). To assess whether this increase resulted primarily from an increase in the selectivity of individual neurons (**Fig. 2b**) or a decrease in the variance of their response (**Fig. 2c**), we computing direction discriminability for two hybrid data sets (see Online Methods). Replacing the variance in naive neurons with that of immature neurons resulted in much higher direction discriminability than using immature tuning with naive variance (**Fig. 7a**). This suggests that within a few days after eye opening, the reduction of variance, albeit modest, has a stronger impact on improvements in direction discriminability than the increase in direction selectivity (MW: $Z(1,998) = -4.10$, $P < 0.001$). Between the immature and the mature stage, single neuron discriminability changed little (**Fig. 7a**) despite an increase in direction selectivity during this period (**Fig. 2b**), which is consistent with the slight increase in variability observed over the same period (**Fig. 2c**).

To determine the degree to which developmental reductions in trial-to-trial variability enhance discriminability in individual neurons, we examined direction discriminability in longitudinally imaged animals. We found a significant increase in discriminability over days (**Supplementary Fig. 6a**,**b**; WSR: $Z(201) = -6.56$, $P < 0.001$), including within a subset of neurons that exhibited reduced selectivity (termed LS for low selectivity; **Supplementary Fig. 6c**). In LS neurons, the response amplitude to both the preferred and null stimuli did not change (**Supplementary Fig. 6d**; WSR: preferred: $Z(30) = 1.46$, $P = 0.145$; null: $Z(30) = 0.79$, $P = 0.432$; $n = 32$), in contrast to that in neurons with increased selectivity (termed HS neurons), in which the preferred response was enhanced and the null suppressed (**Supplementary Fig. 6e**; WSR: preferred: $Z(110) = 5.33$, $P < 0.001$; null: $Z(110) = -5.20$, $P < 0.001$; $n = 112$). Notably, although in both groups variability decreased and this decrease contributed significantly to a rise in discriminability, the decline in variability and corresponding contribution to improved discriminability were both significantly greater in cells lacking an increase in selectivity (**Supplementary Fig. 6f–i**; WSR: variance: LS: $Z(30) = -3.85$, $P < 0.001$, HS: $Z(110) = -4.26$, $P < 0.001$; MW, LS versus HS: $Z(142) = -2.38$, $P = 0.017$; effect of variance on discriminability, WSR: LS: $Z(30) = 4.02$, $P < 0.001$; HS: $Z(110) = 4.23$, $P < 0.001$; MW, LS versus HS: $Z(142) = 3.09$, $P = 0.002$). These results show that developmental reductions in variability are capable of driving increased discriminability in the absence of improved selectivity.

Decrease in noise correlations improves discriminability

While direction discriminability based on the responses of single neurons is enhanced with age, behaviorally relevant direction discriminations are likely to depend on the distribution of activity in populations of cortical neurons where both trial-to-trial variance and noise correlations are contributing factors (**Supplementary Fig. 5b**). Therefore, we sought to assess changes in discriminability in groups of neurons over the course of development and the contribution of a reduction in noise correlation and variance to these changes. As expected, the median discriminability increased with group

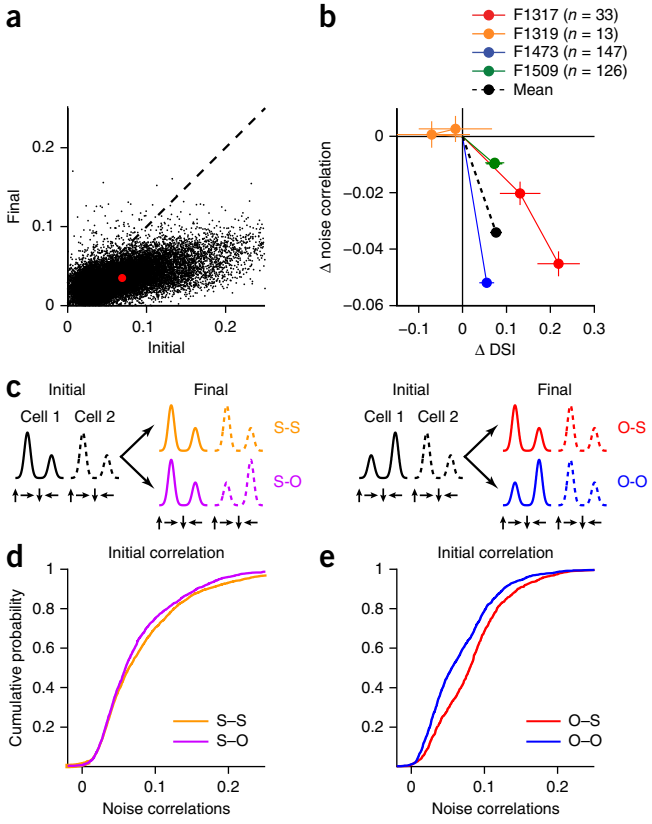

Figure 6 Correlation between decrease in noise correlation and direction selectivity. (**a**) Pairwise noise correlations decreased significantly from the initial to the final imaging session. Red dot indicates mean across all pairs and animals. (**b**) Relationship between change in pairwise noise correlation and direction selectivity (relative to day 0) for 4 animals in which chronic imaging was performed. In 3 of 4 cases, noise correlations exhibited a significant decrease by the final imaging session, whereas direction selectivity significantly increased ($n = 33$, 147 and 126 neurons per experiment). In one experiment (F1319, orange, $n = 13$ neurons), changes in neither DSI nor correlations were significant. Data are shown as mean ± s.e.m. across all neurons. Averaging across all neurons and pairs on the final imaging session (black) reveals a significant decrease in noise correlation and a significant rise in direction selectivity. (**c**) Cartoon depicting possible relationships in pairwise angular preference across imaging sessions. (**d**) Initial pairwise noise correlations were higher for pairs that will maintain similar preferences (S–S pairs) than those that adopt opposite preferences (S–O pairs). (**e**) Initial correlations were higher for pairs with opposite preferences if that pair will adopt matching preferred directions on the final imaging session (O–S) than for pairs that maintain opposite preferences (O–O).

size (**Fig. 7b**,**c**). Consistent with the single cell results, discriminability improved considerably from the naive to the immature stage for group sizes up to $N = 20$ (**Fig. 7b**). To disentangle the different factors contributing to this improvement, we again constructed hybrid data sets combining aspects of both the naive and immature data sets (see Online Methods). Combining direction selectivity of the naive set with the variance and correlation structure of the immature cortex, we obtained levels of discriminability close to those reached by the true immature set (**Fig. 7b**), confirming the single-cell result above that, early on, the increase in direction selectivity is not the prevalent factor in strengthening direction discriminability.

To estimate how much of this enhancement is due to a change in noise correlation, rather than a decrease in (single cell) variance, we generated a second hybrid set, which shared direction tuning and variance with the naive but structure of noise correlations with the immature set. For group sizes of four or larger, discriminability was higher in this set than in the true naive set (**Fig. 7b**, MW: $P < 0.05$; full statistics in **Supplementary Table 4**) and nearly as high as when destroying all noise correlations in the naive set by shuffling (**Fig. 7c**), with the effects of noise correlations increasing with group size (**Fig. 7b**,**d**). For the immature and mature sets, differences between the real and trial-shuffled controls were close to zero for all group sizes. These results indicate that on a population level changes in the structure of correlations contribute considerably to the improvement in direction discrimination early in development.

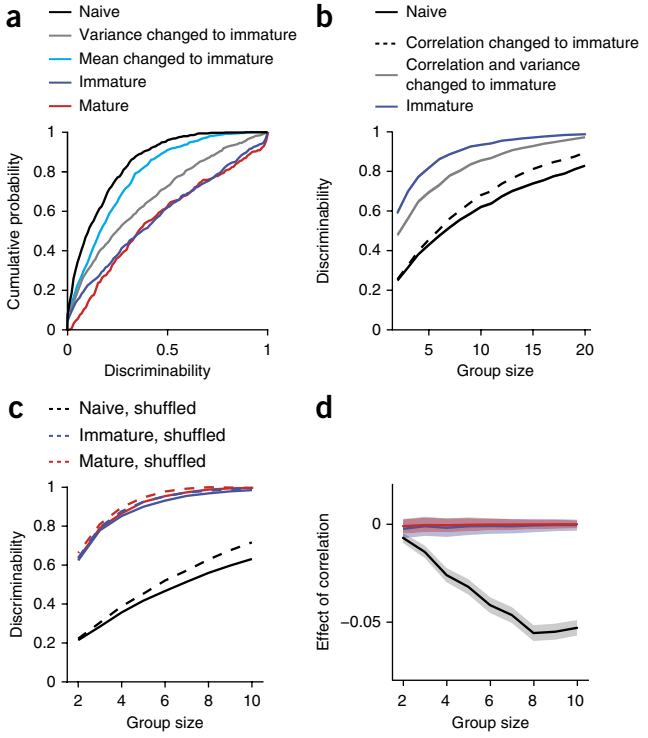

Figure 7 Changes in variance and correlation structure contribute to increased discriminability over development. (**a**) Single-cell direction discriminability increases significantly between the naive and immature ages (black versus dark blue). Combining immature levels of variance with naive levels of selectivity (gray) results in near immature levels of discriminability, whereas the gain in selectivity alone between the naive and immature ages (light blue) has a much smaller effect on discriminability (gray versus light blue). (**b**) Effects of variance and structure of noise correlations on discriminability by groups of cells. The overall change in variance and correlation structure increases discriminability substantially between the naive and immature ages. A considerable fraction (increasing with group size) is due to a change in the structure of noise correlations alone (black dashed line). (**c**) Effect of eliminating correlations by trial shuffling on discriminability by groups of cells. (**d**) Differences between solid and dashed lines from **c**. Shaded region shows ± 1 s.e.m.

Motion training decreases noise correlations

Our results showing changes in variance and noise correlation, and their contributions to improvements in discriminability, raise two important questions: (i) how rapidly can these changes in selectivity, variance and correlation, as well as their contributions to discriminability, occur? and (ii) are these changes affected by the visual experience of the animal?

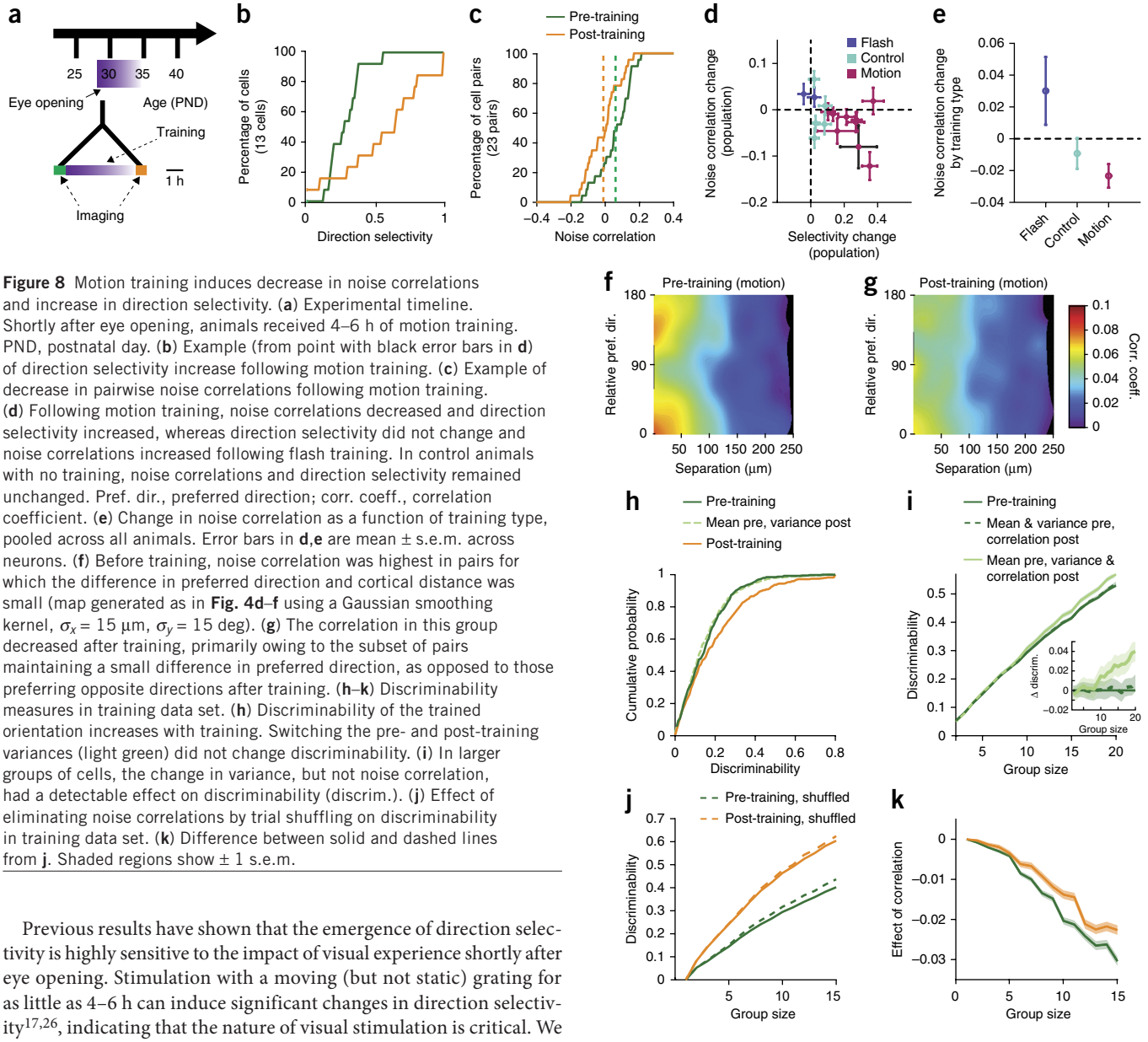

Figure 8 Motion training induces decrease in noise correlations and increase in direction selectivity. (**a**) Experimental timeline. Shortly after eye opening, animals received 4–6 h of motion training. PND, postnatal day. (**b**) Example (from point with black error bars in **d**) of direction selectivity increase following motion training. (**c**) Example of decrease in pairwise noise correlations following motion training. (**d**) Following motion training, noise correlations decreased and direction selectivity increased, whereas direction selectivity did not change and noise correlations increased following flash training. In control animals with no training, noise correlations and direction selectivity remained unchanged. Pref. dir., preferred direction; corr. coeff., correlation coefficient. (**e**) Change in noise correlation as a function of training type, pooled across all animals. Error bars in **d**,**e** are mean ± s.e.m. across neurons. (**f**) Before training, noise correlation was highest in pairs for which the difference in preferred direction and cortical distance was small (map generated as in **Fig. 4d–f** using a Gaussian smoothing kernel, $\sigma_x = 15$ μm, $\sigma_y = 15$ deg). (**g**) The correlation in this group decreased after training, primarily owing to the subset of pairs maintaining a small difference in preferred direction, as opposed to those preferring opposite directions after training. (**h**–**k**) Discriminability measures in training data set. (**h**) Discriminability of the trained orientation increases with training. Switching the pre- and post-training variances (light green) did not change discriminability. (**i**) In larger groups of cells, the change in variance, but not noise correlation, had a detectable effect on discriminability (discrim.). (**j**) Effect of eliminating noise correlations by trial shuffling on discriminability in training data set. (**k**) Difference between solid and dashed lines from **j**. Shaded regions show ± 1 s.e.m.

Previous results have shown that the emergence of direction selectivity is highly sensitive to the impact of visual experience shortly after eye opening. Stimulation with a moving (but not static) grating for as little as 4–6 h can induce significant changes in direction selectivity[17,26], indicating that the nature of visual stimulation is critical. We therefore asked whether exposure to a moving stimulus induces a rapid decrease in variability and noise correlations and, if so, whether these changes improve discriminability beyond that expected from gains in selectivity alone. Data originally reported by Li *et al.*[17] from ferrets that underwent 4–6 h of motion training (**Fig. 8a**) were analyzed for changes in pairwise noise correlations as a function of training. In a population of 13 neurons from an example animal (**Fig. 8b**,**c**), we observed a strong increase in direction selectivity, as would be expected following motion training (**Fig. 8b**, single FOV, $n = 13$ cells, mean difference ± s.e.m.: +0.29 ± 0.11, MW: $Z(24) = 2.46$, $P = 0.014$; for all experiments: $n = 396$ cells, +0.22 ± 0.02, MW: $Z(394) = -11.80$, $P < 0.001$). Among this same population, we also observed a significant decrease in pairwise noise correlations (**Fig. 8c**, single FOV: $n = 14$ pairs, mean difference ± s.e.m.: −0.10 ± 0.04, MW: $Z(26) = -2.37$, $P = 0.018$; for all experiments: $n = 1{,}247$ pairs, −0.022 ± 0.004, MW: $Z(1{,}245) = -4.47$, $P < 0.001$). In contrast, flash training with unmoving stimuli, which fails to elicit changes in direction selectivity[17], not only failed to produce a decrease in noise correlations but actually significantly increased noise correlations (change +0.05 ± 0.01, MW: $Z(253) = 4.78$, $P < 0.001$, $n = 255$ pairs). In control experiments where no training stimulus was given, neither direction selectivity nor noise correlations exhibited significant changes (change −0.006 ± 0.005; MW: $Z(1{,}026) = -1.19$, $P = 0.234$, $n = 1{,}028$ pairs). When analyzed over all experiments, there was a clear relationship between increasing direction selectivity and decreasing noise correlations (**Fig. 8d**,**e**; correlation $r(13) = -0.57$; $P = 0.025$, $n = 15$ experiments (10 motion, 2 flash, 3 control)). These results demonstrate that visual experience with moving stimuli drives rapid changes both in selectivity and in noise correlation, and demonstrate that these changes depend on the nature of visual experience.

Next we tested the degree of specificity by which changes in noise correlations are inter-related with changes in direction selectivity. As during normal development, the largest training-induced decrease in noise correlations occurred in pairs with similar tuning properties (similar preferred direction and small cortical distance) (**Fig. 8f**,**g**; compare **Fig. 4d–f**). However, only the cell pairs with similar tuning before and after training (the prevalent case) showed a significant decrease in

noise correlation, whereas for pairs that changed from same to opposite preferred direction over training, noise correlations remained significantly larger (change in correlation: similar: −0.016 ± 0.007, n = 314; opposite: +0.007 ± 0.009, n = 121; MW: $Z(433) = -2.35$, $P = 0.018$). Similar results were obtained when comparing noise correlations from identified populations in longitudinally imaged animals, where correlations for cell pairs with similar preferences on the final day underwent greater reductions than those for pairs with opposite final-day preferences (**Supplementary Fig. 7**; MW: S–S versus S–O: $Z(4,857) = -3.55$, $P < 0.001$, n = 3,098, 1,761; O–S versus O–O: $Z(2,448) = -2.60$, $P = 0.009$, n = 921, 1,529). As positive noise correlations are most detrimental among pairs with similar tuning[10,27], our results suggest the changes in noise correlations both over normal development and induced by motion training are not simply a general decrease in the strength of all correlations, but may be specifically structured in such a way as to improve discriminability.

Motion training significantly improved discriminability in single-cell responses along the trained orientation (**Fig. 8h**, mean ± s.e.m.: 0.05 ± 0.19, n = 396 cells, MW: $Z(394) = -3.28$, $P < 0.001$). Unlike in our acute data set, we failed to detect a significant change in variance (**Supplementary Fig. 8**; $Z(394) = -0.64$, $P = 0.52$), and the increase in single-cell discriminability could be explained by an increase in selectivity (**Fig. 8h**). However, the impact of altered variability became apparent when examining larger populations (**Fig. 8i**). Removing noise correlations through shuffling resulted in improved discriminability, both before and after training (**Fig. 8j**,**k**, before training: $P < 0.05$ for group size >3, after training: $P < 0.05$ for group size >7; full statistics in **Supplementary Table 5**). However, changes in variance made a larger contribution to improved discriminability than changes in correlation (**Fig. 8i**, MW: for group size >2, $P < 0.05$; full statistics in **Supplementary Table 6**). These results demonstrate not only that changes in noise correlations and discriminability can occur in as little as hours but also that these changes depend on the nature of visual experience.

DISCUSSION

Over a brief period of time following eye opening, the population response in the visual cortex of the ferret undergoes a series of changes that dramatically improve stimulus discriminability. In addition to previously documented increases in direction selectivity[12], we find that in the 2 to 3 weeks following eye opening, the active population undergoes a striking transformation from a highly dense response with complex spatiotemporal wave-like dynamics to a sparse distribution of active neurons. This transformation is characterized by a decline in variance and pairwise noise correlations over this period, occurring with the same time course and in the same neuronal population as the rise in direction selectivity. The high noise correlations present at eye opening limit the direction discriminability of the neuronal population, which improves significantly with the decline in noise correlation. In naive animals, the nature of visual experience appears to play a major role in these processes: experience with moving, but not static, stimuli drives both an enhancement in direction selectivity and a decrease in noise correlations, resulting in improved discrimination.

Dense wave-like responses dominate at eye opening

At eye opening, visual stimulation engages a high percentage of layer 2/3 neurons, and the calcium responses in these neurons exhibit complex spatiotemporal dynamics that have the appearance of propagating waves. Spontaneous wave-like activity appears to be a common feature throughout the developing brain, including the retina, thalamus and cortex, and is thought to help to establish orderly maps of retinotopy and ocular dominance through its ability to synchronize the activity of nearby neurons[28–30]. This suprathreshold activity is distinct from the traveling waves of fast subthreshold depolarization observed in adult visual cortex in response to focal sensory stimulation[31,32].

The wave-like responses and the sparsification with age that we report here are reminiscent of changes observed in the spontaneous calcium activity in the developing mouse cortex, which is especially prominent before eye opening[15]. However, there are several notable differences between the waves we observe and those described previously in mice[15,33–35]. First, wave-like activity in naive ferrets is stimulus specific and local, rather than propagating across the entire cortical surface[33,35]. Second, the propagation direction varies greatly across animals, suggesting the absence of a stereotypic wave pattern[34]. Lastly, the developmental profile of cortical waves appears to differ across species, with waves in mice reflecting, at least in part, the transmission of spontaneous retinal wave activity through the early visual system[33,34], whereas spontaneous retinal wave activity is thought to have diminished by eye opening in the ferret[36].

Regardless of whether the wave-like patterns of activity evoked in the developing ferret cortex are a reflection of retinal waves or downstream processes, it is clear that these patterns of activity reflect an interaction of this input with cortical circuitry[37,38], as they are constrained by the architecture of orientation selective domains. It is possible that these intracolumnar waves serve as a source of local temporal correlation that promotes map refinement and local response coherence.

Response decorrelation with visual experience

In visually naive ferrets, we found that removing noise correlations via shuffling resulted in considerable improvements in direction discriminability. Thus, in the developing ferret cortex, correlated noise limits the performance of the population. As noise correlations decrease with visual experience, these limits are relaxed and the population nears the performance achievable by a completely decorrelated network. Moreover, by factoring out any increase in direction selectivity taking place over the considered period, we found that the observed early changes in noise correlation and variance alone can lead to strong improvements in direction discriminability. The relative contribution of noise correlations becomes more apparent when considering larger groups of neurons ($n < 20$). We expect their impact to be even more pronounced in larger groups or neurons, but a direct assessment was not possible in our data owing to the limited number of stimulus repetitions and the early saturation in discriminability encountered with moving grating stimuli.

One concern that may arise is whether the use of anesthesia for these experiments undermines the significance of the observations. It is possible that the patterns of cortical activity in the awake animal at these ages differ from those encountered in our preparation; however, our results show that animals maintained under the same anesthetic regime at different ages exhibit profoundly different patterns of activity in response to an identical stimulus, which results in sizable differences in discriminability. How the developmental changes in cortical responses observed under these conditions manifest in the awake animal and how these changes contribute to behavioral performance remain important questions for future studies.

Potential circuit mechanisms for response maturation

The decline in noise correlations during the period following eye opening represents a major operational shift for the developing cortex, from a design that is well suited for maximizing the correlations in patterns of spontaneous activity necessary to build circuits to a design that

minimizes correlation in order to improve discrimination of activity patterns evoked by sensory stimulation. A refinement of feedforward inputs along with an increase in the degree of preferential interconnectivity among similarly tuned cells[39,40] could account for many of our results, including the increases in selectivity and sparseness, as well as the decrease in wave-like activity. As noise correlations can reflect common inputs[24,25], such refinement could also drive reductions in noise correlations.

The maturation of intracortical inhibition over this period of development[41,42], which likewise can promote response selectivity[43] and sparsification[44], dampen wave-like activity and decrease noise correlations[45–49], is also likely to contribute. Ultimately, whether the result of a refinement of excitatory connections, a maturation of inhibition or both, population response properties undergo dramatic changes following eye-opening that have profound effects on stimulus discriminability.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We would like to thank D. Ouimet and V. Hoke for technical and surgical assistance, and R. Corlew for administrative support. This research was supported by US National Institutes of Health grants EY011488 (D.F.), EY022001 (G.B.S.), 5T32HG003284 (A.J.S.) and Bernstein Focus Neurotechnology grant 01GQ0840 (M.K.), as well as the Max Planck Florida Institute.

AUTHOR CONTRIBUTIONS

G.B.S., A.S., M.K. and D.F. designed the study, analyzed the results and wrote the paper. G.B.S. performed the acute and longitudinal GCaMP imaging. Y.M.E. developed the method for longitudinal imaging. S.D.V.H. originally acquired the motion training data and prepared these data for the additional analyses reported here.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Tomko, G.J. & Crapper, D.R. Neuronal variability: non-stationary responses to identical visual stimuli. *Brain Res.* **79**, 405–418 (1974).
2. Tolhurst, D.J., Movshon, J.A. & Dean, A.F. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res.* **23**, 775–785 (1983).
3. Tolhurst, D.J. The amount of information transmitted about contrast by neurones in the cat's visual cortex. *Vis. Neurosci.* **2**, 409–413 (1989).
4. Willmore, B. & Tolhurst, D.J. Characterizing the sparseness of neural codes. *Network* **12**, 255–270 (2001).
5. Yen, S.C., Baker, J. & Gray, C.M. Heterogeneity in the responses of adjacent neurons to natural stimuli in cat striate cortex. *J. Neurophysiol.* **97**, 1326–1341 (2007).
6. Weliky, M., Fiser, J., Hunt, R.H., Wagner, D.N. & York, N. Coding of natural scenes in primary visual cortex. *Neuron* **37**, 703–718 (2003).
7. Olshausen, B.A. & Field, D.J. Sparse coding of sensory inputs. *Curr. Opin. Neurobiol.* **14**, 481–487 (2004).
8. Shadlen, M.N. & Newsome, W.T. Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* **4**, 569–579 (1994).
9. Abbott, L.F. & Dayan, P. The effect of correlated variability on the accuracy of a population code. *Neural Comput.* **11**, 91–101 (1999).
10. Averbeck, B.B., Latham, P.E. & Pouget, A. Neural correlations, population coding and computation. *Nat. Rev. Neurosci.* **7**, 358–366 (2006).
11. Chapman, B., Stryker, M.P. & Bonhoeffer, T. Development of orientation preference maps in ferret primary visual cortex. *J. Neurosci.* **16**, 6443–6453 (1996).
12. Li, Y., Fitzpatrick, D. & White, L.E. The development of direction selectivity in ferret visual cortex requires early visual experience. *Nat. Neurosci.* **9**, 676–681 (2006).
13. Chapman, B. & Stryker, M.P. Development of orientation selectivity in ferret visual cortex and effects of deprivation. *J. Neurosci.* **13**, 5251–5262 (1993).
14. Krug, K., Akerman, C.J. & Thompson, I.D. Responses of neurons in neonatal cortex and thalamus to patterned visual stimulation through the naturally closed lids. *J. Neurophysiol.* **85**, 1436–1443 (2001).
15. Rochefort, N.L. *et al.* Sparsification of neuronal activity in the visual cortex at eye-opening. *Proc. Natl. Acad. Sci. USA* **106**, 15049–15054 (2009).
16. Ikezoe, K., Tamura, H., Kimura, F. & Fujita, I. Decorrelation of sensory-evoked neuronal responses in rat barrel cortex during postnatal development. *Neurosci. Res.* **73**, 312–320 (2012).
17. Li, Y., Van Hooser, S.D., Mazurek, M., White, L.E. & Fitzpatrick, D. Experience with moving visual stimuli drives the early development of cortical direction selectivity. *Nature* **456**, 952–956 (2008).
18. Derrington, A.M. & Fuchs, A.F. The development of spatial-frequency selectivity in kitten striate cortex. *J. Physiol. (Lond.)* **316**, 1–10 (1981).
19. Berens, P. CircStat: a MATLAB toolbox for circular statistics. *J. Stat. Softw.* **31**, 1–21 (2009).
20. Bair, W., Zohary, E. & Newsome, W.T. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J. Neurosci.* **21**, 1676–1697 (2001).
21. Ecker, A.S. *et al.* Decorrelated neuronal firing in cortical microcircuits. *Science* **327**, 584–587 (2010).
22. Smith, M.A. & Kohn, A. Spatial and temporal scales of neuronal correlation in primary visual cortex. *J. Neurosci.* **28**, 12591–12603 (2008).
23. Clemens, J.M., Ritter, N.J., Roy, A., Miller, J.M. & Van Hooser, S.D. The laminar development of direction selectivity in ferret visual cortex. *J. Neurosci.* **32**, 18177–18185 (2012).
24. Shadlen, M.N. & Newsome, W.T. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* **18**, 3870–3896 (1998).
25. Kriener, B., Tetzlaff, T., Aertsen, A., Diesmann, M. & Rotter, S. Correlations and population dynamics in cortical networks. *Neural Comput.* **20**, 2185–2226 (2008).
26. Van Hooser, S.D. *et al.* Initial neighborhood biases and the quality of motion stimulation jointly influence the rapid emergence of direction preference in visual cortex. *J. Neurosci.* **32**, 7258–7266 (2012).
27. Jeanne, J.M., Sharpee, T.O. & Gentner, T.Q. Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron* **78**, 352–363 (2013).
28. Blankenship, A.G. & Feller, M.B. Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nat. Rev. Neurosci.* **11**, 18–29 (2010).
29. Huberman, A.D., Feller, M.B. & Chapman, B. Mechanisms underlying development of visual maps and receptive fields. *Annu. Rev. Neurosci.* **31**, 479–509 (2008).
30. Ackman, J.B. & Crair, M.C. Role of emergent neural activity in visual map development. *Curr. Opin. Neurobiol.* **24**, 166–175 (2014).
31. Wu, J.Y., Xiaoying, H. & Chuan, Z. Propagating waves of activity in the neocortex: what they are, what they do. *Neuroscientist* **14**, 487–502 (2008).
32. Sato, T.K., Nauhaus, I. & Carandini, M. Traveling waves in visual cortex. *Neuron* **75**, 218–229 (2012).
33. Ackman, J.B., Burbridge, T.J. & Crair, M.C. Retinal waves coordinate patterned activity throughout the developing visual system. *Nature* **490**, 219–225 (2012).
34. Siegel, F., Heimel, J.A., Peters, J. & Lohmann, C. Peripheral and central inputs shape network dynamics in the developing visual cortex in vivo. *Curr. Biol.* **22**, 253–258 (2012).
35. Stroh, A. *et al.* Making waves: initiation and propagation of corticothalamic Ca2+ waves in vivo. *Neuron* **77**, 1136–1150 (2013).
36. Wong, R.O., Meister, M. & Shatz, C.J. Transient period of correlated bursting activity during development of the mammalian retina. *Neuron* **11**, 923–938 (1993).
37. Hanganu, I.L., Ben-Ari, Y. & Khazipov, R. Retinal waves trigger spindle bursts in the neonatal rat visual cortex. *J. Neurosci.* **26**, 6728–6736 (2006).
38. Colonnese, M.T. *et al.* A conserved switch in sensory processing prepares developing neocortex for vision. *Neuron* **67**, 480–498 (2010).
39. Ko, H. *et al.* Functional specificity of local synaptic connections in neocortical networks. *Nature* **473**, 87–91 (2011).
40. Ko, H. *et al.* The emergence of functional microcircuits in visual cortex. *Nature* **496**, 96–100 (2013).
41. Le Magueresse, C. & Monyer, H. GABAergic interneurons shape the functional maturation of the cortex. *Neuron* **77**, 388–405 (2013).
42. Gao, W.J., Newman, D.E., Wormington, A.B. & Pallas, S.L. Development of inhibitory circuitry in visual and auditory cortex of postnatal ferrets: immunocytochemical localization of GABAergic neurons. *J. Comp. Neurol.* **409**, 261–273 (1999).
43. Van Hooser, S.D., Escobar, G.M., Maffei, A. & Miller, P. Emerging feed-forward inhibition allows the robust formation of direction selectivity in the developing ferret visual cortex. *J. Neurophysiol.* **111**, 2355–2373 (2014).
44. King, P.D., Zylberberg, J. & DeWeese, M.R. Inhibitory interneurons decorrelate excitatory cells to drive sparse code formation in a spiking model of V1. *J. Neurosci.* **33**, 5475–5485 (2013).
45. Bernacchia, A. & Wang, X.J. Decorrelation by recurrent inhibition in heterogeneous neural circuits. *Neural Comput.* **25**, 1732–1767 (2013).
46. Tetzlaff, T., Helias, M., Einevoll, G.T. & Diesmann, M. Decorrelation of neural-network activity by inhibitory feedback. *PLOS Comput. Biol.* **8**, e1002596 (2012).
47. van Vreeswijk, C. & Sompolinsky, H. Chaotic balanced state in a model of cortical circuits. *Neural Comput.* **10**, 1321–1371 (1998).
48. Renart, A. *et al.* The asynchronous state in cortical circuits. *Science* **327**, 587–590 (2010).
49. Renart, A. & van Rossum, M.C. Transmission of population-coded information. *Neural Comput.* **24**, 391–407 (2012).

ONLINE METHODS

Animals. All experimental procedures were approved by the Max Planck Florida Institute for Neuroscience or the Duke University Institutional Animal Care and Use Committee and were performed in accordance with guidelines from the US National Institutes of Health. Female ferret kits were obtained from Marshall Farms and were housed with a jill on a 16 h light/8 h dark cycle. Kits were examined daily to determine the date of eye opening.

Viral injections and GCaMP imaging. Microinjections of AAV2/1.hSyn.GCaMP3.WPRE.SV40 (ref. 50; obtained from University of Pennsylvania Vector Core) were made into the visual cortex approximately 7–14 d before imaging (range 6–22 d) using pulled glass pipettes and aseptic surgical technique. Anesthesia was induced with ketamine (50 mg/kg) and maintained with isoflurane (1–2%) and nitrous oxide (0–50% in oxygen). Atropine (0.2 mg/kg) was given at induction to reduce bronchial secretions. Animals were maintained at approximately 37 °C with a homeothermic heating blanket. Skin and muscle overlying visual cortex were reflected and a small burr hole was made with a hand-held drill (Fordom Electric Co.). Approximately 1 µL of virus was injected over 10 min using a Nanoject-II (WPI). Following the injection, muscle and skin were sutured closed and the animal was recovered and returned to its home cage.

After allowing time for expression of GCaMP3, animals were anesthetized as before, a tracheotomy was performed and an IV catheter was inserted into either the femoral vein or the external jugular vein. Animals were mechanically ventilated and heart rate and end-tidal CO_2 were monitored throughout the experiment. A metal headplate was implanted over the injected region and a craniotomy (~5 mm) was performed. Dura was resected and the brain was stabilized with 2% agarose and a coverslip.

For imaging, isoflurane was reduced to 0.5–1% and animals were paralyzed with vecuronium bromide (2 mg/kg/h in lactated Ringer's, delivered IV). This anesthetic regime produced highly stable heart rates of 280–300 b.p.m. for the duration of imaging, with end-tidal CO_2 levels stably maintained between 35–40 mm Hg. Phenylephrine (5%) and tropicamide (0.5%) were applied to the eyes to retract the nictitating membrane and dilate the pupil, and the cornea was protected with silicon oil.

Visual stimulation and two-photon imaging. Visual stimuli were delivered on an LCD screen placed approximately 25–30 cm in front of the eyes. Stimuli were full-field sinusoidal gratings at 100% contrast, at 0.06–0.08 cycles per degree, drifting at 4 Hz, presented at one of eight directions of motion. Stimuli were randomly interleaved and were presented for 5 s followed by a 5 s gray screen. A 5 s gray screen was used as a blank stimulus and was interleaved with grating stimuli. Stimuli were produced using either Matlab (The MathWorks Inc.) and Psychtoolbox[51,52] or PsychoPy[53].

Two-photon imaging was performed with an Ultima IV microscope (Prairie Technologies) driven by a Mai-Tai DeepSee laser (Spectra Physics) at 910 nm. Images 512 × 512 pixels were collected at 0.6–1.6 Hz. In a subset of experiments, imaging was performed using a resonant scanner (Prairie Technologies). In these experiments, rectangular images (512 pixels × 256 lines with a 1:2 aspect ratio) were collected at 60 Hz and downsampled to 15 Hz by averaging every 4 successive frames. Before analysis, these images were resized to 512 × 512 pixels via bilinear interpolation along the vertical dimension.

Data analysis. Data analysis was performed in ImageJ and Matlab, using MIJ (D. Sage, D. Prodanov, J. Tinevez and J. Schindelin, MIJ: making interoperability between ImageJ and Matlab possible, ImageJ User & Developer Conference, 24–26 October 2012, Luxembourg, http://bigwww.epfl.ch/sage/soft/mij/). In this study, we used fluorescent calcium sensors to report the activity of large populations of neurons simultaneously. The relative change in fluorescence of all three calcium sensors used in this study (GCaMP3, GCaMP6s and OGB) has been shown to be proportional to firing rate, and this relationship is roughly linear over the range of firing rates commonly observed in the visual cortex of juvenile ferrets[13,23,50,54–57]. To extract fluorescence traces from image stacks, ROIs were manually drawn around identified neurons and raw fluorescence for each frame was computed as the mean of all pixels in the ROI. Fluorescence traces were filtered with a first-order high-pass Butterworth filter with a cut-off time (TC) of 300 s. Similar results were obtained using $\Delta F/F$ with F_0 taken as the last 2 s of the inter-stimulus interval immediately preceding stimulus onset. Stimulus-evoked responses were taken as the average high-pass filtered fluorescence over the full stimulus interval. Neurons were considered visually responsive and analyzed further if the response to the preferred stimulus (averaged across trials) was both greater than zero and 2 s.d. above the blank response.

To determine the density of the population response, a neuron was considered responsive on a given trial if its response to any stimulus on that trial was greater than 2 s.d. above the response to a blank stimulus. Only neurons exhibiting at least one response during the imaging session (>90% of neurons) were included in this analysis, to prevent counting poorly labeled or unhealthy neurons as unresponsive. To examine the stimulus specificity of response density, we aligned all FOVs by the dominant stimulus—that which evoked a response in the largest fraction of neurons within the imaging field.

To compute tuning curves and determine the preferred direction (θ_{pref}), responses were averaged across trials and fit with a two-peaked Gaussian function, where the peaks were constrained to be 180° apart. Poorly fit (correlation between fit and actual responses with $P > 0.05$) or unresponsive neurons were excluded from further analysis. Orientation and direction selectivity were computed as described[17]

$$\mathrm{OSI} = \min\left(\frac{\mathrm{Pref}_{ori} - \mathrm{Orth}}{\mathrm{Pref}_{ori} - \mathrm{Blank}}, 1\right) \quad (1)$$

where

$$\mathrm{Pref}_{ori} = \mathrm{mean}(\mathrm{Resp}(\theta_{pref}), \mathrm{Resp}(\theta_{pref} + 180)) \quad (2)$$

$$\mathrm{Orth} = \mathrm{mean}(\mathrm{Resp}(\theta_{pref} + 90), \mathrm{Resp}(\theta_{pref} - 90)) \quad (3)$$

and

$$\mathrm{DSI} = \min\left(\frac{\mathrm{Pref}_{dir} - \mathrm{Opp}}{\mathrm{Pref}_{dir} - \mathrm{Blank}}, 1\right) \quad (4)$$

where

$$\mathrm{Pref}_{dir} = \mathrm{Resp}(\theta_{pref}) \text{ and } \mathrm{Opp} = \mathrm{Resp}(\theta_{pref} + 180) \quad (5)$$

Only neurons with significant orientation selectivity (Hotelling's t^2-test with $P < 0.05$; see ref. 17) were included for analysis of direction selectivity.

Pairwise noise correlations were calculated by first attempting to remove all signal correlations between neurons. This was achieved by filtering the data with a short cut-off time, first-order Butterworth filter (TC = 5 s), then *Z*-scoring the data individually for each stimulus with respect to the mean of that stimulus. The *Z*-scored responses (containing all data points acquired during the stimulus presentation) were then concatenated across all stimuli and the blank condition. Pairwise noise correlations were then computed using the Matlab corrcoef function. To prevent contamination of the fluorescence signal from neighboring neurons, we only considered pairs separated by at least 30 µm. Because of the slower scanning rate (0.5–0.7 Hz) for the training data, two cells were required to be sampled within 100 ms of each other to be included in the pairwise correlation analysis.

Nonparametric statistics were used unless noted, with Kruskal-Wallis (KW) tests followed by *post hoc* Mann-Whitney *U* tests (MW). Wilcoxon signed-rank tests (WSR) were used for paired comparisons. Data are presented as mean ± s.e.m.

Analysis of wave-like responses. Only experiments performed with a resonant scanner were used for the analysis of wave-like responses. To compute temporally colored images (**Fig. 3a**,**b**), images were *Z*-scored pixel-wise relative to the mean and s.d. of images collected during blank stimuli. Images were then smoothed with a mean filter (2 × 2 pixels in *x* and *y* and 2 frames (0.133 s) in *t*). Frames were then binned into 0.26-s intervals, averaged, and pseudocolored by time. Finally, pseudocolored images were summed across the full stimulus period. For calculating the average wave image (**Fig. 3c**), only trials with a significant linear wave (see below) were included.

To determine the presence of a linear wave, we used methods based on those of Siegel *et al.*[34]. First we found the peak response time (t_{peak}) for each responsive neuron on a given trial. For this analysis, a neuron was considered responsive if its fluorescence was >3 s.d. above the response to a blank stimulus for six successive

doi:10.1038/nn.3921

frames. Response onset was defined as the first time during the response that the fluorescence exceeded 1 s.d. above the blank. Peak response times were then fit with a linear traveling wave, where the response time (t_{fit}) is given by

$$t_{fit} = p/v + t_0 \quad (6)$$

where v is the velocity of the wave along its propagation direction and t_0 is the time the wave crossed the origin of the image coordinate system (positioned at the top left pixel). p is the projection of the neuron's x and y coordinates onto a unit vector describing the waves propagation angle:

$$p = (x, y) \cdot (\cos\theta, \sin\theta) \quad (7)$$

where x and y define a neuron's position and θ is the propagation angle of the wave. The error of the fit was defined as the sum of squared errors of t_{fit} relative to t_{peak}, and v, θ and t_0 were adjusted to minimize this error. For each stimulus, we then calculated a goodness of fit by shuffling t_{peak} 100 times, refitting each shuffled data set, and determining an error. The fraction of shuffled responses with errors larger than that of the original data set acts as a waviness index, giving an indication of the waviness of the response. (For a wave-like event, shuffling the response times should produce fits with higher errors.) Only responses with waviness indices >0.9 were considered as waves and included for determining velocity and propagation direction.

Given the frame rate of our microscope and the size of our FOVs, there was an upper limit to the wave velocity that we were able to observe. We estimate that this detection limit is faster than 1,500 μm/s. At such a velocity, a wave would traverse our FOV (512 pixels at 0.77 μm/pixel) in 4 imaging frames, ensuring that the wavefront position could be captured by at least three images. This threshold is approximately 6 times greater than the mean velocity (and 8 times greater than the median velocity) that we observed in naive animals (**Fig. 3g**).

If neurons are distributed in an elongated manner within a FOV, it is possible that our ability to detect waves propagating along the short axis could be reduced. If such an aperture effect is present in our data, then the ability to detect a wave traveling along the short axis should depend on the wave velocity, and fast moving waves should be under-represented along this axis. To determine whether such an effect is present in our data, we computed the velocities of waves traveling within 22.5° of either the long or short axis of a FOV. Considering only naive animals, in which both the frequency and velocity of waves were relatively high, there was no significant difference between velocity distributions for the two propagation directions (mean ± s.e.m. 310.0 ± 63.2 versus 281.2 ± 58.9 μm/s; n = 30 and 17, respectively; MW: P = 0.816). Furthermore, the asymmetry in wave propagation direction we report is present even if we discard from our analysis all waves that moved faster than the short-axis velocity detection limit (taken conservatively to be 400 μm/s; see below), showing that this effect cannot be from a bias in detectability.

To demonstrate the robustness of our detection algorithm, we performed a sensitivity analysis using simulated wavefronts. We constructed FOVs of 200 neurons with positions drawn from a two-dimensional Gaussian distribution with covariance similar to that of our actual data. Propagation directions were drawn from a uniform distribution of 0 to 2π and velocities drawn from a uniform distribution of 2,000–3,000 μm/s. To simulate a wave, response times were calculated for a subset of cells, with the number of cells participating in each wave drawn from a uniform distribution between 5 and 40, with a random selection of cells from within the 200-cell FOV for each wave. Detection is more difficult with low numbers of participating neurons, and this range covers the lower range observed in naive animals (mean ± s.d. 38 ± 28, with quartiles of 16, 31.5 and 53 cells per wave), where we required at least five active neurons to define a wave. After calculating the response times of participating cells, uniform noise (scaled to be 50% of the range of response times on a given wave) was added and response times were binned into intervals of 1/15 s to match the imaging frame rate. Each simulated wave was then fit, and the same detection threshold used for the actual data was applied. In simulated data, our algorithm readily detected waves traveling between 2,000 and 3,000 μm/s (84.5%, 845 of 1,000, with 5–40 active neurons per wave), suggesting that 1,500 μm/s is a conservative estimate of the upper limit for detection.

We also repeated our simulations using highly elliptical FOVs (mean ellipticity >0.84), with wave velocities drawn from a normal distribution with mean and s.d. of 1,000 μm/s. In these FOVs, our detection algorithm identified 92.7% (395 of 426) of waves traveling 1,000–2,000 μm/s, containing on average 23 active neurons per wave, including 87.8% (101 of 115) of waves traveling within 22.5° of the short axis. For waves traveling under 1,000 μm/s, the detection rate along the short axis increased to 98.1%. In our data set, FOVs had much lower ellipticity (<0.61) and the majority of waves had velocity much less than 1,000 μm/s (averaging approximately 250 μm/s). Thus, for realistic ellipticity and wave velocity, we are able to detect waves traveling in all directions and the asymmetries in propagation direction that we observe are not due to an inability to detect certain wave trajectories.

To determine the contribution of wave-like responses to noise correlations in the naive group, we calculated correlations using only trials with linear waves and compared these to correlations on trials with the same stimuli, but without waves. Only stimuli that evoked at least two waves were included in this analysis.

Longitudinal two-photon calcium imaging. Kits were implanted with a custom-designed headplate at P24–25. The headplate design accommodates an 8-mm-diameter coverslip (World Precision Instruments) held in place by a stainless steel retaining ring (McMaster Carr) that fits securely underneath a lip in the headplate, securing the coverslip in place and providing mild pressure (**Fig. 5a**). The headplate also features a protruding tab allowing the headplate to be clamped in a custom-built animal holder during surgery and imaging. The imaging chamber was designed to allow ready and repeated access to the cortex to remove any tissue or neomembrane[58] that may regrow over the imaging field.

During surgery, animals were anesthetized as above, except that animals were intubated and ventilated mechanically. Respiration parameters were adjusted to maintain end-tidal CO_2 at 35–40 mm Hg. During some portions of the surgery (durotomy and microinjection), animals were mildly hyperventilated to ~30 mm Hg to reduce edema. The skull overlying visual cortex was exposed, and a thin layer of cyanoacrylate (Vetbond, 3M) was applied and allowed to dry. The headplate was then positioned in place and attached to the skull with dental cement (C&B Metabond, Parkell Inc.). Once the cement had dried, a second layer of black cement (OrthoJet, Lang Dental mixed with powdered pigment: iron oxide, Dick Blick Art Supply) was applied. A craniotomy (~5 mm diameter) was performed over visual cortex and dura was carefully removed. In two animals (F1317 and F1319), AAV expressing GCaMP3 was injected as above. In two more experiments (F1473 and F1509) conducted after the release of GCaMP6, AAV1.Syn.GCaMP6s.WPRE.SV40 (ref. 57, obtained from University of Pennsylvania Vector Core) was injected in place of GCaMP3. Following the injection, warm agarose (2% in ACSF with Baytril (0.45 mg/mL)) was applied to the cortex and a glass coverslip was quickly inserted. The coverslip was secured with a retaining ring coated in silicone polymer (Kwik-Kast, World Precision Instruments) to seal the chamber. Following surgery, animals were recovered from anesthesia and returned to their home cages.

Approximately 1 week after this surgery, animals were returned to the imaging room, anesthesia was induced with ketamine, atropine was administered and anesthesia was maintained with isoflurane as above. Animals were intubated and ventilated, and an IV catheter was placed in the cephalic vein. In some imaging sessions, it was not possible to catheterize the cephalic vein; in these cases an IP catheter was inserted. If necessary, the chamber was opened under aseptic conditions, any regrown tissue or neomembrane was removed, agarose and a coverslip were reapplied and the chamber was resealed. This was usually necessary about 1–2 weeks after implantation and could typically be performed without causing apparent damage to the underlying cortex. In approximately half of the imaging sessions, the chamber remained optically clear and no tissue regrowth was apparent. Immediately before imaging, animals were paralyzed with vecuronium bromide (0.1 mg/kg/h in lactated Ringer's).

Two-photon imaging was performed as above and lasted approximately 3 h, containing approximately 45 min of visual stimulation. Imaging sessions were kept as short as possible to minimize the potential for inducing training effects, which can be observed in young ferrets following 4–6 h of continuous exposure to moving stimuli[17]. The imaged field of view could be approximately identified on the basis of surface blood vessel patterns, and two-photon z-stacks were taken from the cortical surface through the imaging plane to aid alignment. Following imaging, vecuronium was stopped and paralysis was antagonized with neostigmine (0.06 mg/kg), delivered with atropine (0.2 mg/kg). Isoflurane was discontinued, and animals were removed from the ventilator once spontaneous respiration

was observed. Sessions were repeated every 2–3 d until imaging quality degraded or the imaging FOV could not be conclusively identified.

Alignment of images across sessions was performed with an affine transform and data were analyzed as above. Only neurons that were visible and could be conclusively identified across all imaging sessions were included in this analysis.

To identify neurons with stable and reversing direction preferences, a bootstrapping analysis was performed as described by Li *et al*[17]. Neurons with less than a 10% likelihood of reversing preference were considered stable and neurons with >90% likelihood of reversing were considered reversing neurons. Neurons in which the initial direction preference was highly uncertain[17] (uncertainty > 10%) but became certain (uncertainty < 10%) by the final session were classified as having initially lacked a preferred direction but developed one over sessions. To compare average tuning curves across neurons, curves were aligned on the basis of the preferred direction on the final imaging session.

OGB data and analysis. Data from Li *et al.*[17] were analyzed for changes in noise correlations as described above. Briefly, in these experiments animals were imaged shortly following eye opening after bulk-loading the calcium indicator OGB1 (full methods available in ref. 17). Direction tuning curves were measured both before and after 4–6 h of motion training. Motion training consisted of repeated presentations of a grating drifting in a single orientation (the two directions of motion were interleaved). Another group of animals received flash training, in which a static grating stimulus was shown. As a control, a third group of animals received no training and viewed a static gray screen during the entire training period. To ensure continuity within the studied population, we have restricted analysis to neurons that could be conclusively identified both before and after motion training and that were visually responsive throughout the experiment. We assessed responsiveness by computing the signal-to-noise ratio: the size of response to the preferred direction relative to size of fluctuations during unstimulated periods of each cell. We restrict our analysis to a subpopulation of cells for which the distribution of signal-to-noise ratio did not change significantly over the training period. In this way, we insure that a net change in cell responsiveness did not underlie the change in mean correlation in the population.

Discrimination analysis. For a given set of neurons, we defined discriminability to be 1 minus the normalized overlap between the population response distributions using a Gaussian approximation, which captured the observed responses well (**Supplementary Fig. 9**). More concretely, consider a set of N neurons with $\mathbf{r}_{i,t}$ being the response of N neurons at trial $t = 1 \ldots N_t$ to stimulus $i = 1 \ldots N_S$. We computed $\bar{\mathbf{r}}_i$, the sample mean response to stimulus i, as

$$\bar{\mathbf{r}}_i = \sum_{t=1}^{N_t} \mathbf{r}_{i,t} \tag{8}$$

Covariance was assumed to be stimulus-independent and computed over all eight stimuli after subtracting stimulus means:

$$C = \frac{1}{N_t N_S} \sum_{i=1}^{N_S} \sum_{t=1}^{N_t} (\mathbf{r}_{i,t} - \bar{\mathbf{r}}_i)(\mathbf{r}_{i,t} - \bar{\mathbf{r}}_i)^{\mathrm{T}} \tag{9}$$

With the covariance matrix C and mean responses to stimuli i and j, the discriminability of stimulus pair $\{i,j\}$ is

$$D(i,j) = 1 - \frac{I_{\text{overlap}}(\{\bar{\mathbf{r}}_i, \bar{\mathbf{r}}_j\}, C)}{I_{\text{overlap}}(\{\langle \bar{\mathbf{r}} \rangle, \langle \bar{\mathbf{r}} \rangle\}, C)}$$

with

$$\langle \bar{\mathbf{r}} \rangle = \frac{\bar{\mathbf{r}}_i + \bar{\mathbf{r}}_j}{2} \tag{10}$$

where I_{overlap} is the integrated product of the stimulus-dependent response distributions:

$$I_{\text{overlap}}(\{\bar{\mathbf{r}}_i, \bar{\mathbf{r}}_j\}, C) = \int d^N \mathbf{r} \prod_{k=\{i,j\}} \frac{1}{\sqrt{(2\pi)^N \det C}} \exp\left(-\frac{1}{2}(\mathbf{r} - \bar{\mathbf{r}}_k)^T C^{-1} (\mathbf{r} - \bar{\mathbf{r}}_k)\right) \tag{11}$$

Here we assumed a multivariate Gaussian distribution of cell responses with covariance C. Equation (11) reduces to

$$D(i,j) = 1 - \exp\left(-\frac{1}{4}(\bar{\mathbf{r}}_i - \bar{\mathbf{r}}_j)^T C^{-1} (\bar{\mathbf{r}}_i - \bar{\mathbf{r}}_j)\right) \tag{12}$$

Large overlap yields a small discriminability. If the stimulus means $\bar{\mathbf{r}}_i$ are equal, the discriminability is 0. No overlap yields perfect discriminability (1).

To compute the shuffled discriminability, we computed the covariance matrix from trial-shuffled responses. Means $\bar{\mathbf{r}}_i$ were unchanged. The measure of discriminability used here is related to the sensitivity index d' (ref. 59). For a subset of the data, we computed discriminability using a low-rank approximation of the noise covariance matrix, which was obtained from the singular value decomposition (SVD) of the noise covariance matrix. Results were consistent with the full covariance matrix discriminability. In longitudinally imaged animals, discriminability was only assessed for neurons in which the stimulus orientation evoking the largest response did not change across sessions.

Variance and covariance exchange across data sets. To disentangle the effects of a shift in mean and a shift in variance over development, we computed direction discriminability for two hybrid cell sets: the first was composed by using the average activities (direction selectivities) from the naive data set but taking the variances from the immature set, and the second combined the variances of the naive set with the averages of the immature set. Since cells from the two data sets were not identical (typically from different animals), we assigned cell pairs across sets by matching the rank of their average activity or variance, respectively.

Specifically, we computed discriminability by drawing mean responses from a population of cells in naive animals and variances from populations of cells in immature animals. Mean responses were rank-matched to variance in the following manner. We drew from the available naive (N = 1,159 cells) and immature (N = 763 cells) populations equal-sized subsets of N = 700 cells. For each cell in the naive population, we computed the rank of the mean response to the preferred direction and the rank of the variance. This was repeated for the immature population. We then assigned cell pairs across sets to be rank-matched. For instance, if a cell in the naive population had mean response with rank m, it was matched with the variance of the cell in the immature population that had mean response with rank m. This gives the set of 'cells' with naive means and immature variances. For the reverse, naive variances with immature means, we took the nth-rank variance from the naive population and matched it with the mean response corresponding to the nth-rank variance in the immature population. This method ensures that the statistical relationship between mean activity and its variance is preserved.

The method for exchanging variance and covariance among subsets of cells is a two-step process: first, to match each group of cells from the naive set to a group of cells from the immature set, and second, to determine the cell-by-cell matching between each pair of groups. To construct groups of cells, we rank-matched as described above on the basis of the highest group-average response and the determinant of the covariance matrix. Within each group of N cells, we used rank matching based on the N mean responses to the preferred direction and on the N individual-cell variances. Effectively, this reorders the rows and columns of the covariance matrix. At this stage, we can compute discriminability with naive means and immature variance and covariance. To separate covariance and variance, we computed the correlation matrix r_{ij} from the (immature) covariance matrix (C_{ij}) using immature single-cell variances (σ_i^I)

$$r_{ij} = \frac{C_{ij}}{\sigma_i^I \sigma_j^I} \tag{13}$$

and transformed this back to a covariance matrix using the individual cell variances from the naive data set (σ_i^N):

$$C'_{ij} = r_{ij} \sigma_i^N \sigma_j^N \tag{14}$$

The matrix C'_{ij} was used to compute discriminability with immature correlation structure and naive mean and variance.

doi:10.1038/nn.3921

In the training data set, we had the same populations of cells, so rank matching was not required to match pre- and post-training covariance and means. Variance and covariance were separated as described above, by first converting to a correlation matrix and then using pre-training single-cell variance to compute the covariance matrix from the correlation matrix.

A **Supplementary Methods Checklist** is available.

50. Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nat. Methods* **6**, 875–881 (2009).
51. Brainard, D.H. The Psychophysics Toolbox. *Spat. Vis.* **10**, 433–436 (1997).
52. Pelli, D.G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* **10**, 437–442 (1997).
53. Peirce, J.W. PsychoPy–Psychophysics software in Python. *J. Neurosci. Methods* **162**, 8–13 (2007).
54. Ohki, K., Chung, S., Ch'ng, Y.H., Kara, P. & Reid, R.C. Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature* **433**, 597–603 (2005).
55. Kara, P. & Boyd, J.D. A micro-architecture for binocular disparity and ocular dominance in visual cortex. *Nature* **458**, 627–631 (2009).
56. Hendel, T. *et al.* Fluorescence changes of genetic calcium indicators and OGB-1 correlated with neural activity and calcium in vivo and in vitro. *J. Neurosci.* **28**, 7399–7411 (2008).
57. Chen, T.W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
58. Chen, L.M. *et al.* A chamber and artificial dura method for long-term optical imaging in the monkey. *J. Neurosci. Methods* **113**, 41–49 (2002).
59. Averbeck, B.B. & Lee, D. Effects of noise correlations on information encoding and decoding. *J. Neurophysiol.* **95**, 3633–3644 (2006).