Cerebral Cortex, January 2018;28: 387–394

doi: 10.1093/cercor/bhx299
Advance Access Publication Date: 9 November 2017
Original Article

ORIGINAL ARTICLE

Same Genes, Different Brains: Neuroanatomical Differences Between Monozygotic Twins Discordant for Musical Training

Örjan de Manzano and Fredrik Ullén

Department of Neuroscience, Retzius väg 8, Karolinska Institutet, 17177 Stockholm, Sweden

Address correspondence to Örjan de Manzano, Department of Neuroscience, Retzius väg 8, Karolinska Institutet, 17177 Stockholm, Sweden.
Email: orjan.demanzano@ki.se.

Abstract

Numerous cross-sectional and observational longitudinal studies show associations between expertise and regional brain anatomy. However, since these designs confound training with genetic predisposition, the causal role of training remains unclear. Here, we use a discordant monozygotic (identical) twin design to study expertise-dependent effects on neuroanatomy using musical training as model behavior, while essentially controlling for genetic factors and shared environment of upbringing. From a larger cohort of monozygotic twins, we were able to recruit 18 individuals (9 pairs) that were highly discordant for piano practice. We used structural and diffusion magnetic resonance imaging to analyze the auditory-motor network and within-pair differences in cortical thickness, cerebellar regional volumes and white-matter microstructure/fractional anisotropy. The analyses revealed that the musically active twins had greater cortical thickness in the auditory-motor network of the left hemisphere and more developed white matter microstructure in relevant tracts in both hemispheres and the corpus callosum. Furthermore, the volume of gray matter in the left cerebellar region of interest comprising lobules I–IV + V, was greater in the playing group. These findings provide the first clear support for that a significant portion of the differences in brain anatomy between experts and nonexperts depend on causal effects of training.

Key words: expertise, MRI, music, neuroanatomy, twins

Introduction

Many neuroimaging studies have documented that experts in various domains differ from nonexperts in regional brain anatomy (Ullén et al. 2016). Several of these studies have used musicians as a model group, finding larger volume and cortical thickness of auditory and motor regions, the cerebellum, as well as white-matter structural differences, for example, of the corticospinal tract which carries motor responses from the cortex to the spinal cord, and the corpus callosum which connects the cerebral hemispheres (Schlaug, Jäncke, Huang and Steinmetz 1995; Schlaug, Jäncke, Huang, Staiger, et al. 1995; Gaser and Schlaug 2003; Bengtsson et al. 2005; Bermudez et al. 2009). Animal studies suggest that these differences could arise from various expertise-dependent adaptations, including increased synapse numbers, modified synapse morphology, and increased axonal myelination in task-relevant regions/tracts (reviewed in Markham and Greenough 2004).

Thus, a common assumption is that similar outcomes reflect causal effects of training on brain plasticity. However, meta-analyses indicate that practice only explains between 21% and

36% of the variance in expert performance (Macnamara et al. 2014; Platz et al. 2014); different individuals may require different amounts of practice to reach the same level of skill (Macnamara et al. 2014); and correspondingly, there can still be significant differences in skill at the highest levels of performance despite similar amounts of practice (Macnamara et al. 2016). Furthermore, twin modeling studies demonstrate that both practice itself and its correlation with expertise-related outcomes can be heavily influenced by genetic factors (Mosing, Madison, et al. 2014) and that genetic factors predominantly determine individual differences in gray- and white-matter structure (Peper et al. 2007; Eyler et al. 2012). Hence, expertise likely depends on many variables (Ullén et al. 2016) and therefore, correlations between practice and brain measures does not imply causation. This has important implications for how we interpret previous neuroimaging studies, which have typically relied on cross-sectional and observational longitudinal designs (with self-selected samples) (Moreno and Bidelman 2014), that is, designs where genetic predispositions can influence group compositions and outcomes. Moreover, since no longitudinal study have come close to providing the training required to become an expert, it is unclear how such results correspond to research on professionals. Gray matter in the primary motor cortex, for example, does not expand linearly with practice of fine motor skills, but renormalizes after an initial growth period, which means that the observed volume increases in experts may not show a simple relation to practice per se (Wenger et al. 2017). A critical question is thus to what extent the neuroanatomical differences between experts and nonexperts depend on training or on genetic constitution.

We address this question in a novel way using a monozygotic cotwin control design. Monozygotic twin pairs who are discordant for exposure can be used to approximate a true experiment when experimental approaches are unfeasible (McGue et al. 2010). In the case of musical expertise, a randomized controlled trial would have to extend for many years or decades and involve several thousands of hours of practice. Here, since the twins have a common genotype and have shared a common early rearing environment, we can implicitly use the nonplaying twin to estimate what the playing twin would have looked like if he or she had not practiced. Hence, if musical practice has a long-term causal influence on brain structure, we expect that the playing twins will show higher rates of the outcomes (described below) than their nonplaying cotwins.

From a large cohort of twins (Mosing, Madison, et al. 2014), we identified 18 individuals (i.e., 9 pairs) that matched inclusion criteria: only one sibling still musically active, a within-pair practice difference of at least 1000 h and eligibility for magnetic resonance imaging (MRI). Our aim was to use structural and diffusion magnetic resonance imaging to analyze within-pair differences in cortical thickness, cerebellar regional volumes and white-matter microstructure (fractional anisotropy; FA) while effectively controlling for genes and shared environment of upbringing. We wanted to target the auditory-motor network involved in piano performance, in which neural adaptations are well-replicated. This included key regions of interest (ROIs) of auditory processing—the primary auditory cortex, posterior superior temporal gyrus including the planum temporale (STGp), the temporoparietal junction/Sylvian parietal temporal area (SPT), and inferior temporal gyrus (ITG) (Schlaug, Jäncke, Huang and Steinmetz 1995; Gaser and Schlaug 2003; Bermudez et al. 2009; Hyde et al. 2009; Berkowitz and Ansari 2010); motor processing—the primary motor, dorsal and ventral premotor cortices (M1, PMD, and PMV) (Gaser and Schlaug 2003; Bermudez et al. 2009; Hyde et al. 2009); and executive functions—the inferior frontal gyrus (IFG), dorsolateral prefrontal cortex/rostral middle frontal gyrus (rMFG), and rostral anterior cingulate cortex (rACC) (Sluming et al. 2002; Gaser and Schlaug 2003; Bermudez et al. 2009; Hyde et al. 2009). Two cerebellar ROIs (lobules I–IV + V and lobules VIIIa + VIIIb) were also included, which represent the areas involved in fine hand motor tasks (Stoodley et al. 2012). White-matter tracts of interest were the arcuate fasciculus, inferior longitudinal fasciculus and uncinate fasciculus, that is, tracts through which information from the temporal auditory regions can reach the inferior frontal cortex, directly or indirectly (Schlaug, Jäncke, Huang, Staiger et al., 1995; Bengtsson et al. 2005; Halwani et al. 2011; Dohn et al. 2015), as well as the corticospinal tract and corpus callosum.

Materials and Methods

Participants

The recruitment of participants is described in Eriksson et al. (2016). From a cohort of 10 539 twins (1211 monozygotic twins) who had participated in an extensive web survey (Mosing, Madison, et al. 2014), we identified an initial sample of 83 monozygotic (genetically identical) twin pairs across Sweden, where (1) only one sibling in each pair currently played a keyboard instrument (piano, organ, keyboard) and (2) the within-pair difference in total hours of music practice was at least 1000 h. Out of this sample, 10 twin pairs agreed to participate and come to Stockholm for interviews and testing (Eriksson et al. 2016). The present sample includes the 9 twin pairs (18 individuals; 10 females; age = 31–47 years, M = 37 ± 6 years), in which both twins fulfilled the general inclusion criteria for magnetic resonance imaging. However, due to a scanner malfunction, diffusion-weighted imaging data could not be obtained from one individual (excluding one pair from the FA analysis). We chose to focus on keyboard instrumentalists because keyboard is a common instrument type which has been used frequently in the past. We also wanted to avoid pooling expertise on different musical instruments, for which the behavior and neural implementation could differ substantially (Bangert and Schlaug 2006). The neuroanatomical images were examined by a neuroradiologist, without remark. The interviews explored nongenetic influences on musical engagement and were organized into 5 sections—(1) perceived reasons for the discordance; (2) childhood differences in specific music related variables; (3) strong memories of music; (4) the perceived meaning of music in life and for health; and (5) language interests—and analyzed using response categorization. Relevant results from these interviews are summarized in the Discussion; further details can be found in Eriksson et al. (2016). Travel expenses were covered and the participants were in addition given 2000 SEK in reimbursement. All participants gave informed consent and the study was approved by the Regional Ethical Review Board in Stockholm (Dnr 2011/570-31/5; 2012/788-31/2).

Data Acquisition

Music Practice

The original web-survey (Mosing, Madison, et al. 2014) provided information on music practice. This data consisted of self-reports on starting and (when applicable) ending year of practice, as well as the average number of hours spent on music practice per week in 4 age periods (ages 0–5, 6–11, 12–17, and 18–present). Based on these responses, we calculated hours of early practice (ages 0–11) and the total hours of practice.

Structural and Functional MRI

The MRI data were acquired using a 3-T scanner (Discovery MR750w 3.0 T, GE Healthcare) with a standard 8-channel head coil, at the MR Research Center of the Karolinska Hospital. During the scanning sessions, a separate experiment in which we collected functional MRI (fMRI) data was also conducted (to be presented elsewhere). In the present study, we made use of a subset of that data in order to get functional localizers of the cortical ROIs (see Cortical Thickness Analysis). All imaging parameters can be found in the Supplementary material.

Cortical Thickness Analysis

Data Preprocessing and Definition of ROIs

The cortical thickness analysis was performed with FreeSurfer using the T1w-images (http://surfer.nmr.mgh.harvard.edu/). A brief description of the FreeSurfer preprocessing pipeline can be found in the Supplementary material. The T2w-images were included during preprocessing to optimize refinement of pial surfaces. Parcellation of the cerebral cortex was based on the Desikan–Killiany–Tourville (DKT) atlas (Fischl et al. 2004; Klein and Tourville 2012). Cortical thickness is calculated as the closest distance from the gray/white boundary to the gray/CSF boundary at each vertex on the tessellated surface (Fischl and Dale 2000).

For this analysis, we selected the known key ROIs of the auditory-motor cortical network which are active during piano performance and have consistently been linked to musical expertise in previous studies:

- Heschl's gyrus (the transverse temporal gyrus), containing the primary auditory cortex (Schneider et al. 2002; Gaser and Schlaug 2003; Hyde et al. 2009; James et al. 2014).
- STGp, involved in the perception of complex sounds (Schlaug, Jäncke, Huang and Steinmetz 1995; Bermudez et al. 2009; Fauvel et al. 2014).
- SPT, involved in auditory-motor transformations (Hickok et al. 2003; Berkowitz and Ansari 2010).
- ITG, which plays a role in integrating visual and auditory sensory information (e.g., reading music), in rhythm perception, autobiographical salience and emotional processing (Gaser and Schlaug 2003; Bermudez et al. 2009).
- M1, here more specifically the "hand knob" of the primary motor cortex, involved in the execution of hand/finger movements (Amunts et al. 1997; Gaser and Schlaug 2003; Bermudez et al. 2009; Hyde et al. 2009).
- PMD, involved in planning and organization of movement sequences (Gaser and Schlaug 2003; Bailey et al. 2014).
- PMV, involved in auditory-motor transformations, for example, associating auditory stimuli with corresponding actions (Lahav et al. 2007; Bailey et al. 2014).
- IFG, which plays an essential role in processing musical syntax and semantics as well as more generally—implicit memory retrieval, audio-motor transformations and serial production rules (Sluming et al. 2002; Gaser and Schlaug 2003; Abdul-Kareem et al. 2011; James et al. 2014).
- rMFG, most notably involved in (auditory) working memory and cognitive control (Bermudez et al. 2009; Hyde et al. 2009).
- rACC, important for attention, action monitoring, and response inhibition (Bermudez et al. 2009; Han et al. 2009).

These regions align well with the dorsal and ventral pathways that enable auditory-motor integration in language (Saur et al. 2008). Overall, it appears that music and language do exist "side-by-side in the brain" (Brown et al. 2006), and that even for music, lateralization of function plays an important role in expertise (Schlaug, Jäncke, Huang and Steinmetz 1995). Consequently, we decided to examine the identified network of brain regions in each hemisphere separately.

Anatomical ROIs: From the DKT atlas, we extracted values for the following ROIs bilaterally: rACC, rMFG, ITG, IFG pars opercularis, IFG pars triangularis, IFG pars orbitalis, and Heschl's gyrus.

Functional ROIs: The additional ROIs (hand/finger region of M1, SPT, PMD, PMV, and STGp bilaterally) were constructed from activation data since they are not readily available in current atlases. In order to localize these ROIs, we made use of fMRI data collected in conjunction with the structural imaging data. The details of fMRI data preprocessing and analysis pipelines, which were implemented using the SPM12 software package (Wellcome Department of Imaging Neuroscience, London, UK) and MATLAB, as well as the details concerning the definition of the functional ROIs, can be found in the Supplementary material. In brief, during one condition of the fMRI experiment (out of several), the participants used their right hand to improvise simple melodies (free ordinal and temporal structure), using a 4-button response box and piano tones as auditory feedback (middle C, D, E, F). Contrasting free improvisation to implicit rest using a second-level one-sample *t*-test, provided a way of localizing relevant clusters of activity. We chose a liberal statistical threshold ($P < 0.001$, uncorrected for multiple comparisons), since the objective here was not to pinpoint peak activity, but rather to identify the broader limit of clusters representing the ROIs. Only the musician twins were used in this analysis since the main hypothesis concerned the structure of regions that musicians use during musical performance. The resulting labels were then transformed to the *fsaverage* brain, that is, the standard brain to which all the participant's brains were registered during preprocessing (Fig. 1A). Since the free improvisation was performed using the right hand only, the left hemisphere ROI labels were transformed to the right hemisphere. Lastly, using the inverse subject-to-fsaverage registration parameters, all labels could be transformed to native space and cortical thickness values for each label and participant could be calculated.

Data Analysis

The within-pair difference between playing and nonplaying twins in cortical thickness in the auditory-motor network was analyzed using a linear mixed model implemented in R (package lmertest), where cortical thickness was regressed on musicianship. This approach was adopted to test whether there was, in line with our alternative hypothesis, an overall effect of musical training on brain regions central to auditory and motor processing during musical performance. The mixed model allows for estimating a group effect of musicianship on cortical thickness across ROIs, based on within-pair comparisons. Thus, ROI was included as a fixed effect and pair membership was included as a random effect. As the auditory-motor network is lateralized to some degree, we produced one model for each hemisphere. The model was estimated using restricted maximum likelihood (REML) and the number of degrees of freedom was calculated using Satterthwaite approximations. The group difference was considered significant if $P < 0.05$, one-tailed (i.e., expecting a relatively thicker cortex in the playing twins). Moreover, we correlated the within-pair differences in

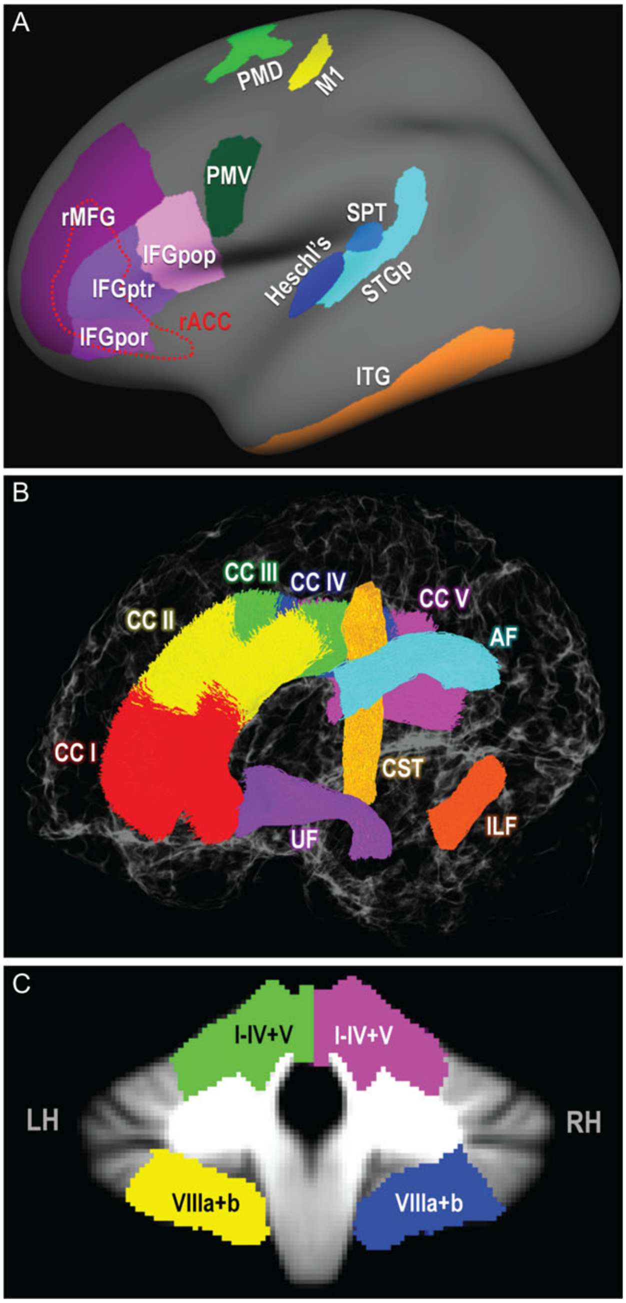

Figure 1. (A) ROIs used in the cortical thickness analysis, illustrated on the left hemispheric surface of the fsaverage inflated brain. Heschl's = Heschl's gyrus; IFGpop = inferior frontal gyrus pars opercularis; IFGpor; inferior frontal gyrus pars orbitalis; IFGptr = inferior frontal gyrus pars triangularis; ITG = inferior temporal gyrus; M1 = primary motor cortex, hand/finger region; PMD = dorsal premotor region; PMV = ventral premotor region; rACC = rostral anterior cingulate cortex (on the medial surface); rMFG = rostral middle frontal gyrus; SPT = Sylvian parietal temporal area; STGp = superior temporal gyrus including planum temporale. (B) The tracts of interest (shown in the left hemisphere, and corpus callosum), based on probabilistic fiber tracking and generation of streamlines in template space. AF = arcuate fasciculus; CC = corpus callosum; CST = corticospinal tract; ILF = inferior longitudinal fasciculus; UF = uncinate fasciculus. (C) Cerebellar ROIs, used in the VBM analysis. LH = left hemisphere; RH = right hemisphere.

mean cortical thickness (across the network) in each hemisphere with the within-pair differences in self-reported music practice (one correlation with total hours of training and one correlation with early practice, age 6–11). The correlation with early practice was added because it is suggested that early practice plays a particular role for brain plasticity (Bengtsson et al. 2005). These supplementary analyses were carried out in STATISTICA. Again, the results were considered significant if $P < 0.05$, one-tailed. Finally, we calculated the within-pair correlation in mean cortical thickness in each hemisphere based on values extracted using FreeSurfer and the full DKT atlas.

FA Analysis

Preprocessing of Diffusion-Weighted Images

Diffusion data could not be obtained from one participant due to technical difficulties during scanning. The FA analysis was consequently limited to 16 individuals (i.e., 8 complete pairs). The preprocessing pipeline can be found in the Supplementary material. Fiber orientation distributions (FODs) were computed using robust constrained spherical deconvolution using the group average response function at $l_{\max} = 6$ (Tournier et al. 2007) as implemented in MRtrix3 (Tournier et al. 2012). A group-specific unbiased FOD template was created (Raffelt et al. 2011, 2012) and the FOD images from all participants were registered to this template. Upsampled versions of the brain mask images were also registered to the FOD template and a group brain mask was created based on the intersection of these images. A group mean FA image was also created.

Definition of ROIs and Probabilistic Streamlines Tractography in Template Space

In this analysis, we targeted the key white-matter tracts which enable auditory-motor integration and motor execution and have been identified in previous studies. Firstly, there are 2 main pathways by which auditory/music related information in the temporal lobe (from regions shown to differ between musicians and nonmusicians, see Cortical thickness analysis) can reach the ventral premotor and inferior frontal cortex: A dorsal pathway, projecting from primarily the auditory regions in the superior posterior temporal lobe and temporoparietal junction via the arcuate fasciculus (AF) (Halwani et al. 2011), and a ventral path which relays information from more inferior temporal regions to the inferior frontal cortex via the inferior longitudinal fasciculus (ILF) and the uncinate fasciculus (UF) (Dohn et al. 2015). Secondly, we wanted to identify the descending motor pathway/corticospinal tract (CST), specifically fibers originating in the hand/finger region of M1 (Bengtsson et al. 2005; Han et al. 2009). Thirdly, studies have shown portions of the corpus callosum (CC) to be larger or more developed in musicians compared with nonmusicians, presumably as a consequence of extensive training of complex bimanual movements and interhemispheric communication (Schlaug, Jäncke, Huang, Staiger, et al. 1995; Schmithorst and Wilke 2002; Steele et al. 2013).

The overall aim of the fiber tractography was to define these tracts of interest and create spatial masks within which to calculate and compare FA between the playing and nonplaying twins. In voxels which contain more than one fiber bundle, it is obviously a nontrivial matter to assign a particular FA value to each fiber bundle. Groeschel and colleagues (Groeschel et al. 2014), while comparing FA in a number of tracts in adolescents born preterm and controls, demonstrated different effects in regions with different underlying fiber architecture. In regions with predominantly single fiber pathways, there were significant differences in FA between the groups; averaging data over entire tracts obscured these differences and resulted in a loss of sensitivity. Thus, the seed and target regions for the CST, AF, ILF, and UF were defined manually on the group FA and FOD template images and placed to truncate the tracts before they would greatly spread across an innervated cortical region (and mingle with local fibers) or greatly cross with other tracts (the regions for the CC were defined in a different manner, see below). This was achieved by (1) creating a seed region at the middle of the tract, (2) performing bidirectional probabilistic fiber tracking, and (3) drawing the final ROIs using the produced streamlines as a reference, in conjunction with the template images. For the CST (in each hemisphere), we created a seed region in the cerebral peduncles above the level of the pontine transverse fibers, a waypoint mask in the posterior limb of the internal capsule and a waypoint/target mask beneath the hand area of M1. For the AF, we defined a seed region just above Heschl's gyrus, a waypoint mask at $y = -10$ (roughly half-way), and a target region before the tract would cross with the UF and forceps minor and further spread out to the different subregions of the IFG. Using the same procedure, seed, waypoint and target regions were created also for the UF and ILF. After all ROIs had been defined, unidirectional probabilistic fiber tracking of each tract was performed using default parameters in template space (Fig. 1*B*). The produced streamlines were visually inspected, ROIs were refined and NOT-gates were added where necessary to restrict streamlines from wandering. In this process, we additionally found that changing the FA-cutoff to 0.2 for the CST, ILF and UF, and to 0.15 for the AF produced more accurate tracts and less spurious streamlines.

For the CC, 5 ROIs (CC I–V) where defined along the midline according to the parcellation scheme developed by Hofer and Frahm (Hofer and Frahm 2006). Bidirectional probabilistic fiber tracking was then performed for each ROI, using an FA-cutoff of 0.2 and the other CC ROIs as NOT-gates (Fig. 1*B*). After examining the produced pathways, the maximum length of produced streamlines was set to 25 mm. This allowed for fiber tracking within the predominantly single-fiber white matter portion of the CC.

Definition of ROIs and Probabilistic Streamlines Tractography in Native Space

All ROIs produced in template space were transformed to the native space of each participant using nearest neighbor interpolation. The transformed ROIs were visually inspected and in the rare case manually edited to ensure accurate anatomical placement (typically by extending a NOT-gate to prevent spurious streamlines).

Probabilistic streamlines tractography was subsequently performed for each tract and participant using their FOD images and ROI as input. The following adjustments were made to the default parameters in line with the fiber tracking performed in template space (see above): the FA-cutoff was set to 0.2 for the CC, CST, ILF and UF, and 0.15 for the AF. Tracking was unidirectional and stopped at the final waypoint, except for the CC pathways, for which tracking was bidirectional with a streamline maximum length of 25 mm. In addition, the number of selected streamlines was fixed to 10^4 for all tracts, which ensured that no bias was introduced with regard to hemispheres and/or participants.

For each produced fiber tract in each participant, a corresponding track density image was created with voxel values equal to the fraction of passing streamlines. An absolute threshold of 0.01 (1% of the total streamline count) was applied to this image to remove voxels with spurious streamlines and create a binary mask image. This resulting image was finally implemented as an inclusive mask when extracting the tract's average FA value.

Data Analysis

The FA analysis was performed using the same approach and software as in the cortical thickness analysis (see above). The within-pair difference in FA between playing and nonplaying twins was analyzed using a linear mixed model, where FA was regressed on musicianship, while including tracts of interest as a fixed effect and pair membership as a random effect. Again, this approach was adopted to test whether there was, in line with our alternative hypothesis, an overall effect of musical training on brain regions, in this case white matter regions, central to auditory and motor processing during musical performance. We similarly produced and estimated one model for each hemisphere using REML and Satterthwaite approximations for the degrees of freedom. Results were considered significant if $P < 0.05$, one-tailed, expecting a relatively greater FA in the playing twins. Lastly, we correlated the within-pair differences in mean FA in each hemisphere and corpus callosum (separately) with the within-pair differences in self-reported music practice (one correlation with total hours of training and one correlation with early practice, age 6–11).

Voxel-Based Morphometry Analysis of the Cerebellum

Data Preprocessing

The voxel-based morphometry (VBM) analysis (Ashburner and Friston 2000) of the cerebellum was performed using the SUIT toolbox (Diedrichsen et al. 2009, 2011) implemented in SPM12. SUIT is a high-resolution atlas template of the human cerebellum and brainstem. The atlas is spatially unbiased and preserves the anatomical detail of cerebellar structures through a nonlinear atlas-generation algorithm. By using automated nonlinear normalization methods, a more accurate intersubject-alignment can be achieved than with current whole-brain methods. The preprocessing pipeline can be found in the Supplementary material.

VBM data for the ROIs were obtained using the summary function in the SUIT toolbox. The function finds for each image all voxels within 28 cerebellar compartments split by lobule and vermis/hemisphere as defined by the supplied probabilistic cerebellar atlas. Among these compartments we selected and extracted data from 2 ROIs in each hemisphere (lobules I–IV + V; and lobules VIIIa + VIIIb), which functionally represent the sensorimotor portions of the cerebellum involved in fine hand motor tasks (Stoodley et al. 2012). Thus data from the 2 compartments representing lobules I–IV and V were averaged to create one ROI in each hemisphere. In a similar way, the data from lobules VIIIa and VIIIb were averaged to create one ROI in each hemisphere. The ROIs are illustrated in Figure 1*C*.

Data Analysis

The regional differences in gray matter volume between playing and nonplaying twins were analyzed with paired *t*-tests in STATISTICA. The results were Bonferroni-corrected for multiple comparisons and reported significant if $P < 0.05$, one-tailed (expecting a larger volume for musicians). We additionally correlated the within-pair differences in mean volume for each of the ROIs with the within-pair differences in self-reported music practice (one correlation with total hours of training and one correlation with early practice, age 6–11).

Results

The results revealed that even in our sample of genetically identical twins, there were within-pair differences in the auditory-motor network between the playing and nonplaying siblings. Using the linear mixed models approach, we found that the playing twins had greater cortical thickness in the left cerebral auditory-motor network ($\beta = 0.068$, standard error (SE) = 0.020, $P < 0.001$) and more developed white matter microstructure in both the left hemisphere ($\beta = 0.022$, $SE = 0.006$, $P < 0.001$), the right hemisphere ($\beta = 0.016$, $SE = 0.005$, $P < 0.002$) and the corpus callosum ($\beta = 0.017$, $SE = 0.006$, $P < 0.004$). Since the linear mixed model, similar to an ANOVA, does not reveal whether the effect is greater in one ROI or another, we additionally performed exploratory paired t-tests (playing vs. nonplaying twins) for each ROI post-hoc. The significant results from the exploratory post-hoc paired t-tests (playing vs. nonplaying twins) for each region of interest (ROI) and tract of interest can be found in the Supplementary material, Tables S1 and S2, respectively. Further, the volume of gray matter in the left cerebellar lobules I–IV + V was greater in the playing group ($t(8) = 4.9$, $P = 0.004$). It can be noted that using two-tailed tests, instead of the hypothesis-driven one-tailed tests, would not change any outcome. The within-pair difference in total practice ranged between 1768 and 9516 h, $M = 4420 \pm 2730$ h. The within-pair difference in early practice ranged between −416 and 1092 h, $M = 225 \pm 489$ h. None of the twins had played music professionally. There was no significant correlation between the within-pair differences in piano practice and any of brain measures. The within-pair correlation of mean cortical thickness was $r = 0.93$, $P < 0.001$ in the left hemisphere and $r = 0.72$, $P = 0.03$ in the right hemisphere. There was no significant paired difference in the mean cortical thickness between the playing and nonplaying twins in either hemisphere.

Discussion

In light of our findings, we conclude that even when controlling for genes and early shared environment, there can be observable neuroanatomical differences in both gray matter and white matter microstructure between individuals that differ vastly in musical training. It appears highly likely, therefore, that causal effects of training underlie a significant portion of the differences in neuroanatomy between experts and nonexperts. This, obviously, does not negate genetic contributions to such differences, which would appear between individuals of different genotypes (e.g., between different twin pairs).

As outlined in the introduction, this was not certain. The influential deliberate practice theory, which states that long-term goal-directed practice is both necessary and sufficient to account for expert performance (Ericsson et al. 1993), and by implication its neural correlates, has been seriously challenged in recent years. With regard to our model behavior, that is, musical skills, Mosing, Madison, et al. (2014) found substantial genetic influence on musical practice. Further, the relation between practice and musical auditory discrimination ability was entirely driven by genetic pleiotropy (i.e., shared genetic factors), rather than causal effects of practice. Thus, correlations between practice and performance can be greatly inflated by genetic factors and cannot be taken as evidence for a causal effect of practice on either performance or associated brain measures. The general conclusion is that expertise relies on variables in addition to practice (Ullén et al. 2016). The practice-independent variance in performance is likely related to individual differences in traits of relevance for the particular domain of expertise. For example, musical auditory discrimination correlates genetically with intelligence (Mosing, Pedersen, et al. 2014).

From an empirical standpoint, the influence of genes on skill acquisition has been the proverbial elephant in the room since neuroimaging on experts began in the 1990s. In our view, this is not due to negligence, but rather the methodological challenges of conducting a randomized controlled trial for years or even decades in order to acquire the required data. As mentioned in the introduction, the existing studies have instead relied on cross-sectional or longitudinal observational data and short-term interventions without any long-term follow-up (Moreno and Bidelman 2014) and are therefore unable to exclude genetic influences as a potential confound. Here, we were able to address this issue in a novel way, by employing a monozygotic cotwin control design, which is essentially elephant-proof. Considering the somewhat modest sample size, the observed expertise-related differences in neuroanatomical measures effects were presumably discernable based on the closely matched comparisons between monozygotic twins. The within-pair correlation of mean cortical thickness in the left hemisphere was $r = 0.93$ and would likely have been even higher were it not for the large discordance in specific sensorimotor training associated with piano practice, which presents as anatomical differences in the network of regions involved in the trained behavior.

The reasons for the within-pair discordance in piano practice were investigated in interviews with the participants (Eriksson et al. 2016). We learned that both twins in each pair had started playing the piano at about the same age (between 7 and 10 years), but one sibling had stopped playing in childhood while the other had continued and was still playing. No systematic reasons for the discordance were revealed. Instead several unique environmental factors were present, which varied across pairs, such as differences in friends' music interests, having or not having the family piano in the bedroom, or exploring different creative outlets. In only one pair, both siblings felt that the playing twin was more "talented" from the outset. Nonetheless, the playing twins from early on showed a higher interest for music, enjoyed music more and showed a higher level of motivation (i.e., experienced more enjoyment when practicing the piano). In several of the pairs, the playing twin was also said to listen more to music from an early age. The discordance may instead be related to unspecific developmental factors as has been suggested by animal research that shows that even in genetically identical individuals reared together, differences in structural brain plasticity and behavior is to some extent an inevitable and potentially unpredictable outcome of development (Freund et al. 2013; Bierbach et al. 2017).

Not accounted for in this study, nor in the previous literature, is the role of gene–environment interactions for skill acquisition. All else being equal, nerve cells function according to the genetic instructions contained within their DNA. This includes how they respond and adapt to certain activity, such as that induced by musical practice. If there are individual differences in the domain general or domain specific genes that regulate such processes, identical exposure to an environmental factor will result in characteristic differences in brain plasticity, in terms of the quality, magnitude or specificity of neuroanatomical adaptations. Consequently, even in a randomized longitudinal study where all participants are complete novices and matched on relevant variables at the outset, outcomes after the intervention might vary with a positively skewed distribution. The reason would be individual differences in genes that synergize with practice, which enable some participants to excel beyond their peers. Such effects could also account for mean differences between experts and controls. The success of the intervention would likely be attributed to practice alone, but in reality a favorable genetic predisposition among some participants would also be responsible. In the present sample, there were no correlations between the within-pair differences in practice and the within-pair differences in any of the associated brain measures, that is, the relationship between training and effect was not equal across pairs. Although we obviously have limited power to detect such relations in the present sample, one plausible explanation is that there were differences in genetic predisposition between pairs that modulated the influence of practice on brain plasticity. An alternative or additional explanation, is that there were between-pair differences in the quality of practice. It can be assumed that within this sample, there exists a larger variation in both contextual and psychological practice-related variables, such as frequency and regularity of sessions, personal goals, commitment, musical content etc. It should also be noted that even though all playing twins had practiced the piano for more than 25 years, none were professional musicians. The fact that the playing twins were amateur pianists might also explain why they did not have larger cortical thickness than their siblings in the right hemisphere, though this was included in our hypotheses. Compared with professionals, amateur pianists would likely use their left hand for simpler chord progressions and less complex accompaniment, with lower demands on bimanual coordination, independent finger movements, motor planning and sequencing. Consequently, neural adaptations in the right cortical auditory-motor network might be less pronounced and would perhaps require a larger sample to be observed.

In conclusion, by ruling out pure genetic influences as a causal factor for the first time, the findings reported here give legitimacy to the notion that extensive musical training furthers the development of the auditory-motor network to such an extent that we can observe the effects even at a macroanatomical level.

Supplementary Material

Supplementary data is available at *Cerebral Cortex* online.

Funding

The Bank of Sweden Tercentenary Foundation (M11–0451:1), the Sven and Dagmar Salén Foundation, the Kjell and Märta Beijer Foundation and Hjärnfonden (FO2015–0267).

Notes

We thank Dr Miriam Mosing for assistance during experiments, Dr Rita Almeida for consultations on statistics and Dr Laura Babcock for reviewing the manuscript. *Conflict of Interest*: None declared.

References

Abdul-Kareem IA, Stancak A, Parkes LM, Sluming V. 2011. Increased gray matter volume of left pars opercularis in male orchestral musicians correlate positively with years of musical performance. J Magn Reson Imaging. 33:24–32.

Amunts K, Schlaug G, Jäncke L, Steinmetz H, Schleicher A, Dabringhaus A, Zilles K. 1997. Motor cortex and hand motor skills: structural compliance in the brain. Hum Brain Mapp. 5:206–215.

Ashburner J, Friston KJ. 2000. Voxel-based morphometry—the methods. NeuroImage. 11:805–821.

Bailey JA, Zatorre RJ, Penhune VB. 2014. Early musical training is linked to gray matter structure in the ventral premotor cortex and auditory-motor rhythm synchronization performance. J Cogn Neurosci. 26:755–767.

Bangert M, Schlaug G. 2006. Specialization of the specialized in features of external human brain morphology. Eur J Neurosci. 24:1832–1834.

Bengtsson SL, Nagy Z, Forsman L, Forssberg H, Skare S, Ullén F. 2005. Extensive piano practicing has regionally specific effects on white matter development. Nat Neurosci. 8: 1148–1150.

Berkowitz AL, Ansari D. 2010. Expertise-related deactivation of the right temporoparietal junction during musical improvisation. NeuroImage. 49:712–719.

Bermudez P, Lerch JP, Evans AC, Zatorre RJ. 2009. Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. Cereb Cortex. 19:1583–1596.

Bierbach D, Laskowski KL, Wolf M. 2017. Behavioural individuality in clonal fish arises despite near-identical rearing conditions. Nat Commun. 8:15361.

Brown S, Martinez MJ, Parsons LM. 2006. Music and language side by side in the brain: a PET study of the generation of melodies and sentences. Eur J Neurosci. 23:2791–2803.

Diedrichsen J, Balsters JH, Flavell J, Cussans E, Ramnani N. 2009. A probabilistic MR atlas of the human cerebellum. NeuroImage. 46:39–46.

Diedrichsen J, Maderwald S, Kuper M, Thurling M, Rabe K, Gizewski ER, Ladd ME, Timmann D. 2011. Imaging the deep cerebellar nuclei: a probabilistic atlas and normalization procedure. NeuroImage. 54:1786–1794.

Dohn A, Garza-Villarreal EA, Chakravarty MM, Hansen M, Lerch JP, Vuust P. 2015. Gray- and white-matter anatomy of absolute pitch possessors. Cereb Cortex. 25:1379–1388.

Ericsson KA, Krampe RT, Tesch-Römer C. 1993. The role of deliberate practice in the acquisition of expert performance. Psychol Rev. 100:363–406.

Eriksson H, Harmat L, Theorell T, Ullén F. 2016. Similar but different: Interviewing monozygotic twins discordant for musical practice. Music Sci. 21:1–17.

Eyler LT, Chen CH, Panizzon MS, Fennema-Notestine C, Neale MC, Jak A, Jernigan TL, Fischl B, Franz CE, Lyons MJ, et al. 2012. A comparison of heritability maps of cortical surface area and thickness and the influence of adjustment for whole brain measures: a magnetic resonance imaging twin study. Twin Res Hum Genet. 15:304–314.

Fauvel B, Groussard M, Chetelat G, Fouquet M, Landeau B, Eustache F, Desgranges B, Platel H. 2014. Morphological brain plasticity induced by musical expertise is accompanied by modulation of functional connectivity at rest. NeuroImage. 90:179–188.

Fischl B, Dale AM. 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc Natl Acad Sci USA. 97:11050–11055.

Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, et al. 2004. Automatically parcellating the human cerebral cortex. Cereb Cortex. 14:11–22.

Freund J, Brandmaier AM, Lewejohann L, Kirste I, Kritzler M, Kruger A, Sachser N, Lindenberger U, Kempermann G. 2013. Emergence of individuality in genetically identical mice. Science. 340:756–759.

Gaser C, Schlaug G. 2003. Brain structures differ between musicians and non-musicians. J Neurosci. 23:9240–9245.

Groeschel S, Tournier JD, Northam GB, Baldeweg T, Wyatt J, Vollmer B, Connelly A. 2014. Identification and interpretation of microstructural abnormalities in motor pathways in adolescents born preterm. NeuroImage. 87:209–219.

Halwani GF, Loui P, Ruber T, Schlaug G. 2011. Effects of practice and experience on the arcuate fasciculus: comparing singers, instrumentalists, and non-musicians. Front Psychol. 2:156.

Han Y, Yang H, Lv Y-T, Zhu C-Z, He Y, Tang H-H, Gong Q-Y, Luo Y-J, Zang Y-F, Dong Q. 2009. Gray matter density and white matter integrity in pianists' brain: a combined structural and diffusion tensor MRI study. Neurosci Lett. 459:3–6.

Hickok G, Buchsbaum B, Humphries C, Muftuler T. 2003. Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. J Cogn Neurosci. 15:673–682.

Hofer S, Frahm J. 2006. Topography of the human corpus callosum revisited—comprehensive fiber tractography using diffusion tensor magnetic resonance imaging. NeuroImage. 32: 989–994.

Hyde KL, Lerch J, Norton A, Forgeard M, Winner E, Evans AC, Schlaug G. 2009. Musical training shapes structural brain development. J Neurosci. 29:3019–3025.

James CE, Oechslin MS, Van De Ville D, Hauert CA, Descloux C, Lazeyras F. 2014. Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks. Brain Struct Funct. 219:353–366.

Klein A, Tourville J. 2012. 101 labeled brain images and a consistent human cortical labeling protocol. Front Neurosci. 6:171.

Lahav A, Saltzman E, Schlaug G. 2007. Action representation of sound: audiomotor recognition network while listening to newly acquired actions. J Neurosci. 27:308–314.

Macnamara BN, Hambrick DZ, Oswald FL. 2014. Deliberate practice and performance in music, games, sports, education, and professions: a meta-analysis. Psychol Sci. 25:1608–1618.

Macnamara BN, Moreau D, Hambrick DZ. 2016. The relationship between deliberate practice and performance in sports: a meta-analysis. Perspect Psychol Sci. 11:333–350.

Markham JA, Greenough WT. 2004. Experience-driven brain plasticity: beyond the synapse. Neuron Glia Biol. 1:351–363.

McGue M, Osler M, Christensen K. 2010. Causal inference and observational research: the utility of twins. Perspect Psychol Sci. 5:546–556.

Moreno S, Bidelman GM. 2014. Examining neural plasticity and cognitive benefit through the unique lens of musical training. Hear Res. 308:84–97.

Mosing MA, Madison G, Pedersen NL, Kuja-Halkola R, Ullén F. 2014. Practice does not make perfect: no causal effect of music practice on music ability. Psychol Sci. 25:1795–1803.

Mosing MA, Pedersen NL, Madison G, Ullén F. 2014. Genetic pleiotropy explains associations between musical auditory discrimination and intelligence. PLoS One. 9:e113874.

Peper JS, Brouwer RM, Boomsma DI, Kahn RS, Hulshoff Pol HE. 2007. Genetic influences on human brain structure: a review of brain imaging studies in twins. Hum Brain Mapp. 28: 464–473.

Platz F, Kopiez R, Lehmann AC, Wolf A. 2014. The influence of deliberate practice on musical achievement: a meta-analysis. Front Psychol. 5:646.

Raffelt D, Tournier JD, Crozier S, Connelly A, Salvado O. 2012. Reorientation of fiber orientation distributions using apodized point spread functions. Magn Reson Med. 67:844–855.

Raffelt D, Tournier JD, Fripp J, Crozier S, Connelly A, Salvado O. 2011. Symmetric diffeomorphic registration of fibre orientation distributions. NeuroImage. 56:1171–1180.

Saur D, Kreher BW, Schnell S, Kummerer D, Kellmeyer P, Vry MS, Umarova R, Musso M, Glauche V, Abel S, et al. 2008. Ventral and dorsal pathways for language. Proc Natl Acad Sci USA. 105:18035–18040.

Schlaug G, Jäncke L, Huang Y, Staiger JF, Steinmetz H. 1995. Increased corpus callosum size in musicians. Neuropsychologia. 33:1047–1055.

Schlaug G, Jäncke L, Huang Y, Steinmetz H. 1995. In vivo evidence of structural brain asymmetry in musicians. Science. 267:699–701.

Schmithorst VJ, Wilke M. 2002. Differences in white matter architecture between musicians and non-musicians: a diffusion tensor imaging study. Neurosci Lett. 321:57–60.

Schneider P, Scherg M, Dosch HG, Specht HJ, Gutschalk A, Rupp A. 2002. Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. Nat Neurosci. 5:688–694.

Sluming V, Barrick T, Howard M, Cezayirli E, Mayes A, Roberts N. 2002. Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. NeuroImage. 17:1613–1622.

Steele CJ, Bailey JA, Zatorre RJ, Penhune VB. 2013. Early musical training and white-matter plasticity in the corpus callosum: evidence for a sensitive period. J Neurosci. 33: 1282–1290.

Stoodley CJ, Valera EM, Schmahmann JD. 2012. Functional topography of the cerebellum for motor and cognitive tasks: an fMRI study. NeuroImage. 59:1560–1570.

Tournier J-D, Calamante F, Connelly A. 2012. MRtrix: diffusion tractography in crossing fiber regions. Int J Imag Syst Tech. 22:53–66.

Tournier JD, Calamante F, Connelly A. 2007. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. NeuroImage. 35:1459–1472.

Ullén F, Hambrick DZ, Mosing MA. 2016. Rethinking expertise: a multifactorial gene-environment interaction model of expert performance. Psychol Bull. 142:427–446.

Wenger E, Kuhn S, Verrel J, Martensson J, Bodammer NC, Lindenberger U, Lövden M. 2017. Repeated structural imaging reveals nonlinear progression of experience-dependent volume changes in human motor cortex. Cereb Cortex. 27: 2911–2925.