The photosensitive phase acts as a sensitive window for seasonal multisensory neuroplasticity in male and female starlings

Jasmien Orije¹, Emilie Cardon¹, Julie Hamaide¹, Elisabeth Jonckers¹, Veerle M. Darras², Marleen Verhoye¹, Annemie Van der Linden¹

¹ Bio-Imaging Lab, University of Antwerp, Belgium

² Laboratory of Comparative Endocrinology, Biology Department, KU Leuven, Belgium

Abstract

Traditionally, research unraveling seasonal neuroplasticity in songbirds has focused on the male song control system and testosterone. We longitudinally monitored the song and neuroplasticity in male and female starlings during multiple photoperiods using Diffusion Tensor and Fixel-Based techniques. These exploratory data-driven whole-brain methods resulted in a population-based tractogram uncovering microstructural sexual dimorphisms in the song control system and beyond. Male brains showed microstructural hemispheric asymmetries, whereas females had higher interhemispheric connectivity, which could not be attributed to brain size differences. Only females with large brains sing but differ from males in their song behavior by showing involvement of the hippocampus. Both sexes experienced multisensory neuroplasticity in the song control, auditory and visual system, and the cerebellum, mainly during the photosensitive period. This period with low gonadal hormones might represent a ‘sensitive window’ during which different sensory and motor systems in telencephalon and cerebellum can be seasonally re-shaped in both sexes.

1 ABBREVIATIONS

A	Arcopallium
AD	Axial diffusivity
CA	<i>Commissura anterior</i> / anterior commissure
CbL	<i>Nucleus cerebellaris lateralis</i> / lateral cerebellar nucleus
CO	<i>Chiasma opticum</i> / optic chiasm
CP	<i>Commissura posterior</i> / posterior commissure
CSt	Caudal part of the lateral striatum
DLM	<i>Nucleus dorsolateralis anterior thalami pars medialis</i> / medial part of the dorsolateral nucleus of the anterior thalamus
DMA	Dorsomedial nucleus of the anterior thalamus
DMP	Dorsomedial nucleus of the posterior thalamus

DSD	<i>Decussatio supraoptica dorsalis</i>
E	Entopallium
FA	Fractional anisotropy
FC	Fiber-bundle cross-section
FD	Fiber density
FOD	Fiber orientation distribution
FPL	<i>Fasciculus prosencephali lateralis</i> / lateral forebrain bundle
GP	Globus pallidus
HA	<i>Hyperpallium apicale</i> / apical part of the hyperpallium
HD	<i>Hyperpallium densocellulare</i> / densocellular part of the hyperpallium
HVC	Used as a proper name (High vocal center)
LAD	<i>Lamina arcopallialis dorsalis</i> / dorsal arcopallial lamina
LaM	<i>Lamina mesopallialis</i> / mesopallial lamina
LFM	<i>Lamina frontalis suprema</i> / supreme frontal lamina
LFS	<i>Lamina frontalis superior</i> / superior frontal lamina
LMAN	<i>Nucleus lateralis magnocellularis nidopallii anterioris</i> / lateral magnocellular nucleus of the anterior nidopallium
LPS	<i>Lamina pallio-subpallialis</i> / pallial-subpallial lamina
LSt	<i>Striatum laterale</i> / lateral striatum
M	Mesopallium
MD	Mean diffusivity
MLD-DM-ICO	Intercollicular nucleus complex
MMAN	<i>Nucleus medialis magnocellularis nidopallii anterioris</i> / medial magnocellular nucleus of the anterior nidopallium
MRI	Magnetic Resonance Imaging
MSt	<i>Striatum mediale</i> / medial striatum
N	Nidopallium
NCM	<i>Nidopallium caudomediale</i> / caudomedial nidopallium
OM	<i>Tractus occipito-mesencephalicus</i> / occipito-mesencephalic tract
Ov	Nucleus ovoidalis
POA	<i>Area praeoptica</i> / pre-optic area
Pt	<i>Nucleus pretectalis</i> / pretectal nucleus
RA	<i>Nucleus robustus arcopallialis</i> / robust nucleus of arcopallium
RARE	Rapid acquisition with relaxation enhancement
RD	Radial diffusivity
Rt	<i>Nucleus rotundus</i>
surr	<i>surroundings</i>
TeO	<i>Tectum opticum</i> / optic tectum
tFA	<i>Tractus fronto-arcopallialis</i> / fronto-arcopallial tract
TFM	<i>Tractus thalamo-frontalis</i>
TnA	<i>Nucleus taeniae amygdalae</i> / nucleus taeniae of the amygdala
TrO	<i>Tractus opticus</i> / optic tract
TSM	<i>Tractus septopallio-mesencephalicus</i> / septopallio-mesencephalic tract

2 INTRODUCTION

Although various songbird species demonstrate different levels of female song, in one of the most studied song bird species, the zebra finch, female song is absent (Odom et al., 2014). Observations like these led to the discovery of one of the largest sexual dimorphisms in the brain of vertebrate species (Bernard et al., 1993; Nottebohm and Arnold, 1976). Such sexual dimorphisms in singing behavior and brain structure drove a bias in research of the song control system towards male songbirds. Besides the large sexual dimorphism in song and song control nuclei, seasonal songbirds expose an additional peculiarity as they display naturally re-occurring seasonal cycles of singing and neuroplasticity of the song control system. However, the male focus proceeded in the research of seasonal neuroplasticity, since this process was shown to be largely driven by photoperiod-induced increases in testosterone, which results in a pre-optic area (POA) mediated increase in motivation to sing and subsequent singing activity induced neuroplasticity (Alward et al., 2013). The largest differences in song control nuclei volumes are found between breeding and non-breeding season (Riters et al., 2002). This led seasonal neuroplasticity studies to focus mainly on the breeding season or photostimulated phase, which they often compare to the non-breeding season or photorefractory and photosensitive phases. The main emphasis on the role of photoperiodic-induced increase in testosterone has led to many manipulation studies involving castration and/or testosterone implantation and its effect on song production and the song control system (e.g. (Hall and Macdougall-Shackleton, 2012; Orije et al., 2020; Stevenson and Ball, 2010)).

Prior studies have also pointed out that testosterone implants only have an effect when the birds have experienced a photosensitive phase with short day length exposure (Bernard and Ball, 1997; Rouse et al., 2015). Although little research is devoted to brain neuroplasticity during this state (Larson et al., 2019; Riters et al., 2002), multiple studies have shown that during the photosensitive period, testes size (Cornez et al., 2017) and GnRH expression in the POA slowly start to increase (Bentley et al., 2013; Hurley et al., 2008; Stevenson et al., 2012), while some song control nuclei slowly start to grow (Riters et al., 2002). The photosensitive period might be seen as a sensitive window, defined as a period in which experiences lead to long-lasting changes (that remain in the subsequent photostimulated phase or breeding season) in the brain and song behavior (Knudsen, 2004). The photosensitive period plays a crucial role in creating permissive circumstances for enhanced neuroplasticity resulting in enlarged song control nuclei as observed during the photostimulated stage.

The studies on sexual dimorphisms and seasonal neuroplasticity in songbirds have been preoccupied with vocal control areas as the neural substrate for the behavioral sexual dimorphism in song performance. However, *in vivo* Magnetic Resonance Imaging (MRI) studies from our team focusing on

the entire songbird brain have demonstrated seasonal neuroplasticity and functional changes in both auditory and olfactory systems of starlings (De Groof et al., 2017; De Groof et al., 2010; De Groof et al., 2013). This outcome hints towards a seasonal re-occurring cycle of multisensory neuroplasticity and raises the question whether the photosensitive stage perhaps represents a renewed ‘sensitive window’ for heightened neuroplasticity and rewiring of various sensory and sensorimotor systems.

The current study aims to fill the gaps highlighted above and answers the following questions with an extensive data-driven exploratory brain-wide structural MRI study of both male and female starlings throughout different seasons: (1) Is **sexual dimorphism** limited to the song control system in starlings, as in zebra finches where only the male sings (Hamaide et al., 2017) or does this dimorphism extend to other sensory systems? Can it be attributed to a difference in total brain size between the sexes, similar to what has been observed in mammals? (2) Do female and male starlings display parallel or different **seasonal structural neuroplasticity**? Is this neuroplasticity limited to the song control system or does it extend to other neural circuits? When do these (neuroplasticity) changes take place, already during the photosensitive period when gonadal hormone levels are still very low? (3) How does seasonal neuroplasticity correlate -at the group and individual level - with **gonadal hormone levels and with song-production**?

Since MRI is an *in vivo* method, we were able to measure the same subjects repeatedly and monitor the structural neuroplasticity longitudinally in thirteen male and twelve female starlings as they undergo different subsequent photoperiodic phases: at baseline photorefractory state (PR), after four, eight and twelve weeks of short days during the photosensitive state (SD4, SD8, SD12), after four weeks of photostimulation (LD4) and after sixteen weeks on long days when the birds are photorefractory again (LD16) (figure 1A). At each time point we: collected blood samples; measured body weight; recorded song; and acquired structural MRI data including 3D anatomical scans and diffusion-weighted images (DWI). Using voxel-based diffusion tensor imaging (DTI) and fixel-based analysis (De Groof et al., 2006; Raffelt et al., 2017), we established spatio-temporal statistical maps revealing microstructural changes in grey and white matter, specifically in fiber connections throughout the different seasons in the entire brain in both sexes. To learn about the contribution of singing activity induced neuroplasticity in the seasonal neuroplasticity, voxel-based multiple regression between seasonal song rate and corresponding DTI parameter maps was performed to identify the neuronal correlates of song behavior in both sexes (Hamaide et al., 2020; Sagi et al., 2012).

Figure 1. Simplified overview of the experimental setup and various processing steps of the DTI and fixel-based analysis. (A) Male and female starlings were measured repeatedly as they went through different photoperiods. At each time point, their songs were recorded, blood samples were collected and T2-weighted 3D anatomical and diffusion-weighted images (DWI) were acquired. The 3D anatomical images were used to extract whole-brain volume. (B) Using the DWI, we performed a common DTI analysis and a fixel-based analysis. By doing this, the extracted DTI metrics (such as fractional anisotropy (FA)) and fixel-based metrics (fiber density (FD) and fiber-bundle cross-section (FC)) complemented each other during interpretation of the results. Insets present an example DWI image at each respective processing step. (C) Large insets are at the level of the commissura anterior and OM tract and show the fiber orientation distribution (FOD) and Diffusion Tensor map used in fixel-based and DTI-based analysis respectively as described in table 1 colored according to the standard red-green-blue code (red = left-right orientation (L-R), blue = dorso-ventral (D-V) and green = rostro-caudal (R-C)). (D) All measures were normalized to a population template, a population-based tractogram and fixel template mask were created which are necessary for the voxel- and fixel-based statistical analysis (E). The material and method section contains the extended explanation on the various processing steps. Abbreviations: LD, long days; PR, photorefractory state; SD short days.

3 RESULTS

3.1 POPULATION-BASED TRACTOGRAM OF THE STARLING BRAIN USING FIXEL-BASED ANALYSIS

In the current study, we analyzed the DWI scans in two distinct ways: 1) using the common approach of diffusion tensor derived metrics such as fractional anisotropy (FA) and; 2) using a novel method of **fiber orientation distribution (FOD)** derived fixel-based analysis. Both techniques infer the microstructural information based on the diffusion of water molecules, but they are conceptually different (table 1). Figure 1B-E gives a comprehensive overview of the processing steps (see material and method section for detailed outline). Common DTI analysis estimates for each voxel the diffusion tensor from which several parameters can be extracted, which are sensitive to various microstructural changes specified in table 1. Since this is a brain-wide voxel-based method, both white and grey matter can be investigated. Fixel-based analysis on the other hand relies on a population-based tractogram generated from FOD information and is therefore inherently sensitive only to changes in white matter structures; both microscopic changes in apparent fiber density (FD) or macroscopic changes in fiber-bundle cross-section (FC) (table 1). As a voxel can hold multiple fiber bundles, FD is estimated by the intensity of the FOD before reorientation for each fiber bundle separately (illustrated by FD1 and FD2 in table 1). This allowed us to examine crossing fibers within a voxel, which is a major advantage of

the fixel-based analysis. FC is based on the transformation parameters or Jacobian determinants perpendicular to the fixel orientation. For group statistical analysis, the FC values were logarithmically transformed to log FC in order to center the data on zero and ensure a normal distribution. Positive values indicate expansion, whereas negative values reflect shrinkage of a fiber bundle relative to the template (Raffelt et al., 2017).

Table 1. Overview of the parameters studied by common DTI analysis and novel fixel-based analysis, including a comprehensive, but not exhaustive, list of some of the microstructural changes to which they are sensitive.

	Metric	Level	Parameter	Measures	Sensitive to changes in :
DTI analysis		Voxel	FA	Directionality of water diffusion	Axon number and density, axon diameter, myelination, fiber organization
			MD	Average diffusion across all directions	Cell size, cell spacing, cell density, dendrite branching, extracellular space
			AD	Diffusion along the fiber direction	Axon number and density
			RD	Diffusion perpendicular to the fiber direction	Axon diameter and myelination
fixel-based analysis		Fixel	FD	Microscopic density within fiber population	Axon number and density, axon diameter
			FC	Macroscopic change in cross-sectional area perpendicular to fiber bundle	Fiber bundle size due to changes in extra-axonal space and/or myelination, number of axons

Abbreviations: FA, fractional anisotropy; MD, mean diffusivity; AD, axial diffusivity; RD, radial diffusivity; FD, fiber density; FC, fiber-bundle cross-section. Fixel-based analysis can distinguish between different fiber bundles within a voxel and estimate FD for each fiber bundle separately (FD1 and FD2). For a complete description and interpretation of these measures we refer to the following reviews of DTI analysis (Beaulieu, 2002; Beaulieu, 2014; Song et al., 2002; Zatorre et al., 2012) and fixel based analysis (Genc et al., 2020; Raffelt et al., 2017)

It is important to note that fixel-based analysis was mainly created to study the human or mammalian brain. The mammalian cortex has a laminar organization with large white matter tracts interconnecting both hemispheres through the corpus callosum, whereas the avian telencephalon has a nuclear organization with distinct white matter tracts organized in lamina (Jarvis et al., 2005). Avian interhemispheric connections run mainly through the commissure anterior and posterior, as the corpus callosum is an exclusively mammalian feature (Suarez et al., 2018). Certain connections

between avian nuclei cross the grey matter like pallium and organize in bundles such as the HVC-RA tract (Karten et al., 2013). Since the connections within the avian brain are not organized in large white matter tracts, but in lamina and sparse fibers crossing brain structures, a lower FOD threshold is applied to ensure continuity in our fiber tracking. This holds the potential risk of introducing noisy fixels, especially within the grey matter, which we will keep in mind during interpretation of the results. However, a higher threshold may exclude many of the genuine white matter tracts, which is unfavorable. The result is a population-based template in which many of the avian anatomical structures can be identified (figure 2). We recognize many of the white matter structures such as the different lamina, occipito-mesencephalic tract (OM) and optic tract (TrO) among others. Interestingly, many of the nuclei within the song control system (i.e. HVC, robust nucleus of the arcopallium (RA), lateral magnocellular nucleus of the anterior nidopallium (LMAN), and Area X), auditory system (i.e. intercollicular nucleus complex, nucleus ovoidalis) and visual system (i.e. entopallium, nucleus rotundus) are identified by the empty spaces between tracts. The population-based tractogram provides a clear delineation of various brain regions and highlights the connections throughout the brain. The applied fixel-based approach cannot report on the microstructure within different brain nuclei; but rather sheds light on the fiber tracts surrounding and interconnecting them. As such, it provides an excellent tool to investigate neuroplasticity of different brain networks without focusing on changes within the nuclei, but still allowing defining them through their surrounding fibres, referred to as HVC surr, RA surr and Area X surr.

159

160

Figure 2. Overview of the population based tractogram of male and female starlings over the seasons created for the fixel-based analysis with indications of different lamina, interconnecting tracts, nuclei and brain regions displayed on axial (A) and coronal slices (B) throughout the brain. The different intervals of the coronal and axial sectioning are indicated on the sagittal inset in the left corner of each panel. The coronal slices do not follow a strict interval to visualise small nuclei such as DLM and Ov. The crosshair indicates the color-coding of the tractogram following the standard red-green-blue code (red = left-right orientation (L-R), blue = dorso-ventral (D-V) and green = rostral-caudal (R-C)). For abbreviations see abbreviation list.

3.2 DATA-DRIVEN ANALYSIS OF SEXUAL DIMORPHISMS IN BRAIN STRUCTURE

We established the general structural differences between male and female starling brains by applying a full-factorial design voxel-based analysis to our DTI data (figure 3, 4, Table 2). This includes data from all time points with time as a within-subject factor that needs to be averaged over (McFarquhar, 2019). This analysis confirmed the well-established sexual dimorphism of vocal control areas in songbirds (Bernard et al., 1993; Hamaide et al., 2017; Nottebohm and Arnold, 1976), where male starlings have higher FA in several parts of the song control system (the surroundings of HVC, Area X and RA) and auditory system (Field L and caudomedial nidopallium (NCM)) compared to female starlings. Interestingly, even though most differences occur bilaterally, these structures are visible as sub peaks within a single significant cluster. This cluster was asymmetric towards the left hemisphere and comprised large parts of the mesopallium and nidopallium, but not the lamina mesopallialis (LaM) which divides these regions. This means that not only the fibers surrounding the song control nuclei display different FA values, but also a larger part of the left hemisphere. In general, FA provides a correlate for a set of underlying microstructural characteristics related to structure and spatial organization of axons, dendrites and even cells, but cannot pinpoint a single underlying cellular event, as this requires dedicated histological studies (Sagi et al., 2012; Zatorre et al., 2012). Importantly, females generally have higher FA in tracts interconnecting both hemispheres (commisura anterior and posterior) and in bilateral tracts interconnecting nuclei within a hemisphere including the TrO, OM, septopallio-mesencephalic tract (TSM) and fronto-arcopallial tract (tFA). In addition, females have higher FA within the cerebellum and caudal part of the lateral striatum (CSt) compared to males. These findings are further confirmed by fixel-based analysis; as many of the sexual dimorphisms in FA match with the FD statistical analysis (figure 4 A, B). Certain subregions of the cerebellum, the molecular layer of cerebellar lobules II, III, IV have higher FD in females, whereas the cerebellar lobule VIII has higher FD in males. This diversifies the finding of increased FA in the cerebellum of female starlings. The molecular layer consists of parallel fibers emitted by the granule cells of the cerebellar cortex (D'Angelo, 2018). Fixel-based analysis is able to detect significant changes in this highly organized fibre structure, which is depicted as the red parallel fibers running left to right in figure 4 – figure supplement 1B. In addition to these microscopic changes, log FC statistical analysis revealed some macroscopic differences between male and female starlings (figure 4 C, D). Male starlings generally have a larger volume in the rostral part of the left hemisphere at the level of the hyperpallium apicale, which further supports the asymmetry towards the left hemisphere. Furthermore, the left HVC surroundings and cerebellar lobule VIII are larger in males compared to females. However, certain regions within the superior frontal lamina (LFS), nidopallium and mesopallium are larger in females, compared to males.

Alongside the well-established difference in the song control system, the applied data-driven approach sheds new light on sexual dimorphisms in the rest of the starling brain, including a microstructural hemispheric asymmetry towards the left hemisphere in males and a more structurally organized cerebellum as well as several white matter tracts in females.

Figure 3: Overview of the general structural sexual dimorphism within the starling brain attributed to differences in fractional anisotropy (FA). For each sex comparison, the statistical parametric maps are displayed on axial and coronal sections throughout the brain (upper and lower row respectively). The results are displayed with $p_{\text{uncorr}} < 0.001$ and $k_E \geq 10$ voxels, and overlaid on the population FA map. The T-values are colour-coded according to the scale on the right. In males, FA was higher in a large cluster lateralized to the left hemisphere and covering parts of the nidopallium and mesopallium, including several regions of the auditory system and surrounding the song control system. Females, on the other hand, have higher FA at the level of the cerebellum and several tracts such as OM, TSM and TrO. For abbreviations see abbreviation list.

Table 2: Clusters displaying a sex difference in fractional anisotropy (FA)

Main sex difference for FA	Cluster	Hemisphere	Cluster		Peak	
			p_{FWE}	k_E	p_{FWE}	T
male > female	Nidopallium / mesopallium	Left	0.000	36936	0.000	14.25
	HVC surr	Left			0.000	8.35
		Right			0.000	7.53
	Field L dorsal	Left			0.000	10.00
		Right			0.000	6.93
	NCM	Left			0.000	7.72
		Right			0.000	6.72
	Area X caudal surr	Left			0.001	5.69
		Right			0.001	6.07
	RA surr	Left			0.000	7.86
	HVC-RA tract	Left	0.000	441	0.000	6.68
	RA surr	Right			0.001	6.08
	HVC-RA tract	Right			0.000	6.42

	TeO (superior)	Left	0.000	189	0.000	9.44
		Right	0.000	125	0.000	7.47
	LaM	Right	0.000	165	0.000	6.97
	Cerebellum	Lobule VII	0.011	101	0.000	6.54
female > male	TrO	Right	0.000	24846	0.000	12.08
		Left			0.000	8.84
	TSM	Left			0.000	9.62
		Right			0.000	9.74
	Commisura anterior	Center			0.222	4.85
	Commisura posterior	Center			0.066	5.19
	OM	Left			0.000	8.77
		Right			0.000	7.78
	CSt	Left			0.000	8.50
		Right			0.000	7.64
	Cerebellum	Lobule VIII			0.000	9.12
		Lobule VII			0.000	9.68
		Lobule VI			0.000	7.77
		Lobule V			0.000	7.12
		Lobule IV			0.000	6.58
	tFA	Right			0.000	10.50
		Left	0.006	113	0.000	8.64
	Area X	Left	0.139	57	0.000	6.34
		Right	0.000	254	0.000	6.91

Table summarizing the significance at cluster level and peak level. Some clusters are large and cover multiple regions, indicated by the single statistic at cluster level. For each sub-region within this large cluster the peak significance (pFWE) and T-value (T) are reported. P-values are FWE corrected. K_E indicates the number of continuous voxels within a cluster. For abbreviations see abbreviation list.

Figure 4: Overview of the general structural sexual dimorphism within the starling brain attributed to microscopic differences in fiber density (FD) (A, B) and macroscopic differences in fiber-bundle cross-section (log FC) (C, D). For each sex comparison, the statistical parametric maps are displayed on axial and coronal sections throughout the brain (upper and lower row respectively). The results are displayed with $p_{\text{uncorr}} < 0.001$, and overlaid on the population based tractogram. Only significant tracks are displayed in a color representing the T-value. For abbreviations see abbreviation list.

interconnectivity found in females. Even though there was a small difference in brain size between male and female starlings (respectively, Mean male= 1815, SE= 8 mm³ and Mean female= 1770, SE= 13 mm³), it was not significant ($F(1, 23)=1.78$, $p=0.195$). Next, we artificially divided each sex in two groups based on their brain size using a median split similar to the analysis in Kurth et al. (2018) (respectively, females: N=6 large and N=6 small, males N=6 large, N=7 small).

These artificially generated groups were used for a voxel-based analysis with brain size as a fixed factor to assess the effect of brain size on the FA, FD and log FC analysis (figure 5, 6, figure 6- figure supplement 1, table 5). This revealed that only a few sex differences could be attributed to a difference in brain size. Large brain male and female starlings had higher FA values in specific sections of the caudal surroundings of the right Area X, surroundings of RA and HVC compared to small brain male and female starlings (figure 5A). The brain size difference in FA in right Area X surroundings and cerebellum was matched with a difference in FD (figure 6A), whereas the higher FA in the surroundings of HVC of large brain starlings, is complemented by a higher log FC in the surroundings of HVC (figure 6C).

Small brain starlings had higher FA in left NCM, a small part of the right LaM and TrO (figure 5A), which closely matches the FD brain size difference in the left NCM and right LAM (figure 6B) and the log FC difference in TrO (figure 6D). Following the hypothesis of neuronal interconnectivity, the voxel-based statistical analysis for brain size is in line with the statistical analysis for sex at the level of the TrO, but not in the NCM where both male and small brain starlings exhibit high FA. This might be partially explained by the finding of higher log FC values in the NCM of large compared to small brain starlings (figure 6C). Furthermore, some regions such as the hippocampus in large brain starlings and ventricle in small brain birds were not detected in the voxel-based sex difference analysis and are solely attributed to a difference in brain size. In the log FC analysis for brain size differences, more regions unrelated to a sex difference showed a significant larger log FC value in large brains compared to smaller brains, including parts of the pallial-subpallial lamina, lateral forebrain bundle, medial and lateral surroundings of entopallium.

Interestingly, many of the fiber tracts that had a pronounced sex difference, such as the anterior and posterior commissure, TSM and OM, were not found in the brain size statistical maps. Furthermore, the microstructural hemispheric asymmetry covering the nidopallium and mesopallium detected in the voxel-based sex difference analysis can also not be attributed to a mere difference in brain size. Importantly, further statistical analysis is required to ensure that the brain size effect is not confounded by other factors such as song behavior differences.

Table 3: Clusters displaying a difference in brain size versus a difference in sex in fractional anisotropy (FA)

	Cluster	Hemisphere	Cluster		Peak		Relevant difference
			p_{FWE}	k_E	p_{FWE}	T	
Large > small brain	HVC surr	Left	0.000	227	0.000	6.43	M>F
	Hippocampus	Left	0.000	155	0.000	5.63	↑Singing F
	LaM/LFS	Left			0.000	7.51	
	HVC surr	Right	0.000	417	0.000	6.61	M>F
	Hippocampus	Right			0.000	6.94	
	Area X caudal surr	Right	0.000	190	0.002	5.99	M>F ↑Singing F
	LPS	Left	0.004	126	0.006	5.78	
	LPS/LFS	Right	0.000	337	0.042	5.29	
	tFA	Right	0.925	20	0.029	5.39	F>M
	TrO	Right	0.001	160	0.000	7.29	F>M
	RA surr	Left	0.000	451	0.193	4.88	M>F
	Cerebellum	Lobule VI			0.000	7.50	
	Cerebellum	Lobule VI	0.259	48	0.010	5.64	
	Cerebellum	Lobule IV	0.114	62	0.003	5.94	
	Cerebellum	Lobule VII	0.101	64	0.041	5.30	
Small > large brain	NCM	Left	0.006	116	0.000	6.39	M>F
	LaM	Right	0.014	100	0.025	5.43	
	Ventricle rostral	Left	0.217	51	0.000	7.21	
		Right	0.000	246	0.000	6.83	
	Ventricle caudal	Center	0.002	138	0.003	5.90	
		Center	0.004	122	0.214	4.84	
	TrO	Center	0.011	105	0.232	4.82	F>M

Table summarizing the significance at cluster level and peak level. Some clusters are large and cover multiple regions, indicated by the single statistic at cluster level. For each sub-region within this large cluster the peak significance (p_{FWE}) and T-value (T) are reported. P-values are FWE corrected. k_E indicates the number of continuous voxels within a cluster. The final column reports if a region with a difference in brain size is analogue to other relevant differences such as the general sexual dimorphism in FA (M>F or F>M) or the correlation between song rate and FA in females (↑Singing F).

Figure 5: Overview of general structural difference between large and small brain starling in fractional anisotropy (FA). The statistical parametric maps are displayed on axial and coronal sections throughout the brain (upper and lower row respectively). The results are displayed with $p_{\text{uncorr}} < 0.001$ and $k_E \geq 10$ voxels, and overlaid on the population FA map. The T-values are colour-coded according to the scale on the right.

Figure 6: Overview of general structural difference between large and small brain starling in microscopic differences in fiber density (FD) (A, B) and macroscopic differences in fiber-bundle cross-section (log FC) (C, D). The statistical parametric maps are displayed on axial and coronal sections throughout the brain (upper and lower row respectively). The results are displayed with $p_{\text{uncorr}} < 0.001$, and overlaid on the population-based tractogram. Only significant tracks are displayed and colored according to their significance by T-value.

Figure 6- figure supplement 1: Overview of general structural difference between large and small brain starling in microscopic differences in fiber density (FD) and macroscopic differences in fiber-bundle cross-section (log FC) using fiber orientation color coding. The statistical parametric maps are displayed on axial and coronal sections throughout the brain (upper and lower row respectively). The results are displayed with $p_{\text{uncorr}} < 0.001$, and overlaid on the population based tractogram. Only significant tracks are displayed and are colored according to the fiber orientation using the standard red-green-blue code as indicated by the crosshair (red = left-right orientation (L-R), blue = dorso-ventral (D-V) and green = rostro-caudal (R-C)). For abbreviations see abbreviation list.

3.4 SEXUAL DIMORPHISM IN SEASONAL NEUROPLASTICITY

After establishing the general structural sexual dimorphisms, we investigated whether seasonal neuroplasticity occurs differently between sexes. Therefore, we performed a voxel-based flexible factorial analysis looking at the general FA changes over time and the interactions between time and sex. These results are summarized in figure 7 and table 4. We extracted the mean FA values from the significant clusters and plotted them to determine how FA changes over time (figure 8). FA changes can be caused by a multitude of underlying microstructural changes, however FA is most sensitive to changes in axon number and density, axon diameter, myelination (see table 1). Additionally, we extracted the mean of the other diffusion parameters (MD, AD and RD) and fixel-based measures from these ROIs to provide more insight into the basis of the FA change (figure 8 - figure supplements 1-5). For the statistical analysis of the extracted parameters, we conducted a linear mixed model analysis with Tukey's Honest Significant Difference (HSD) multiple comparison correction during *post hoc* statistics.

3.4.1 Sexual dimorphism in seasonal neuroplasticity is lateralized and restricted to specific regions of the song control and auditory system

Voxel-based analysis picked up a significant interaction between sex and time in the surroundings of HVC (bilateral), in a rostro-lateral part of left RA surroundings and in left NCM (figure 7, 8). At the level of HVC surroundings and left NCM, males show a significant increase in FA during the photosensitive period at SD8 compared to photorefractory phases PR (left HVC $t(5, 54.7) = -4.09$, $p = 0.0019$, right HVC $t(5, 55.0) = -3.09$, $p = 0.0349$) and LD16 (left HVC $t(5, 54.7) = 3.66$, $p = 0.0071$, right HVC $t(5, 55.0) = 3.71$, $p = 0.0061$, left NCM $t(5, 54.3) = -3.71$, $p = 0.0061$). In contrast, in females FA slightly decreases over time during the photosensitive and photostimulated phase. The change in FA in the male HVC surroundings is also reflected in other diffusion parameters (MD, AD, RD), but not in FD (figure 8 - figure supplement 1). Interestingly, the surroundings of left HVC were larger in males compared to females, as indicated by log FC sex difference ($F(1, 23.2) = 13.9$, $p = 0.0011$). The sex difference in FA in left NCM ($F(1, 21.6) = 11.9$, $p = 0.0024$) can be attributed to a sex difference in RD, where RD is lower in males compared to females ($F(1, 23.2) = 109.2$, $p < 0.0001$), suggesting a difference in myelin content or axon diameter (see table 1).

In the rostro-lateral surroundings of left RA, males and females do not differ from each other except on SD4 where males have significantly higher FA compared to the later time points; in contrast, females do not present any change over time in this specific part of RA surroundings. Even though there is no sex difference in FA, males do have lower AD and RD compared to females in this region (respectively, $F(1, 22.6) = 39.6$, $p < 0.0001$ and $F(1, 23.4) = 60.3$, $p < 0.0001$).

Whereas male FA increases significantly in specific parts of the surroundings of main song control nuclei (HVC and left RA) and of the auditory system (left NCM), in females FA values remain constant or even decrease over time. Male FA increase during the photosensitive period in NCM and HVC surroundings is reflected in increased diffusion in all directions (MD, AD and RD all increase over time (figure 8 – figure supplement 1).

3.4.2 Similar seasonal neuroplasticity in both sexes mostly occurs in the photosensitive stage.

While some specific regions of the song control and auditory circuit present sexually dimorphic neuroplasticity, most neuroplasticity occurs in parallel in males as in females; not only in regions related to song control but also other sensory systems and the cerebellum (figure 7, 8). Within the song control system, sexually analogous neuroplasticity occurs in the surroundings of right rostral HVC, RA, Area X (rostral and caudal surroundings) and within left LMAN. More specifically the earliest change was seen at four weeks of short days, in the caudal and rostral regions surrounding Area X. FA values in Area X surroundings peaked around the end of the photosensitive period, sometimes showing a slight dip at LD4 but generally remaining high even at LD16. The caudal part of left RA surroundings presents increased FA values as early as 8 weeks of short days in both sexes, while in the rostral surroundings of right HVC males increase FA at SD8 and females increase FA at SD12. In contrast, the rostro-medial surroundings of right RA only shows a significant increase in FA at photostimulation (male PR vs LD4 $t(5, 55.7) = -3.05$, $p=0.0386$, females PR vs LD4 $t(5, 54.0) = -4.91$, $p=0.0001$). However, both in left and right RA surroundings, FA values remain elevated after sixteen weeks of long days. In the right HVC rostral surroundings, FA remains high in the male starlings, whereas females show a slight dip in FA during the photostimulated phase before increasing again during the photorefractory phase (female LD4 vs LD16 $t(5, 53.9) = -3.74$, $p=0.0057$). The LMAN is one of the last song control nuclei to change with increased FA levels at SD12 and LD4 relative to PR, following a similar pattern in males and females.

In several song control nuclei, FA starts to increase early in the photosensitive period in both sexes at a similar pace, and remains high during the photostimulated and photorefractory period. These FA changes can be mostly attributed to an increase in AD over time particularly in males (figure 8-figure supplement 2-3). Female starlings have a rapid decrease in RD during the photosensitive period in right HVC and Area X surroundings, whereas males present no, or limited, changes in RD (figure 8-figure supplement 2-3). This suggests that myelination plays a role in the neuroplasticity of females during the photosensitive period, whereas males display more changes in axon number or axon density.

In line with the voxel-based analysis for sex differences, only the left RA and Area X surroundings present FA values that are higher in males compared to females. However, both left and right Area X surroundings of males have lower diffusion in all directions (MD, AD and RD) compared to females, suggesting a higher tissue density or fiber organization in males (figure 8-figure supplement 2). Both left and right RA surroundings have lower RD in males compared to females, suggesting males have more myelin content or smaller axon diameter. However, in the right RA surroundings, AD is also significantly lower in males, suggesting a lower axon number or axon density. Interestingly, female starlings have larger MD, AD and log FC within LMAN compared to male starlings. This is in line with the fixel-based sex difference in log FC that encompasses parts of LFS in close proximity to LMAN.

Besides the song control system, many other regions including several white matter tracts and the cerebellum changed over time as well. Some of these structures are involved in song behavior such as the OM, LaM and CSt. OM conveys projections from the arcopallium, including connections from RA to downstream regions (Reiner et al., 2004), whereas CSt holds projections going to the auditory system (Vates et al., 1996). Females already show significant changes as early as four weeks of short days in OM and CSt, whereas males only have significant FA change at SD4 in the left OM and at SD12 in the right OM. In males, FA values of OM peak at SD12 and slightly decrease during photostimulation (SD12 vs LD4 left $t(5, 55.6) = 4.04$, $p = 0.0022$, right $t(5, 56.1) = 4.09$, $p = 0.0019$). In females the FA values in OM remain high during photostimulation and only gradually decrease in the subsequent photorefractory state in left OM (SD12 vs LD16 $t(5, 54.1) = 4.60$, $p = 0.0004$). These changes over time in FA are closely matched by the changes in FD. In line with the voxel-based sex differences, female starlings have higher FA and FD in OM and CSt compared to male starlings (figure 8, figure 8 - figure supplement 3). The CSt distinguishes itself from the other neuroplastic regions, as FA and FD keep on gradually increasing over time, even when switching from the photostimulated to the photorefractory phase.

LaM is the dividing lamina between the mesopallium and nidopallium and carries connections of HVC to Area X and of LMAN to RA (Nottebohm et al., 1982). In the left LaM, FA increased significantly after eight weeks of short days in both sexes (females PR vs SD8 ($t(5, 54.1) = -3.72$, $p = 0.0061$), males PR vs SD8 ($t(5, 55.5) = -2.99$, $p = 0.0454$)). FA levels remained elevated until SD12 in males and until LD4 in females, but did not reach baseline levels at LD16. FD in the LaM shows a similar increase during the photosensitive period, especially in males. Conversely, the increase in FA in females is mostly contributed to a decrease in RD during the photosensitive period.

In other white matter structures such as the TrO, TSM and tFA most neuroplasticity changes start during the photosensitive phase, but are only significant around SD12. FA values in these tracts often

remain high during the photostimulated phase, and afterwards return to photorefractory state levels. As the name indicates, the optic tract is involved in visual processing. TSM is also involved in the visual pathway as it conveys projections from the wulst to sites within the brain stem (Reiner et al., 2004). tFA holds projections of the trigeminal sensorimotor pathway between the basorostral nucleus (in the rostral pallium) and arcopallium (Reiner et al., 2004; Wild and Farabaugh, 1996). Only in the right tFA and TSM do FA levels remain high at LD16. The FD changes in TrO, TSM and left tFA closely match the changes over time in FA and confirm the sex difference detected in the fixel-based analysis (figure 8 - figure supplement 4).

Interestingly, cerebellar lobules VII-VIII, which are part of the oculomotor cerebellum, significantly increase in FA by the end of the photosensitive period, peak in FA during the photostimulated phase in both sexes (PR vs LD4 males $t(5, 54.4)=-6.32$, $p<0.0001$, PR vs LD4 females $t(5, 54.1)=-7.98$, $p<0.0001$) and return to baseline values at LD16. Similarly to the surroundings of Area X and right HVC, the FA changes over time in females coincide with a decrease in RD during the photosensitive period, but increased AD values in males (figure 8 - figure supplement 4). Overall, neuroplasticity in structures related to song and auditory processing starts early into the photosensitive period (SD4 and SD8). Structures related to other sensory systems, such as the TrO, TSM, tFA and cerebellum, only show significant increases by the end of the photosensitive phase (SD12) and often have decreased in FA again by the next photorefractory phase. Even though the majority of these neuroplasticity changes occur similarly in males and females, they sometimes differ in their approach. FA changes in males are mostly due to changes in AD; however, females show changes in RD over time.

Figure 7: Time effect and interaction between sex and time in fractional anisotropy (FA) The statistical maps were assessed at $p_{\text{uncorr}} < 0.001$ and $k_E \geq 10$ voxels with a small volume correction including regions of the white matter tracts, auditory and song control system. Main time effect of the cerebellum cluster was assessed without the small volume correction. Color scale represents significance by F-value. Abbreviations: C, caudal; RL, rostro-lateral; R, rostral; surr, surroundings.

Table 4: Clusters displaying a significant change over time and the interaction between sex and time in fractional anisotropy (FA)

Effect	Cluster	Hemisphere	Cluster		Peak	
			p_{FWE}	k_E	p_{FWE}	F
Interaction time * sex	HVC surr	Left	0.001	61	0.013	8.75
		Right	0.003	50	0.136	7.31
	RA rostro-lateral surr	Left	0.013	36	0.139	7.29
	NCM	Left	0.048	26	0.331	6.65
Main time effect	HVC rostral surr	Right	0.032	29	0.526	6.24
	RA caudal surr	Left	0.000	86	0.009	8.93
	RA rostro-medial surr	Right	0.006	42	0.017	8.57
	Area X caudal surr	Right	0.000	222	0.000	14.96
		Right	0.128	19	0.016	8.61
		Left	0.000	131	0.129	7.35
		Left	0.148	18	0.039	8.07
	Area X rostral surr	Left	0.000	76	0.007	9.09
		Left	0.037	28	0.144	7.27
		Right	0.000	176	0.000	11.05
	LMAN	Left	0.000	140	0.022	8.40
	OM	Left	0.000	513	0.000	16.44
		Right	0.000	158	0.002	9.75
	CSt	Left	0.000	82	0.003	9.53
		Right	0.001	61	0.048	7.95
		Right	0.025	31	0.242	6.89
	TSM	Right	0.009	39	0.009	8.97
		Left	0.001	56	0.074	7.70
	tFA	Left	0.000	91	0.000	12.73
		Right	0.015	35	0.042	8.03
	TrO	Left	0.000	302	0.000	14.18
		Right	0.001	58	0.264	6.83
	LaM	Left	0.037	28	0.120	7.40
	Cerebellum	Lobule VII	0.000	1779	0.000	15.26
		Lobule VIII			0.000	13.17

Table summarizing the significance at cluster level and peak level. Some clusters are large and cover multiple regions, indicated by the single statistic at cluster level. For each sub-region within this large cluster the peak significance (p_{FWE}) and T-value (T) are reported. P-values are FWE corrected. k_E indicates the number of voxels within a cluster. Grey values indicate clusters that are significant at cluster level but not at peak level. Abbreviations: surr, surroundings

Figure 8: Summary of the longitudinal changes over time of FA extracted from ROI-based clusters that showed a significant interaction (first row), a significant change over time in the surroundings of song control nuclei (row 2 and 3) or a significant change over time in the fiber tracts (row 4 - 6). The grey area indicates the entire photosensitive period of short days (8L:16D). Significant sex differences are reported with their p-value under the respective ROI-based cluster. Different letters denote significant differences by comparison with each other in post-hoc t-tests with $p < 0.05$ (Tukey's HSD correction for multiple comparisons) comparing the different time points to each other. If two time points share the same letter, the FA values are not significantly different from each other. Abbreviations: C, caudal; RL, rostro-lateral; R, rostral; surr, surroundings.

3.5 PHOTOPERIODIC VARIATION IN HORMONE LEVELS AND SONG BEHAVIOR

In order to determine how the changes in neuroplasticity are related to changes in hormones and singing behavior across the different photoperiods, we collected blood samples and recorded songs at each time point. A linear mixed model analysis was performed to determine differences between sex, time and identify their interaction with HSD multiple comparison correction. This way, we validated whether the artificial photoperiod schedule induced hormonal changes comparable to the natural seasonal gonadal cycle in starlings.

3.5.1 Hormone levels

Plasma testosterone levels change significantly over time in both male and female starlings ($F(5, 98.0)=11.7$, $p<0.0001$), but differed between sexes, as illustrated by a significant interaction ($F(5, 98.0)=4.82$, $p=0.0006$) (figure 9A, B). In males, testosterone levels peak after four weeks of photostimulation ($M= 1.67$, $SE=0.24$) compared to photorefractory state PR ($M= 0.03$, $SE= 0.26$, $t(5, 43.8)=-4.73$, $p=0.0003$). Two out of the thirteen male starlings already showed elevated plasma testosterone levels above 1ng/ml at the end of the photosensitive phase, but this increase was not significant at group level. In females, plasma testosterone levels are significantly elevated at the end of the photosensitive phase SD12 ($M= 0.43$, $SE= 0.11$) compared to LD16 ($M= 0.01$, $SE= 0.10$, $t(5, 54.2)=2.99$, $p=0.0461$). Even though both sexes increase their circulating testosterone levels, female starlings do not reach the same levels as male starlings do.

Since testosterone is not the main reproductive hormone in female starlings, we also looked at changes in circulating estrogen levels (figure 9C). Similar to testosterone in males, plasma estrogen levels change significantly over time ($F(5, 48.3)=19.5$, $p<0.0001$) with the highest concentrations measured after four weeks of photostimulation ($M= 67.8$, $SE=3.26$) compared to the photorefractory state PR ($M= 47.7$, $SE=3.12$, $t(48.3)=-4.88$, $p=0.0002$). These increased circulating hormones at photostimulation confirm the reproductive maturation associated with the breeding season in starlings. Furthermore, we show that circulating testosterone levels already gradually increase during the photosensitive period in both sexes.

Figure 9: Overview of the seasonal changes in plasma testosterone (A, B) and estrogen (C) concentrations in male and female starlings. Grey lines depict individual animals, while blue and red lines represent the group average with the standard errors of the mean error bars. The grey area indicates the photosensitive period of short days (8L:16D). Different letters denote significant differences by comparison with each other in post-hoc t-tests with $p < 0.05$ (Tukey's HSD correction for multiple comparisons) comparing the different time points to each other. If two time points share the same letter, the testosterone level or estrogen level are not significantly different from each other.

3.5.2 Seasonal variation in song behavior and the role of brain size

Under artificial conditions, male starlings sing with a consistent song rate during every photoperiod. In contrast, female starlings do show seasonal variation in song rate ($F(5, 42.1)=3.24, p=0.0145$). They sang most during the photosensitive phase (PR vs SD12 $t(5, 42.1)=-3.05, p=0.0428$). During the photostimulated phase, the song rate in females is drastically lower compared to males ($t(1, 14.0)=5.39, p=0.0358$) (figure 10A). In female starlings, the song bout length increased gradually over time from (22.7 ± 1.5) s ($M \pm SE$) at the photorefractory phase to (28.6 ± 0.8) s at the end of the photosensitive phase ($t(5, 336)=-2.97, p=0.0368$), when their song rate and testosterone levels were the highest (figure 10B). Male starlings presented a similar increase in song bout length over the sequential photoperiods from (24.3 ± 1.3) s at SD4 to (33.2 ± 1.1) s in the photostimulated phase ($t(5, 382.7)=-7.19, p<0.0001$). Interestingly, song bout length was positively correlated at the within-subject level to testosterone plasma levels in male starlings (rmcorr $r=0.632, p=0.037$). Such a repeated measures correlation was not present in female starlings; neither did the song rate correlate with testosterone levels in either sex.

We further investigated the effect of brain size on song behavior, using brain size (defined in result section 3.3) as a fixed factor in the analysis of the song behavior for each sex separately. Interestingly, the song rate of female starlings presented a significant effect of brain size and interaction between brain size and time (respectively ($F(1, 10.8) = 5.74, p=0.0358$; $F(5, 37.6)= 3.14, p=0.0181$). *Post hoc* analysis quickly revealed that female starlings with a small brain rarely sing during any of the time points (figure 10C). Large brain females were the most responsible in determining the seasonal change in song rate and song bout length. Further analysis revealed that the brain size of the female birds is proportional to their body weight, evident by positive Pearson's correlation between the averaged brain size and averaged body weight ($r=0.908, p<0.0001$), which is not the case in the male starlings (figure 10D). Female starlings with a large brain have a higher body weight and are capable of singing with a high song rate during the photosensitive period unlike their smaller female counterparts. Importantly, the interpretation of the sexual dimorphisms attributed to brain size is further

complicated, since the difference between large and small brain female starlings is confounded by their difference in singing behavior.

Figure 10: Overview of the seasonal changes in song rate (A) and song bout length (B) in male and female starlings. Dividing the starlings based on brain size showed that only large brain female starlings sing (C). Females present a significant correlation between the average brain size and body weight (D). (A-C) Blue and red lines represent the group average with the standard errors of the mean error bars. The grey area indicates the photosensitive period of short days (8L:16D). Different letters denote significant differences by comparison with each other in post-hoc t-tests with $p < 0.05$ (Tukey's HSD correction for multiple comparisons) comparing the different time points to each other. If two time points share the same letter, the song rate or song bout length are not significantly different from each other. (D) The Pearson's correlation and its significance are reported for males and females.

3.6 NEURAL CORRELATES OF SONG BEHAVIOR

Having established the seasonal changes in neuroplasticity and song behavior, we performed a voxel-based multiple regression analysis to determine the neural correlate of song behavior. Neither song bout length nor the male song rate presented a significant voxel-based correlation. However, female song rate correlated positively to FA in the left hippocampus and the surrounding fibers of RA and Area X (figure 11, Table 5). This correlation was not only linked to the overall difference between singing and non-singing females, but also to the increase of song rate within a subject as evidenced by the significant repeated measures correlation with FA of the caudal surroundings of Area X and left

hippocampus. Thus, we conclude that not only are large female starlings capable of producing more songs, this increased song rate is also correlated to improved microstructural organization of some major song control nuclei and the hippocampus.

Figure 11: Overview of structural neural correlates of song rate to FA in female starlings. The statistical maps were assessed at $p_{\text{uncorr}} < 0.001$ and $k_E \geq 10$ voxels with a small volume correction including regions of the song control system and other white matter structures. Below each statistical parametric map, the identified correlations were further explored with repeated measures correlation. Solid colored lines show the best linear fit for the within-subject correlation using parallel regression lines for individual animals. The dashed line represents the linear fit of the overall Pearson correlation representing the between-subject correlations. Abbreviations: C, caudal; RM, rostro-medial; R, rostral; surr, surroundings.

Table 5: Summary of the voxel-based correlation analysis of song rate vs FA in female starlings.

Cluster	Hemisphere	Cluster		Peak		Pearson's correlation	Rmcorr	
		p_{FWE}	k_E	p_{FWE}	T		r	P
Area X caudal surr	Left	0.000	168	0.002	6.05	0.5396	0.444	0.0023
	Right	0.001	116	0.015	5.45	0.5045	0.399	0.0066
Area X rostral surr	Left	0.041	45	0.427	4.24	0.4587	0.291	0.0525
RA rostro-medial surr	Right	0.015	59	0.039	5.13	0.3969	0.263	0.0806
Hippocampus	Left	0.132	30	0.029	5.22	0.3883	0.343	0.0211

This table summarizes for each significant cluster-based ROI the voxel-based FWE corrected p-value at cluster and peak level, next to the overall Pearson's correlation, the repeated measures correlation (rmcorr) and its significance of the averaged FA values versus the respective song characteristic. K_E indicates the cluster size. T reflects the t-value of the maximal peak located within the cluster-based ROI. Voxel-based correlations in grey are significant at the cluster level, but not at peak level. Significant repeated measures correlations are indicated in bold. Abbreviations: surr, surroundings.

4 DISCUSSION

The present data-driven longitudinal study confirms the well-established sexual dimorphism of the songbird brain, showing that male starlings generally have a more structured song control system compared to female starlings- as is the case in zebra finches where only the male sings (Hamaide et al., 2017). We also demonstrated new ones, including the finding that females have stronger tracts interconnecting hemispheres and nuclei, while males have a microstructural asymmetry towards the left hemisphere. Interestingly, brain size does not completely explain the sexually dimorphic hemispheric asymmetry of males or increased interconnectivity of female brains; only some of the sex differences in certain song control nuclei could be attributed to brain size. However, the difference in brain size is confounded by the difference in singing behavior of female starlings. Female starlings with small brains rarely sing during any of the time points. In contrast, females with large brains are capable of singing at a high song rate and strengthen their song control network during the photosensitive period. In general, both male and female starlings display similar structural neuroplasticity not only in the surrounding fibers of the song control system, but also in tracts related to other sensory systems such as the auditory and visual network. Even specific parts of the cerebellum, related to the processing of sensory information, display seasonal neuroplasticity. However, males and females show slight differences in the way neuroplasticity is displayed. Male starlings, for example, demonstrated significant neuroplasticity in HVC surroundings and left NCM, which was absent in female starlings. Importantly, most of these neuroplasticity changes already start during the photosensitive period, when circulating testosterone levels are low but gradually begin to increase in both sexes.

4.1 GENERAL SEXUAL DIMORPHISM IN THE STARLING BRAIN'S TRACTOGRAPHY

In vivo DTI and fixel-based analysis, although focusing on fibers surrounding and interconnecting the nuclear brain regions in the songbird brain, could capture the well-established structural sexual dimorphism at the level of the song control system in songbirds (Bernard et al., 1993; Hamaide et al., 2017; Nottebohm and Arnold, 1976). Even though the sexual dimorphism in the song control system

is bilateral, we demonstrated a strong asymmetry towards the left hemisphere in male starlings covering the left nidopallium and mesopallium. Furthermore, analysis of seasonal neuroplasticity confirmed this lateralization, as many clusters found in the left hemisphere (i.e. surroundings of left HVC, RA, rostral and caudal surroundings of Area X) had higher FA values in males compared to females. However, in many of the surroundings of song control nuclei, males had lower RD compared to females, which implies that males have smaller axonal diameters or more myelinated fibers in these structures (Song et al., 2005). In the rostral and caudal surroundings of Area X, diffusion in all directions (MD, AD and RD) is larger in females compared to males, which suggests that females have a less structured fiber organization or tissue density in this region (Beaulieu, 2002; Zatorre et al., 2012). These observations might guide histological studies to shed more light on the underlying structural differences causing these differences in FA. Left-right differences are not always examined in histological studies. Most of the studies concerning hemispheric asymmetry or lateralization focus on the lateralization of functions involved in the processing of auditory, visual and olfactory information (Rogers, 2012). Often these studies use electrophysiology or functional MRI (fMRI) to quantify the neural response to certain stimuli (De Groof et al., 2017; Yang and Vicario, 2015). Following the Hebbian hypothesis (neurons that fire together, wire together), a lateralized firing pattern suggests lateralized structural microstructure. Human imaging studies have combined fMRI and tract-tracings, and found that white matter asymmetries may be one of the factors underlying functional hemispheric asymmetries (Ocklenburg et al., 2013). Our study contributes to this hypothesis, showing a large microstructural hemispheric asymmetry of the male starling brain, which might be underlying some of the functional hemispheric asymmetries in starlings. Furthermore, this lateralization might be species-specific; zebra finches by contrast exhibit no left-right difference in the volume of song control nuclei (Nixdorf-Bergweiler, 1996). Prior DTI studies in zebra finches also did not find a sex difference in cerebral asymmetry of song control nuclei (Hamaide et al., 2017).

Unexpectedly, females have higher FA in several tracts interconnecting nuclei of the visual and song control system and in interhemispheric connections such as the commissure anterior and posterior. This is in contrast to zebra finches where males have a larger white matter volume compartment compared to females (Hamaide et al., 2018a). However, stronger interhemispheric connectivity is present in female mammals, though this has been attributed to brain size rather than a real effect of sex (Hanggi et al., 2014). Even though male starlings had a slightly -but not significantly- larger brain compared to females, matching our findings in zebra finches (Hamaide et al., 2018a), our study could not confirm the hypothesis of neuronal interconnectivity (Hanggi et al., 2014). The voxel-based analysis for brain size could not explain the hemispheric asymmetry in males or the increased interhemispheric connections in females. Brain size could only explain some sexual dimorphisms in

specific parts of the song control system, such as the caudal surroundings of the right Area X, surroundings of RA and HVC. However, in females the difference in brain size is confounded by the difference in song behavior. Our data reveal that female starlings with a large brain sing more, especially during the photosensitive phase, compared to female starlings with a relatively smaller brain. Furthermore, female song rate is correlated to FA in the same song control regions, e.g. the surroundings of Area X and RA. Brain size is affected by multiple factors including age, circulating hormone levels, environment and experience (Healy and Rowe, 2007). Of note, because we used wild caught starlings we cannot control for their exact age or developmental history. The larger females could be older compared to the smaller female starlings. In summation, these confounding factors make it unlikely that the brain size difference is the sole explanation for the sexual dimorphisms.

Sexual structural dimorphisms of the brain might also be explained by sexual differences in behavior. Take, for example, the sexual dimorphism of the TrO and TSM that carries projections to the wulst (Reiner et al., 2004). Enhancement of the visual system in females could be related to the visual cues like ultraviolet plumage colors female starlings take into consideration during mate selection (Bennett et al., 1997). Moreover, this role of the visual system during the breeding season is further corroborated by the neuroplasticity changes in the visual system, which occur at the end of the photosensitive period, right before the onset of the photostimulated period and return to baseline levels after sixteen weeks of long days. This is consistent with a prior DTI study, where the optic chiasm of male starlings had lower FA values during the photorefractory state compared to the breeding season (De Groof et al., 2008).

Interestingly, because of the limited research in the cerebellum of songbirds, we are the first to report clear sexual dimorphisms in the fiber characteristics of specific lobules of the cerebellum. Females had higher FA values in a large part of the cerebellum covering both anterior and posterior lobules IV-VIII, and higher FD in anterior lobules II-IV. In males, mostly the posterior lobule VII had higher FA, whilst lobule VIII had higher FD and log FC. The avian cerebellum has a strict topographic organization where each cerebellar lobule has a distinct input and output (Arends and Zeigler, 1991). Anterior lobules (I-V) contain the somatotopic representation of the tail, leg and wings (Whitlock, 1952) and are more pronounced in avian species with strong hind limbs (Iwaniuk et al., 2007). Cerebellar lobules VII and VIII are part of the oculomotor cerebellum (VI-VIII) and receive visual input (Clarke, 1977), input from the pretecto-ponto-cerebellar system involved in flight (Wylie et al., 2018) and have a somatotopic representation of the cochlea and retina (Whitlock, 1952). The pronouncement of the oculomotor cerebellum in males versus females could reflect the integration of visual and auditory sensory input, including the more developed song behavior of males. Language tasks in humans mostly involve the lateral lobules VI and crus I/II, which are implicated in initiating the motor sequence of phonological

content, whereas VIIb/VIII support the phonological storage during the maintenance of verbal information (Marien et al., 2014). However, there is only limited information on the specific cerebellar lobules implicated in song production of songbirds. Recently a cerebello-thalamic-basal ganglia pathway was identified that influences song processing, where lateral cerebellar nucleus (CbL) projects to dorsal thalamic regions and coincides with retrograde label of Area X and medial striatum (Person et al., 2008; Pidoux et al., 2018). Furthermore, lesioning Area X caused structural remodelling within the cerebello-thalamo-basal ganglia pathway including the dorsal thalamic zone, CbL and extended further in cerebellar lobules VIb and VII (Hamaide et al., 2018b). These lesions independently affected the duration of syllables (Hamaide et al., 2018b; Pidoux et al., 2018), which indicates the importance of the cerebello-thalamo-basal ganglia pathway in the timing of birdsong (Konopka and Roberts, 2016). The cerebellum generally works as a forward controller, where it compares intention with execution and uses the error information to correct future events (D'Angelo, 2018; Strick et al., 2009). This conceptualization of the cerebellum as a learning machine was developed as a model for motor control (Miall and Reckess, 2002), but can also be applied in other learning behaviors, such as vocal learning in birds, since the forward control process of the cerebellum is especially implicated in the timing of combined stimuli. Moreover, the oculomotor cerebellum exhibits neuroplasticity with a similar temporal profile to the multisensory neuroplasticity observed in the cerebrum. The molecular layer of the cerebellum contains the parallel fibers emitted by the granule cells and is known to maintain a high level of neuroplasticity throughout life (Ito, 2006; Medina et al., 2002). Furthermore, steroids can affect cerebellar Purkinje firing related to locomotor activity (Dean and McCarthy, 2008; Smith, 1995). Together, these findings imply that the plasticity in song control system is not limited to the known song circuitry but also involves the cerebellum.

The general sexual dimorphism in the starling brain extends well beyond the song control system, and whilst for mammals this was suggested to be related to the difference in brain size, we do not find evidence for that in avian species. The few sex differences in the song control system that could be attributed to brain size are confounded by song behavior. Furthermore, many of the other sexual dimorphisms involving the visual system and cerebellum could be explained by behavioral differences between sexes. It can thus be reasonably assumed that the general structural sexual dimorphisms in starlings are related to differences in behavior rather than brain size.

4.2 SEASONAL VARIATION IN GONADAL HORMONES, SONG BEHAVIOR AND ITS NEURAL CORRELATES

The seasonal variation in hormone levels (Dawson, 1983; Riters et al., 2002; Williams et al., 2004) and song behavior during artificial photoperiod switching is in good agreement with the seasonal changes

observed in natural conditions. In male starlings on a natural cycle, hypothalamic GnRH expression is already high in February under short days, which coincides with the start of the testes growth and results in a gradual increase of circulating testosterone (Bentley et al., 2013; Dawson, 1983). Female birds also produce testosterone, mainly in the ovaries and the adrenal gland (Tanabe et al., 1979). Circulating testosterone levels are high in female starlings in the pre-laying phase and decrease after ovulation, whereas estrogens are high during the breeding season in March and peak around ovulation (Dawson, 1983; Williams et al., 2004). It is well known that estrogens are fundamental in the regulation of female reproduction, whereas the role of testosterone in female reproduction is less studied. However, female starlings are the most aggressive during the pre-laying period, when their testosterone levels are high (Williams et al., 2004). Our study confirms that both male and females present a gradual increase in circulating testosterone during the photosensitive phase. In males, this was shown to coincide with an increase in hypothalamic GnRH expression (Bentley et al., 2013). Furthermore, the current study showed that in male's testosterone plasma levels were positively correlated to the song bout length on a within-subject level. We suggest that beside the effect of testosterone on reproductive functions, it could also be important in the regulation of neuroplasticity observed during the photosensitive phase.

4.3 MALE AND FEMALE STARLINGS BOTH EXPERIENCE NEUROPLASTICITY IN THEIR SONG CONTROL SYSTEM, BUT DIFFER IN THEIR SONG BEHAVIOR AND THE UNDERLYING MECHANISMS

Although structural neuroplasticity of the song control system is broadly similar in males and females, there are some notable differences. We revealed that seasonal neuroplasticity differs between males and females in specific parts of the HVC surroundings and left NCM, where only males experienced seasonal neuroplasticity. Females by contrast only showed an FA increase in the more rostral part of the right HVC surroundings after twelve weeks of short days. During photostimulation, when the female singing behavior is lower, FA values in the right HVC surroundings dropped slightly.

After four weeks of short days, our earliest time point of the photosensitive period, females presented a significant FA change in the OM, rostral and caudal Area X surroundings and the CSt, whereas males only had significant FA changes in the left OM, left RA surroundings (interaction cluster) and caudal surroundings of right Area X at that time point. Most of the neuroplasticity of song control nuclei like HVC and Area X surroundings in males occurred only after eight weeks of short days. It seems as if the song-related seasonal neuroplasticity occurs slightly sooner in females compared to males, especially in the rostral surroundings of Area X, which is a part of the anterior forebrain pathway, and the CSt, which is an auditory region. Both structures could play a role in the auditory feedback loop that is associated with vocal learning (Bosíková et al., 2010). The female singing behavior was most

prominent during the photosensitive period in our experiment, which is in good agreement with Pavlova et al. (2007). In contrast, males do not change their song rate over different photoperiods, in agreement with Adret-Hausberger (1994), but do show seasonal variation in their song bout length, in line with Van Hout et al. (2009). Sexual dimorphism at the level of the song control system suggests that male circuitry already has a stronger microstructural organization compared to the females. Male neuroplasticity changes in FA are mostly attributed to changes in AD, sensitive to changes in axon number or axon density. Female starlings conversely demonstrated decreased RD during the photosensitive period, indicating changes in myelin content or axon diameter, which was minimal in males.

Furthermore, FA values in the Area X surroundings and hippocampus were positively correlated to the female song rate on a within-subject level. The role of hippocampus in song behavior is not widely understood. Previous work has mainly focused on the role of the hippocampus in spatial memory performance (Macphail, 2002). Prior studies demonstrated that males generally outperform females in spatial memory performance tasks, but rely on different strategies in learning such cognitive tasks in both mammals and songbirds (Kosarussavadi et al., 2017; Rensel et al., 2015; Tropp and Markus, 2001). Sexually different learning mechanisms have also been suggested in vocal learning in songbirds (Riebel, 2016). Few studies have looked at the role of the hippocampus in song behavior in zebra finches. Hippocampal lesions during song learning and adulthood in male zebra finches did not affect song learning, singing or song structure. However, in females, hippocampal lesions affected preference for their father's song (Bailey et al., 2009). Interestingly, the hippocampus has a high aromatase expression and is a site of immediate early gene expression in response to conspecific zebra finch song, predominantly in female zebra finches (Bailey and Saldanha, 2015; Bailey and Wade, 2003; Gobes et al., 2009). These findings suggest that the hippocampus might play a role in specific characteristics of song perception, like identity and temporal context. To our knowledge, we are the first to report on a potential role of the hippocampus in the song behavior of female starlings.

Interestingly, many of the seasonal changes in brain structures related to the singing behavior do not return to baseline levels after sixteen weeks of long days. In males, the surroundings of HVC, left RA and left NCM were not different between the two sample points in the photorefractory phases in our study. Some structures such as the left OM show a gradual decrease in FA values at LD16 compared to their peak at SD12. A potential explanation is that these singing related structures require more time to return to baseline levels. Alternatively, certain regions related to song behavior may never fully regress to baseline photorefractory levels, but instead increase year after year. Bernard et al. (1996) demonstrated that in yearling starlings, HVC and RA volume was lower compared to older adult starlings, which was linked to the increasing song repertoire of starlings over the years. Furthermore,

CSt related to the auditory system distinguishes itself from other structures as it keeps on increasing in FA over subsequent photoperiods. This could represent a maturation upon aging, independent of the seasonal changes.

Together these findings suggest that males and females, just like in spatial memory, use different learning strategies in vocal learning as suggested by Riebel (2016). They rely on different parts of the song control system and use different mechanisms to manifest neuroplasticity in their respective song control system. Female starlings relied particularly on the auditory feedback loop and show involvement of the hippocampus. Males, on the other hand, display greater utilization of the NCM, left RA and HVC surroundings in their song control related neuroplasticity.

4.4 THE PHOTSENSITIVE PHASE IS A SENSITIVE WINDOW WHERE BOTH SEXES EXPERIENCE MULTISENSORY NEUROPLASTICITY

One of the most important findings of this study is that most of the neuroplasticity changes start during the photosensitive phase. Despite widespread structural sexual dimorphism, the majority of the structural seasonal neuroplasticity occurs similarly in males and females. A second major observation is that the SCS, various sensory systems and the cerebellum display neuroplasticity in both sexes. This finding fits with a prior histological study that reported volume changes in song control nuclei during the photosensitive phase of male starlings (Riters et al., 2002). During the photosensitive period, aromatase expression is increased and 5 β -reductase (which inactivates testosterone) is decreased in the telencephalon (Riters et al., 2001). This indicates that local changes in testosterone metabolism in the brain, could contribute to singing behavior and neuroplasticity changes during the photosensitive period, even though circulating testosterone levels are still low. In line with male starlings, male photosensitive canaries' HVC volume was not significantly different from photostimulated males, which the authors related to a gradual increase in GnRH expression in the POA during the photosensitive period (Hurley et al., 2008).

Just like starlings, canaries are open-ended learners. However, canaries modify their song between breeding seasons (when their testosterone levels are low); once they reach a stable song it remains stable for the rest of the breeding season (Nottebohm et al., 1986). Recently, it was shown that this song stabilization is associated with an increase in the perineuronal nets (PNN) expression in RA and Area X of adult male canaries prior to the breeding season (Cornez et al., 2020). They hypothesized that the changes in PNN expression were associated with the closing and re-opening of a sensitive period for vocal learning. We hypothesize that a similar pattern of closing and re-opening of sensitive windows exists in starlings. Among open-ended learners, starlings are more plastic compared to canaries, that is, they can learn new song elements throughout the year and can incorporate them in

their song up to eighteen months after first hearing them (Böhner et al., 1990; Marler et al., 1994). However, these experiences only result in long-lasting structural changes in the brain during the photosensitive period, which persist during the photostimulated phase and dissipate again in the photorefractory state. As in canaries, male starlings have higher percentage of PNN surrounding PV in the HVC during the photosensitive and photostimulated phase compared to the photorefractory state (Cornez et al., 2017). This indicates that PNNs could play a role in ending a sensitive period. However, compared to other open-ended learners (canaries) and closed-ended learners (zebra finches), starlings exhibit dramatically less PNN expression in their song control system (Cornez et al., 2017). This could be linked to the starling's ability to change its repertoire throughout the year. The relatively low PNN density in the starling brain does not necessarily mean that they cannot use other mechanisms to end their sensitive period. Myelination is another relevant mechanism by which the connectivity of a circuit is consolidated when a sensitive window closes (Knudsen, 2004; Xin and Chan, 2020). Using in vivo DTI, which is sensitive to changes in myelination, we demonstrated that both sexes experienced differential and parallel-heightened multisensory neuroplasticity in auditory and visual system and cerebellum mainly during the photosensitive period. This often-neglected period in which gonadal hormones are still very low might represent a 'sensitive window', during which different sensory and motor systems in both telencephalon and cerebellum are seasonally re-shaped. These findings open new vistas for investigating reopening of sensitive or critical periods of neuroplasticity.

5 MATERIAL & METHODS

5.1 SUBJECTS & EXPERIMENTAL SETUP

All experiments were performed on thirteen male and twelve female starlings (*Sturnus vulgaris*) that were wild caught as adults in Normandy (France) and Nicosia (Cyprus) in November 2014 and March 2017, respectively. All subjects were housed in large indoor aviaries (L x W x H: 2.2 x 1.4 x 2.1 m) at the University of Antwerp with food and water available *ad libitum* and an artificial light-dark cycle mimicking the natural photoperiod. All birds were kept on a long day photoperiod (16h light: 8h dark, 16L:8D) in order to induce photorefractory state at the start of experiment. When all birds molted and were fully photorefractory, a first baseline measurement was made. This was in May 2015 for the male starlings, and December 2017 for the female starlings. Both males and females followed the same photoperiodic regime: after the photorefractory baseline scan, they were switched from long to short (8L:16D) days to induce photosensitivity. After twelve weeks of short days, the photosensitive starlings were switched back to long days (16L:8D) to induce photostimulation. They remained on long days, so that they became photorefractory once more. We monitored neuroplasticity repeatedly at six

different time points. The first time point was at the end of the photorefractory state (PR). After switching to short days, we measured every four weeks to follow up the SEASONAL plasticity during the photosensitive period (SD4, SD8, SD12). Additionally, we measured after four weeks on long days (LD4), when starlings were fully photostimulated. Finally, we measured after sixteen weeks on long days (LD16), when starlings are photorefractory again. At each time point, birds were weighed, songs were recorded, blood samples were taken, and MRI (DTI and 3D scans) was acquired. The housing and experimental procedures were performed in agreement with the European directives, Belgian and Flemish laws and were approved by the Committee on Animal Care and Use of the University of Antwerp, Belgium (2014-52).

5.2 TESTOSTERONE AND ESTROGEN ASSAY

Blood samples were collected to quantify plasma testosterone concentration in all birds. In female starlings, we also assayed estradiol in addition to testosterone, since this is one of the primary female sex hormones. At each time point, blood collection occurred within a 2-hour window at noon between 11:30h and 13:00h in order to limit variation of plasma levels due to a circadian rhythm. Starlings were housed individually for song recording, which allowed fast capture and blood sampling (within 5 min. of capture) to minimize the increase in testosterone as a result of acute stress (Van Hout et al., 2010). The alar wing vein was punctured with a 25-gauge needle to collect 300–500 μ l of blood into heparin-coated capillary tubes. The blood was transferred to tubes and centrifuged for 10 min at 2060g while cooling to 4°C. Plasma was collected and stored at –20 °C prior to assay. Estradiol and testosterone concentrations were quantified by Radioimmunoassay (RIA) using a commercial double antibody system (MP Biomedicals, Solon, Ohio). The estradiol RIA does not cross-react significantly with other estrogens beside estradiol-17 β (Estrone: 20% and Estriol 1.51%; all other steroids: <1%) and the intra-assay coefficient of variation is 4.7–10.6%. The testosterone RIA does not significantly cross-react with other androgens beside testosterone (5 α -dihydrotestosterone: 3.4%; 5 α -androstane-3 β , 17 β -diol: 2.2%; 11-oxo-testosterone: 2%; all other steroids: <1%) and the intra-assay coefficient of variation is 4.6–9.1%.

5.3 SONG RECORDING AND ANALYSIS

At each MRI imaging time point, songs of all birds were recorded for 2 days, so that we could accurately monitor the evolution of the song behavior in relation to the hormone and neuroplasticity changes. Since isolated starlings do not sing as they would do in groups, we temporarily housed the starlings individually in cages that were placed next to the aviary. Each cage was equipped with a Behringer C-2 condensator microphone placed on top, which was connected to a PR8E multi-channel

preamplifier (SM Pro audio), ‘Delta Series’ Breakout box (M-audio), and a personal computer with Sound Analysis Pro (SAP) software version 2011.104 (<http://soundanalysispro.com/>) for automated song detection. Due to technical issues, songs of PR and SD12 of male birds were lost.

For each bird at each time point, song recordings of four consecutive hours in the morning were analyzed using Raven Pro 1.5 (Cornell Lab of Ornithology, Ithaca, NY). We created sonograms that were acoustically and visually inspected to identify individual songs (N=881) from the background noise of the aviary based on their intensity and the spectral definition. In line with prior song processing studies (Eens, 1997; Van Hout et al., 2012), we defined a song bout as a period of at least 5 s of song with pauses no longer than 1.5 s. Starling song consists of a series of distinct song elements or phrases, which can be allocated to 4 categories of phrase types that are mostly performed in a fixed order. The song often starts with one or several whistles (whistle phrase type), followed by a section of complex phrases (variable phrase type) and rapid series of clicks or rattles (rattle phrase type), before ending the song with high frequency song elements (high frequency phrases type). Song bouts containing at least three different phrase types were labeled as ‘complete song bouts’. The number of complete song bouts sung per hour within the 4-hour period was taken as a measure of the song rate. We used the first 20 complete song bouts sung within this 4-hour period to calculate the song bout length. This way, we avoided the overrepresentation of songs during time points where the song rate was high.

5.4 MRI DATA ACQUISITION

The birds were initially anesthetized using 2% Isoflurane (Isoflo[®], Abbot Laboratories Ltd.) with a mixture of oxygen and nitrogen at flow rates of 100 and 200 cm³/min, respectively. Throughout the imaging procedure, the anesthesia was gradually reduced to 1% Isoflurane in response to their respiration rate, which was monitored with a small pneumatic sensor (SA Instruments, NY, USA) positioned under the bird. Body temperature was monitored with a cloacal temperature probe and kept within a narrow physiological range (41.0 ± 0.2 °C) using a warm air system with a feedback unit (SA Instruments, NY, USA).

All MRI measurements were performed on a 7T horizontal MR system (Pharmascan 70/16 US, Bruker Biospin, Germany). In each imaging session, we acquired a T2-weighted 3D anatomical Rapid acquisition with relaxation enhancement (RARE) scan (3D; TR: 2000 ms; TE_{eff}: 44 ms; RARE factor: 8; acquisition matrix of (256x92x64) giving a voxel resolution of (0.089 x 0.250 x 0.250) mm³). Subsequently, a 4 shot spin-echo echo-planar imaging (SE-EPI) diffusion-weighted (DW) scan (TR: 7000 ms; TE: 23 ms; δ 4ms, Δ 12ms; b-value 670 s/mm²; 60 diffusion gradient directions; spatial resolution: (0.179 x 0.179 x 0.350) mm³; 28 coronal slices) was acquired. The 60 diffusion gradient directions were

divided over 3 scans, each starting with the acquisition of 7 b0 images. Total acquisition time of all MRI scans amounted to 50 min. After the scanning session, birds were left to recover in a warmed recovery box before returning to the song recording cages next to the aviary.

5.5 MRI DATA PROCESSING

Whole-brain volume was manually delineated on the T2-weighted 3D anatomical RARE scan, which covered the entire brain, including telencephalon, diencephalon, mesencephalon and metencephalon. These volumes were used as a measure of brain size, used in further statistical analysis.

DW-images were prepared for voxel-based analysis using MRtrix3 version 3.0 (Tournier et al., 2012). We used an in-house algorithm to convert the Bruker 2dseq files to nifti files, which are compatible with other software programs such as SPM and MRtrix3. This step includes a signal scale correction, since the DWI data is acquired in three separate sequential scans. Furthermore, all DW-images were scaled by factor a 10 in 3 dimensions, to enable proper processing of small brains in software programs designed to process human data like SPM and MRtrix3. Voxel-based analysis requires that all images are spatially normalized to the same template, to enable voxel-wise comparisons. A simplified overview of the different MRI data processing steps is given in figure 1B-D.

5.5.1 DTI processing

Preprocessing the diffusion data for voxel-based analysis was performed using MRtrix3 (Tournier et al., 2012). First, diffusion gradient orientations were checked and automatically corrected to match the coordinate frame of MRtrix3 and ensure the best global 'connectivity' (Jeurissen et al., 2014a). Since the DTI data is acquired in three separate sequential scans, we further corrected for intensity differences between scans by rescaling the diffusion scans based on their b0 images. Preprocessing of the individual DW-images included the following steps: denoising (Veraart et al., 2016), correction for Gibbs ringing (Kellner et al., 2016), motion and distortion correction using FSL (Andersson and Sotiropoulos, 2016; Jenkinson et al., 2012), bias field correction using ANTS (Advanced Normalization Tool; (Avants et al., 2010)), creating an automated whole-brain mask for whole-brain extraction that were manually checked, upsampling to isotropic voxels of 1.75 mm. These preprocessed diffusion-weighted images were used to calculate individual diffusion maps (FA, MD, AD, and RD). The transformation parameters derived from building the FOD template (see 5.5.2) were applied to the diffusion maps to warp them into the template space to perform voxel-based analysis. Next, these images were smoothed to double voxel size (3.5 x 3.5 x 3.5 mm³). Finally, all normalized diffusion maps were averaged to create an FA template that is used as a background to display the statistical results.

5.5.2 Fixel based analysis

For calculation of the fiber-based metrics, we followed the preprocessing steps as defined in (Raffelt et al., 2017). Fixel based analysis follows the same preprocessing steps as DTI processing up until the bias field correction. Apparent fiber density analysis differs from DTI analysis by the fact that it is related to the diffusion-weighted signal intensity within a given voxel. Therefore, global intensity normalization is performed to ensure robust voxel-wise comparison across subjects. Within this step, we used the default FA threshold of 0.4 to create an approximate WM mask, which is used to normalize the median white matter $b=0$ intensity across all subjects (Raffelt et al., 2012). For each image, a white matter response function was estimated for spherical deconvolution using the unsupervised Dhollander algorithm (Dhollander et al., 2019). Next, the average of all individual response functions was calculated and used for constrained spherical deconvolution to estimate FOD images (Jeurissen et al., 2014b). These FOD images were normalized to create an unbiased study-based FOD template, which involves linear and non-linear registration (Raffelt et al., 2011).

Next, the fixels in the FOD template are thresholded at 0.15, identifying the template white matter fixels to be included in further analysis. This threshold is lower than the default threshold of 0.25 for the human brain, as this threshold is too high for the songbird brain and excludes many of the genuine white matter fibers. We are aware that choosing a lower threshold comes with the risk of introducing noisy fixels, especially within grey matter, and consider this in the interpretation of the results.

The estimated transformation parameters or warps of each subject to the template were used to transform the individual FOD maps into template space without FOD reorientation, so that the apparent FD can be estimated prior to reorienting the fixels. In the next step, we compare the fixels within a spatially matching voxel of the individual subject and the template white matter fixels, to identify which fixels correspond to each other and subsequently assign the corresponding FD value (Raffelt et al., 2017).

Next to FD, we also computed a fixel-based metric related to the macroscopic morphology in fiber bundle cross-section (FC). Fiber bundle cross-section (FC) information relies solely on the transformation parameters or Jacobian determinants generated during the construction of the population template, similar to other morphometry analyses like voxel-based morphometry (Ashburner and Friston, 2000). Morphological differences in the plane perpendicular to the fixel orientation could reflect differences in the number of axons, myelination, or the spacing between axons. For group statistical analysis, the FC values were logarithmically transformed to log FC to ensure that the data are centered on zero and normally distributed. Positive values indicate then expansion,

whereas negative values reflect shrinkage of a fiber bundle relative to the template (Raffelt et al., 2017).

5.5.3 Tractography

A whole-brain tractogram was generated from the FOD template, which is required for statistical analysis using connectivity-based fixel enhancement (CFE) (Raffelt et al., 2015). We created 20 million tracks using dynamic seeding, and maximum angle of 45°. We used Spherical-deconvolution informed filtering (SIFT) to filter our tractogram to ensure that the streamline densities approximate the biological fiber densities estimated by spherical deconvolution. SIFT2 creates for each streamline a weighting coefficient based on the underlying FOD (Smith et al., 2015). Currently, track-weights created in SIFT2 cannot be implicated in statistical analysis. Therefore, in our statistical analyses we used SIFT, which selectively filters out streamlines based on the spherical deconvolution to reduce the tractogram to 2 million tracks (Smith et al., 2013).

5.6 STATISTICAL ANALYSIS

Voxel-based statistical analyses allow for unbiased determination of sexual dimorphisms, longitudinal changes over time and the neural correlates of song behavior (performed using SPM12). To examine group level repeated measurements in a voxel-based manner using the standard general linear model approach as implemented in SPM and FSL, we followed the recommendations of McFarquhar (2019). Effects of sex and brain size are assessed using full factorial design, whereas the effect of time and the interaction between sex and time are assessed using a flexible factorial design in SPM.

Voxel-based clusters are further explored by extracting the average values from these statistical ROI clusters and investigate either their change over time or their correlation to a certain song parameter. Traditional correlation techniques like Pearson correlation require independence of data. Since we have repeated measures, this assumption of independence is violated when performing a multiple regression on longitudinal data. Traditionally this is solved by using data from only one time point. However, a lot of information on temporal changes is lost this way. Therefore, we use repeated measures correlation instead of traditional correlation techniques (Bakdash and Marusich, 2017). This way, we can assess the within-subject correlation, which gives a meaningful interpretation of how testosterone, song and diffusion parameters correlate to each other on a within-subject level.

Hormone levels, brain size and song parameters were analyzed using linear mixed model in JMP Pro 13 (SAS Institute Inc., Cary, NC, 1989-2007) to assess the main effect of time, sex and their interaction.

5.6.1 Song analysis and plasma testosterone and estrogen concentrations

We investigated whether brain size, song behavior, testosterone and estrogen levels changed significantly over time using linear mixed model ('time point', 'sex' and the interaction 'sex * time point' as fixed factors and 'subject ID' as a random factor) via JMP Pro 13 (SAS Institute Inc., Cary, NC, 1989-2007). Tukey's Honest Significant Difference (HSD) was used for *post hoc* statistical testing to correct for multiple comparison.

Subsequently, within-subject correlations between testosterone and song parameters were assessed by using the repeated measures correlation package (rmcorr in R) (Bakdash and Marusich, 2017). Differences were considered significant for $p < 0.05$.

5.6.2 Statistical analysis brain size

To investigate the role of brain size on the general sexual dimorphism and on song behavior, we artificially divided the starlings by their brain size using a median split analogous to the analysis in Kurth et al. (2018). Birds with a brain size larger than the median brain size of 1805 mm³ were considered as large brain starlings. This way, we divided females in large and small brain groups (respectively, females: N=6 large and N=6 small, males N=6 large, N=7 small). In order to test whether song or testosterone change differently in large versus small brain birds, we performed a linear mixed model analysis for each sex separately with time, brain size (categorical factor of 2 levels) and their interaction as fixed factors and subject as a random factor. Post hoc testing in the linear mixed model analysis was corrected with HSD for multiple comparison using JMP Pro 13. Brain size (continuous variable), body weight and song parameters were averaged over time to perform a Pearson correlation between these factors for each sex separately in JMP Pro 13.

5.6.3 Voxel-based analyses

To identify the general sexual dimorphisms between male and female starlings, we first performed a full factorial voxel-based analysis within the entire brain using SPM12. The same full factorial design was used in the voxel-based analysis for categorical brain size difference with the caveat that brain size was substituted as a factor instead of sex. In a separate analysis, we determined the longitudinal neuroplasticity changes over time within the song control system, auditory system and their interconnecting tracts using a flexible factorial analysis. To investigate where within the starling brain microstructural properties relate to song behavior parameters, we executed voxel-based multiple regression analyses for FA. To this end, the averaged song behavior parameters were used for this analysis. Clusters were only considered significant if their peak value reached the significance level of $p < 0.05$ corrected for the familywise error rate (FWE) and contained at least (k_E) 10 (contiguous) voxels. Flexible factorial and multiple regression analyses were performed with a small volume mask including

the song control nuclei, auditory system and their interconnecting tracts. This way, we ensured we did not miss significant changes in relevant regions that might be obscured by the stringent correction for multiple comparisons in whole-brain analysis. Results shown are from the voxel-based analysis with the latter small volume correction. Results were displayed with a significance threshold of $p < 0.001$ (uncorrected for multiple comparison) and overlaid to a group-average FA map (average of all spatially normalized FA maps, including all subjects and time points).

Significant clusters were converted to ROIs (cluster-based ROIs) of which the average diffusion parameters (FA, MD, AD and RD) and fixel based parameters (FD and log FC) were extracted. In order to determine the extent of the overall association between diffusion parameters and song rate, Pearson's correlation was calculated using JMP Pro 13 (SAS Institute Inc., Cary, NC, 1989-2007). Further in depth analysis was performed to determine if this overall association was also caused by a within-subject correlation by using rmcrr in R (Bakdash and Marusich, 2017).

5.6.4 Fixel-based analysis

Fixel-based statistical analysis was performed using the connectivity-based fixel enhancement (CFE) described by (Raffelt et al., 2015). A connectivity matrix was calculated using the fixel template and SIFT filtered tractogram. This connectivity matrix was used to perform connectivity-based smoothing where only fixels belonging to the same fiber tract are smoothed, instead of the neighboring voxels as in a voxel-based analysis. A fixel-based analysis was performed to assess the difference between sexes and brain size, both examined in separate tests. In the CFE statistical test, we used whole block exchangeability to look for changes between subjects and included the time point information in a variance group. Similarly to the voxel-based DTI analysis, statistical results were visualized with a significance threshold of $p < 0.001$ (uncorrected for multiple comparison) and color-coded according to their significance T-value or the physical direction of the tract.

6 ACKNOWLEDGMENTS

The authors wish to thank Ben Jeurissen (Vision Lab of the University of Antwerp) and Robert Smith (The Florey Institute of Neuroscience and Mental Health) for their help with MRtrix3. This research was funded by a grant from the Research Foundation – Flanders (FWO, project Nr G030213N)) and Interuniversity Attraction Poles (IAP) ("PLASTOSCINE": P7/17) to Annemie Van der Linden. Jasmien Orije is a PhD fellow of the FWO (Nr 1115217N) and Elisabeth Jonckers received Postdoctoral fellowships of the FWO (Nr 12R1917N).

7 REFERENCES:

- Adret-Hausberger, M., 1994. Seasonal variation in the whistles of the starling *sturnus vulgaris*. *Ibis* 126, 372-378.
- Alward, B.A., Balthazart, J., Ball, G.F., 2013. Differential effects of global versus local testosterone on singing behavior and its underlying neural substrate. *Proc Natl Acad Sci U S A* 110, 19573-19578.
- Andersson, J.L.R., Sotiropoulos, S.N., 2016. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 125, 1063-1078.
- Arends, J.J.A., Zeigler, H.P., 1991. Organization of the cerebellum in the pigeon (*Columba livia*): I. Corticonuclear and corticovestibular connections. *Journal of Comparative Neurology* 306, 221-244.
- Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry--the methods. *Neuroimage* 11, 805-821.
- Avants, B.B., Yushkevich, P., Pluta, J., Minkoff, D., Korczykowski, M., Detre, J., Gee, J.C., 2010. The optimal template effect in hippocampus studies of diseased populations. *Neuroimage* 49, 2457-2466.
- Bailey, D.J., Saldanha, C.J., 2015. The importance of neural aromatization in the acquisition, recall, and integration of song and spatial memories in passerines. *Hormones and behavior* 74, 116-124.
- Bailey, D.J., Wade, J., 2003. Differential expression of the immediate early genes FOS and ZENK following auditory stimulation in the juvenile male and female zebra finch. *Molecular Brain Research* 116, 147-154.
- Bailey, D.J., Wade, J., Saldanha, C.J., 2009. Hippocampal lesions impair spatial memory performance, but not song--a developmental study of independent memory systems in the zebra finch. *Dev Neurobiol* 69, 491-504.
- Bakdash, J.Z., Marusich, L.R., 2017. Repeated Measures Correlation. *Front Psychol* 8, 456.
- Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR Biomed* 15, 435-455.
- Beaulieu, C., 2014. The Biological Basis of Diffusion Anisotropy. 155-183.
- Bennett, A.T., Cuthill, I.C., Partridge, J.C., Lunau, K., 1997. Ultraviolet plumage colors predict mate preferences in starlings. *Proc Natl Acad Sci U S A* 94, 8618-8621.
- Bentley, G.E., Tucker, S., Chou, H., Hau, M., Perfito, N., 2013. Testicular growth and regression are not correlated with Dio2 expression in a wild male songbird, *sturnus vulgaris*, exposed to natural changes in photoperiod. *Endocrinology* 154, 1813-1819.
- Bernard, D.J., Ball, G.F., 1997. Photoperiodic condition modulates the effects of testosterone on SCS nuclei volumes in male starlings. *General and comparative endocrinology* 105, 276-283.
- Bernard, D.J., Casto, J.M., Ball, G.F., 1993. Sexual dimorphism in the volume of song control nuclei in European starlings: Assessment by a Nissl stain and autoradiography for muscarinic cholinergic receptors. *The Journal of comparative neurology* 334, 559-570.
- Bernard, D.J., Eens, M., Ball, G.F., 1996. Age- and behavior-related variation in volumes of song control nuclei in male European starlings. *J Neurobiol* 30, 329-339.
- Böhner, J., Chaiken, M.L., Ball, G.F., Marler, P., 1990. Song acquisition in photosensitive and photorefractory male European starlings. *Hormones and behavior* 24, 582-594.
- Bosíková, E., Kostal, L., Kubikova, L., 2010. Birdsong: From behaviour to brain. *Biologia* 65, 379-387.
- Clarke, P.G.H., 1977. Some visual and other connections to the cerebellum of the pigeon. *Journal of Comparative Neurology* 174, 535-552.
- Cornez, G., Collignon, C., Müller, W., Ball, G.F., Cornil, C.A., Balthazart, J., 2020. Seasonal changes of perineuronal nets and song learning in adult canaries (*Serinus canaria*). *Behavioural Brain Research* 380, 112437.
- Cornez, G., Madison, F.N., Van der Linden, A., Cornil, C., Yoder, K.M., Ball, G.F., Balthazart, J., 2017. Perineuronal nets and vocal plasticity in songbirds: A proposed mechanism to explain the difference between closed-ended and open-ended learning. *Dev Neurobiol* 77, 975-994.
- D'Angelo, E., 2018. Physiology of the cerebellum. *Handb Clin Neurol* 154, 85-108.

- Dawson, A., 1983. Plasma gonadal steroid levels in wild starlings (*Sturnus vulgaris*) during the annual cycle and in relation to the stages of breeding. *General and comparative endocrinology* 49, 286-294.
- De Groof, G., Balthazart, J., Cornil, C.A., Van der Linden, A., 2017. Topography and Lateralized Effect of Acute Aromatase Inhibition on Auditory Processing in a Seasonal Songbird. *J Neurosci* 37, 4243-4254.
- De Groof, G., Gwinner, H., Steiger, S., Kempenaers, B., Van der Linden, A., 2010. Neural correlates of behavioural olfactory sensitivity changes seasonally in European starlings. *PLoS One* 5, e14337.
- De Groof, G., Poirier, C., George, I., Hausberger, M., Van der Linden, A., 2013. Functional changes between seasons in the male songbird auditory forebrain. *Front Behav Neurosci* 7, 196.
- De Groof, G., Verhoye, M., Van Meir, V., Balthazart, J., Van der Linden, A., 2008. Seasonal rewiring of the songbird brain: an in vivo MRI study. *Eur J Neurosci* 28, 2475-2485; discussion 2474.
- De Groof, G., Verhoye, M., Van Meir, V., Tindemans, I., Leemans, A., Van der Linden, A., 2006. In vivo diffusion tensor imaging (DTI) of brain subdivisions and vocal pathways in songbirds. *Neuroimage* 29, 754-763.
- Dean, S.L., McCarthy, M.M., 2008. Steroids, sex and the cerebellar cortex: implications for human disease. *Cerebellum* (London, England) 7, 38-47.
- Dhollander, T., Mito, R., Raffelt, D., Connelly, A., 2019. Improved white matter response function estimation for 3-tissue constrained spherical deconvolution. *Proc Intl Soc Mag Reson Med* 27.
- Eens, M., 1997. Understanding the Complex Song of the European Starling: An Integrated Ethological Approach. 26, 355-434.
- Genc, S., Tax, C.M.W., Raven, E.P., Chamberland, M., Parker, G.D., Jones, D.K., 2020. Impact of b-value on estimates of apparent fibre density. *Hum Brain Mapp* 41, 2583-2595.
- Gobes, S.M.H., ter Haar, S.M., Vignal, C., Vergne, A.L., Mathevon, N., Bolhuis, J.J., 2009. Differential responsiveness in brain and behavior to sexually dimorphic long calls in male and female zebra finches. *Journal of Comparative Neurology* 516, 312-320.
- Hall, Z.J., Macdougall-Shackleton, S.A., 2012. Influence of testosterone metabolites on song-control system neuroplasticity during photostimulation in adult European starlings (*Sturnus vulgaris*). *PLoS One* 7, e40060.
- Hamaide, J., De Groof, G., Van Ruijssevelt, L., Lukacova, K., Van Audekerke, J., Verhoye, M., Van der Linden, A., 2018a. Volumetric development of the zebra finch brain throughout the first 200 days of post-hatch life traced by in vivo MRI. *Neuroimage* 183, 227-238.
- Hamaide, J., De Groof, G., Van Steenkiste, G., Jeurissen, B., Van Audekerke, J., Naeyaert, M., Van Ruijssevelt, L., Cornil, C., Sijbers, J., Verhoye, M., Van der Linden, A., 2017. Exploring sex differences in the adult zebra finch brain: In vivo diffusion tensor imaging and ex vivo super-resolution track density imaging. *Neuroimage* 146, 789-803.
- Hamaide, J., Lukacova, K., Orije, J., Keliris, G.A., Verhoye, M., Van der Linden, A., 2020. In vivo assessment of the neural substrate linked with vocal imitation accuracy. *Elife* 9.
- Hamaide, J., Lukacova, K., Van Audekerke, J., Verhoye, M., Kubikova, L., Van der Linden, A., 2018b. Neuroplasticity in the cerebello-thalamo-basal ganglia pathway: A longitudinal in vivo MRI study in male songbirds. *Neuroimage* 181, 190-202.
- Hanggi, J., Fovenyi, L., Liem, F., Meyer, M., Jancke, L., 2014. The hypothesis of neuronal interconnectivity as a function of brain size-a general organization principle of the human connectome. *Front Hum Neurosci* 8, 915.
- Healy, S.D., Rowe, C., 2007. A critique of comparative studies of brain size. *Proc Biol Sci* 274, 453-464.
- Hurley, L.L., Wallace, A.M., Sartor, J.J., Ball, G.F., 2008. Photoperiodic induced changes in reproductive state of border canaries (*Serinus canaria*) are associated with marked variation in hypothalamic gonadotropin-releasing hormone immunoreactivity and the volume of song control regions. *General and comparative endocrinology* 158, 10-19.
- Ito, M., 2006. Cerebellar circuitry as a neuronal machine. *Progress in Neurobiology* 78, 272-303.

1129 Iwaniuk, A.N., Hurd, P.L., Wylie, D.R., 2007. Comparative morphology of the avian cerebellum: II. Size
1130 of folia. *Brain Behav Evol* 69, 196-219.

1131 Jarvis, E.D., Gunturkun, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W.J., Medina, L., Paxinos, G.,
1132 Perkel, D.J., Shimizu, T., Striedter, G., Wild, J.M., Ball, G.F., Dugas-Ford, J., Durand, S.E., Hough, G.E.,
1133 Husband, S., Kubikova, L., Lee, D.W., Mello, C.V., Powers, A., Siang, C., Smulders, T., Wada, K., White,
1134 S.A., Yamamoto, K., Yu, J., Reiner, A., Butler, A.B., 2005. Avian brains and a new understanding of
1135 vertebrate brain evolution. *Nat Rev Neurosci* 6, 151-159.

1136 Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M., 2012. *Fsl. Neuroimage*
1137 62, 782-790.

1138 Jeurissen, B., Leemans, A., Sijbers, J., 2014a. Automated correction of improperly rotated diffusion
1139 gradient orientations in diffusion weighted MRI. *Med Image Anal* 18, 953-962.

1140 Jeurissen, B., Tournier, J.D., Dhollander, T., Connelly, A., Sijbers, J., 2014b. Multi-tissue constrained
1141 spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *Neuroimage* 103,
1142 411-426.

1143 Karten, H.J., Brzozowska-Precht, A., Lovell, P.V., Tang, D.D., Mello, C.V., Wang, H., Mitra, P.P., 2013.
1144 Digital atlas of the zebra finch (*Taeniopygia guttata*) brain: a high-resolution photo atlas. *The Journal*
1145 *of comparative neurology* 521, 3702-3715.

1146 Kellner, E., Dhital, B., Kiselev, V.G., Reiser, M., 2016. Gibbs-ringing artifact removal based on local
1147 subvoxel-shifts. *Magn Reson Med* 76, 1574-1581.

1148 Knudsen, E.I., 2004. Sensitive Periods in the Development of the Brain and Behavior. *Journal of*
1149 *Cognitive Neuroscience* 16, 1412-1425.

1150 Konopka, G., Roberts, T.F., 2016. Insights into the Neural and Genetic Basis of Vocal Communication.
1151 *Cell* 164, 1269-1276.

1152 Kosarussavadi, S., Pennington, Z.T., Covell, J., Blaisdell, A.P., Schlinger, B.A., 2017. Across sex and
1153 age: Learning and memory and patterns of avian hippocampal gene expression. *Behavioral*
1154 *neuroscience* 131, 483-491.

1155 Kurth, F., Thompson, P.M., Luders, E., 2018. Investigating the differential contributions of sex and
1156 brain size to gray matter asymmetry. *Cortex* 99, 235-242.

1157 Larson, T.A., Thatra, N.M., Hou, D., Hu, R.A., Brenowitz, E.A., 2019. Seasonal changes in neuronal
1158 turnover in a forebrain nucleus in adult songbirds. *The Journal of comparative neurology* 527, 767-
1159 779.

1160 Macphail, E., 2002. The role of the avian hippocampus in spatial memory. *Psicológica: Revista de*
1161 *metodología y psicología experimental*, ISSN 0211-2159, Vol. 23, Nº 1, 2002 (Ejemplar dedicado a:
1162 Spatial Learning and Cognition), pages. 93-108 23.

1163 Marien, P., Ackermann, H., Adamaszek, M., Barwood, C.H., Beaton, A., Desmond, J., De Witte, E.,
1164 Fawcett, A.J., Hertrich, I., Kuper, M., Leggio, M., Marvel, C., Molinari, M., Murdoch, B.E., Nicolson,
1165 R.I., Schmammann, J.D., Stoodley, C.J., Thurling, M., Timmann, D., Wouters, E., Ziegler, W., 2014.
1166 Consensus paper: Language and the cerebellum: an ongoing enigma. *Cerebellum* 13, 386-410.

1167 Marler, P., BÖHNER, J., Chaiken, M., 1994. Repertoire Turnover and the Timing of Song Acquisition in
1168 European Starlings. *Behaviour* 128, 25-39.

1169 McFarquhar, M., 2019. Modeling Group-Level Repeated Measurements of Neuroimaging Data Using
1170 the Univariate General Linear Model. *Frontiers in neuroscience* 13, 352.

1171 Medina, J.F., Repa, J.C., Mauk, M.D., LeDoux, J.E., 2002. Parallels between cerebellum- and
1172 amygdala-dependent conditioning. *Nat Rev Neurosci* 3, 122-131.

1173 Miall, R.C., Reckess, G.Z., 2002. The Cerebellum and the Timing of Coordinated Eye and Hand
1174 Tracking. *Brain and Cognition* 48, 212-226.

1175 Nixdorf-Bergweiler, B.E., 1996. Divergent and parallel development in volume sizes of telencephalic
1176 song nuclei in male and female zebra finches. *The Journal of comparative neurology* 375, 445-456.

1177 Nottebohm, F., Arnold, A.P., 1976. Sexual dimorphism in vocal control areas of the songbird brain.
1178 *Science* 194, 211-213.

1179 Nottebohm, F., Kelley, D.B., Paton, J.A., 1982. Connections of vocal control nuclei in the canary
1180 telencephalon. *The Journal of comparative neurology* 207, 344-357.

1181 Nottebohm, F., Nottebohm, M.E., Crane, L., 1986. Developmental and seasonal changes in canary
1182 song and their relation to changes in the anatomy of song-control nuclei. *Behavioral and Neural*
1183 *Biology* 46, 445-471.

1184 Ocklenburg, S., Strockens, F., Gunturkun, O., 2013. Lateralisation of conspecific vocalisation in non-
1185 human vertebrates. *Laterality* 18, 1-31.

1186 Odom, K.J., Hall, M.L., Riebel, K., Omland, K.E., Langmore, N.E., 2014. Female song is widespread and
1187 ancestral in songbirds. *Nat Commun* 5, 3379.

1188 Orije, J., Cardon, E., de Groof, G., Hamaide, J., Jonckers, E., Van Massenhoven, S., Darras, V.,
1189 Balthazart, J., Verhoye, M., Van der Linden, A., 2020. In vivo online monitoring of testosterone-
1190 induced neuroplasticity in a female songbird. *Hormones and behavior* 118.

1191 Pavlova, D., Pinxten, R., Eens, M., 2007. Seasonal singing patterns and individual consistency in song
1192 activity in female European starlings (*Sturnus vulgaris*). *Behaviour* 144, 663-680.

1193 Person, A.L., Gale, S.D., Farries, M.A., Perkel, D.J., 2008. Organization of the songbird basal ganglia,
1194 including area X. *The Journal of comparative neurology* 508, 840-866.

1195 Pidoux, L., Le Blanc, P., Levenes, C., Leblois, A., 2018. A subcortical circuit linking the cerebellum to
1196 the basal ganglia engaged in vocal learning. *Elife* 7.

1197 Raffelt, D., Tournier, J.D., Fripp, J., Crozier, S., Connelly, A., Salvado, O., 2011. Symmetric
1198 diffeomorphic registration of fibre orientation distributions. *Neuroimage* 56, 1171-1180.

1199 Raffelt, D., Tournier, J.D., Rose, S., Ridgway, G.R., Henderson, R., Crozier, S., Salvado, O., Connelly, A.,
1200 2012. Apparent Fibre Density: a novel measure for the analysis of diffusion-weighted magnetic
1201 resonance images. *Neuroimage* 59, 3976-3994.

1202 Raffelt, D.A., Smith, R.E., Ridgway, G.R., Tournier, J.D., Vaughan, D.N., Rose, S., Henderson, R.,
1203 Connelly, A., 2015. Connectivity-based fixel enhancement: Whole-brain statistical analysis of
1204 diffusion MRI measures in the presence of crossing fibres. *Neuroimage* 117, 40-55.

1205 Raffelt, D.A., Tournier, J.D., Smith, R.E., Vaughan, D.N., Jackson, G., Ridgway, G.R., Connelly, A., 2017.
1206 Investigating white matter fibre density and morphology using fixel-based analysis. *Neuroimage* 144,
1207 58-73.

1208 Reiner, A., Perkel, D.J., Bruce, L.L., Butler, A.B., Csillag, A., Kuenzel, W., Medina, L., Paxinos, G.,
1209 Shimizu, T., Striedter, G., Wild, M., Ball, G.F., Durand, S., Gunturkun, O., Lee, D.W., Mello, C.V.,
1210 Powers, A., White, S.A., Hough, G., Kubikova, L., Smulders, T.V., Wada, K., Dugas-Ford, J., Husband,
1211 S., Yamamoto, K., Yu, J., Siang, C., Jarvis, E.D., *Avian Brain Nomenclature*, F., 2004. Revised
1212 nomenclature for avian telencephalon and some related brainstem nuclei. *The Journal of*
1213 *comparative neurology* 473, 377-414.

1214 Rensel, M.A., Ellis, J.M.S., Harvey, B., Schlinger, B.A., 2015. Sex, estradiol, and spatial memory in a
1215 food-caching corvid. *Hormones and behavior* 75, 45-54.

1216 Riebel, K., 2016. Understanding Sex Differences in Form and Function of Bird Song: The Importance
1217 of Studying Song Learning Processes. *Frontiers in Ecology and Evolution* 4.

1218 Ritters, L.V., Baillien, M., Eens, M., Pinxten, R., Foidart, A., Ball, G.F., Balthazart, J., 2001. Seasonal
1219 Variation in Androgen-Metabolizing Enzymes in the Diencephalon and Telencephalon of the Male
1220 European Starling (*Sturnus vulgaris*). *Journal of neuroendocrinology* 13, 985-997.

1221 Ritters, L.V., Eens, M., Pinxten, R., Ball, G.F., 2002. Seasonal changes in the densities of alpha(2)
1222 noradrenergic receptors are inversely related to changes in testosterone and the volumes of song
1223 control nuclei in male European starlings. *The Journal of comparative neurology* 444, 63-74.

1224 Rogers, L.J., 2012. The two hemispheres of the avian brain: their differing roles in perceptual
1225 processing and the expression of behavior. *Journal of Ornithology* 153, 61-74.

1226 Rouse, M.L., Jr., Stevenson, T.J., Fortune, E.S., Ball, G.F., 2015. Reproductive state modulates
1227 testosterone-induced singing in adult female European starlings (*Sturnus vulgaris*). *Hormones and*
1228 *behavior* 72, 78-87.

1229 Sagi, Y., Tavor, I., Hofstetter, S., Tzur-Moryosef, S., Blumenfeld-Katzir, T., Assaf, Y., 2012. Learning in
1230 the fast lane: new insights into neuroplasticity. *Neuron* 73, 1195-1203.

1231 Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2013. SIFT: Spherical-deconvolution informed
1232 filtering of tractograms. *Neuroimage* 67, 298-312.

1233 Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2015. SIFT2: Enabling dense quantitative
1234 assessment of brain white matter connectivity using streamlines tractography. *Neuroimage* 119,
1235 338-351.

1236 Smith, S.S., 1995. Sensorimotor-correlated discharge recorded from ensembles of cerebellar Purkinje
1237 cells varies across the estrous cycle of the rat. *Journal of Neurophysiology* 74, 1095-1108.

1238 Song, S.-K., Sun, S.-W., Ramsbottom, M.J., Chang, C., Russell, J., Cross, A.H., 2002. Dysmyelination
1239 Revealed through MRI as Increased Radial (but Unchanged Axial) Diffusion of Water. *NeuroImage* 17,
1240 1429-1436.

1241 Song, S.K., Yoshino, J., Le, T.Q., Lin, S.J., Sun, S.W., Cross, A.H., Armstrong, R.C., 2005. Demyelination
1242 increases radial diffusivity in corpus callosum of mouse brain. *Neuroimage* 26, 132-140.

1243 Stevenson, T.J., Ball, G.F., 2010. Photoperiodic differences in a forebrain nucleus involved in vocal
1244 plasticity: enkephalin immunoreactivity reveals volumetric variation in song nucleus IMAN but not
1245 Nlf in male European starlings (*Sturnus vulgaris*). *Dev Neurobiol* 70, 751-763.

1246 Stevenson, T.J., Hahn, T.P., Ball, G.F., 2012. Variation in gonadotrophin-releasing hormone-1 gene
1247 expression in the preoptic area predicts transitions in seasonal reproductive state. *Journal of*
1248 *neuroendocrinology* 24, 267-274.

1249 Strick, P.L., Dum, R.P., Fiez, J.A., 2009. Cerebellum and nonmotor function. *Annu Rev Neurosci* 32,
1250 413-434.

1251 Suarez, R., Paolino, A., Fenlon, L.R., Morcom, L.R., Kozulin, P., Kurniawan, N.D., Richards, L.J., 2018. A
1252 pan-mammalian map of interhemispheric brain connections predates the evolution of the corpus
1253 callosum. *PNAS* 115, 9622-9627.

1254 Tanabe, Y., Nakamura, T., Fujioka, K., Doi, O., 1979. Production and secretion of sex steroid
1255 hormones by the testes, the ovary, and the adrenal glands of embryonic and young chickens (*Gallus*
1256 *domesticus*). *General and comparative endocrinology* 39, 26-33.

1257 Tournier, J.D., Calamante, F., Connelly, A., 2012. MRtrix: Diffusion tractography in crossing fiber
1258 regions. *International Journal of Imaging Systems and Technology* 22, 53-66.

1259 Tropp, J., Markus, E.J., 2001. Sex differences in the dynamics of cue utilization and exploratory
1260 behavior. *Behavioural Brain Research* 119, 143-154.

1261 Van Hout, A.J., Eens, M., Balthazart, J., Pinxten, R., 2009. Complex modulation of singing behavior by
1262 testosterone in an open-ended learner, the European Starling. *Hormones and behavior* 56, 564-573.

1263 Van Hout, A.J., Eens, M., Darras, V.M., Pinxten, R., 2010. Acute stress induces a rapid increase of
1264 testosterone in a songbird: implications for plasma testosterone sampling. *General and comparative*
1265 *endocrinology* 168, 505-510.

1266 Van Hout, A.J., Pinxten, R., Darras, V.M., Eens, M., 2012. Testosterone increases repertoire size in an
1267 open-ended learner: an experimental study using adult male European starlings (*Sturnus vulgaris*).
1268 *Hormones and behavior* 62, 563-568.

1269 Vates, G.E., Broome, B.M., Mello, C.V., Nottebohm, F., 1996. Auditory pathways of caudal
1270 telencephalon and their relation to the song system of adult male zebra finches. *The Journal of*
1271 *comparative neurology* 366, 613-642.

1272 Veraart, J., Novikov, D.S., Christiaens, D., Ades-Aron, B., Sijbers, J., Fieremans, E., 2016. Denoising of
1273 diffusion MRI using random matrix theory. *Neuroimage* 142, 394-406.

1274 Whitlock, D.G., 1952. A neurohistological and neurophysiological study of afferent fiber tracts and
1275 receptive areas of the avian cerebellum. *Journal of Comparative Neurology* 97, 567-635.

1276 Wild, J.M., Farabaugh, S.M., 1996. Organization of afferent and efferent projections of the nucleus
1277 basalis prosencephali in a passerine, *Taeniopygia guttata*. *Journal of Comparative Neurology* 365,
1278 306-328.

- 1279 Williams, T.D., Kitaysky, A.S., Vezina, F., 2004. Individual variation in plasma estradiol-17beta and
1280 androgen levels during egg formation in the European starling *Sturnus vulgaris*: implications for
1281 regulation of yolk steroids. *General and comparative endocrinology* 136, 346-352.
- 1282 Wylie, D.R., Gutierrez-Ibanez, C., Gaede, A.H., Altshuler, D.L., Iwaniuk, A.N., 2018. Visual-Cerebellar
1283 Pathways and Their Roles in the Control of Avian Flight. *Frontiers in neuroscience* 12, 223.
- 1284 Xin, W., Chan, J.R., 2020. Myelin plasticity: sculpting circuits in learning and memory. *Nature Reviews*
1285 *Neuroscience* 21, 682-694.
- 1286 Yang, L.M., Vicario, D.S., 2015. Exposure to a novel stimulus environment alters patterns of
1287 lateralization in avian auditory cortex. *Neuroscience* 285, 107-118.
- 1288 Zatorre, R.J., Fields, R.D., Johansen-Berg, H., 2012. Plasticity in gray and white: neuroimaging
1289 changes in brain structure during learning. *Nat Neurosci* 15, 528-536.

1290

1291

1292 8 SUPPLEMENTARY FIGURES

Figure 8 – Figure supplement 1: Summary of the longitudinal changes over time in FA, MD, AD, RD, FD and log FC extracted from ROI-based clusters that had a significant interaction, including HVC surroundings (surr), left RA rostro-lateral surroundings (RLsurr) and NCM.

Figure 8 – Figure supplement 2: Summary of significant the longitudinal changes over time in FA, MD, AD, RD, FD and log FC extracted from ROI-based clusters surrounding several song control nuclei including right HVC rostral surroundings (Rsurr), left RA caudal surroundings (Csurr), Area X rostral surroundings and caudal surroundings.

Figure 8 – Figure supplement 3: Summary of the significant longitudinal changes over time in FA, MD, AD, RD, FD and log FC extracted from ROI-based clusters at level of LMAN and several tracts related to song and auditory processing including OM, CSt and LaM.

Figure 8 – Figure supplement 4: Summary of the significant longitudinal changes over time in FA, MD, AD, RD, FD and log FC extracted from ROI-based clusters at level of TrO, TSM, tFA and cerebellum.