Seismic anisotropy of Precambrian lithosphere: Insights from Rayleigh wave tomography of the eastern Superior Craton

Laura Petrescu¹,², Fiona Darbyshire³, Ian Bastow⁴, Eoghan Totten¹,⁴, and Amy Gilligan¹,⁵

¹Department of Earth Science and Engineering, Imperial College London, London, UK, ²National Institute for Earth Physics, Magurele, Romania, ³Centre de recherche GEOTOP, Université du Québec à Montréal, Montréal, Quebec, Canada, ⁴Now at Department of Earth Sciences, University of Oxford, Oxford, UK, ⁵Now at School of Geosciences, University of Aberdeen, Aberdeen, UK

Abstract The thick, seismically fast lithospheric keels underlying continental cores (cratons) are thought to have formed in the Precambrian and resisted subsequent tectonic destruction. A consensus is emerging from a variety of disciplines that keels are vertically stratified, but the processes that led to their development remain uncertain. Eastern Canada is a natural laboratory to study Precambrian lithospheric formation and evolution. It comprises the largest Archean craton in the world, the Superior Craton, surrounded by multiple Proterozoic orogenic belts. To investigate its lithospheric structure, we construct a frequency-dependent anisotropic seismic model of the region using Rayleigh waves from teleseismic earthquakes recorded at broadband seismic stations across eastern Canada. The joint interpretation of phase velocity heterogeneity and azimuthal anisotropy patterns reveals a seismically fast and anisotropically complex Superior Craton. The upper lithosphere records fossilized Archean tectonic deformation: anisotropic patterns align with the orientation of the main tectonic boundaries at periods ≤110 s. This implies that cratonic blocks were strong enough to sustain plate-scale deformation during collision at 2.5 Ga. Cratonic lithosphere with fossil anisotropy partially extends beneath adjacent Proterozoic belts. At periods sensitive to the lower lithosphere, we detect fast, more homogenous, and weakly anisotropic material, documenting postassembly lithospheric growth, possibly in a slow or stagnant convection regime. A heterogeneous, anisotropic transitional zone may also be present at the base of the keel. The detection of multiple lithospheric fabrics at different periods with distinct tectonic origins supports growing evidence that cratonic processes may be episodic and are not exclusively an Archean phenomenon.

Plain Language Summary The roots of the oldest parts of the continents are unusually thick (~250 km) and preserve structures dating back to the Earth’s most ancient geological period, >2.5 billion years ago. The structure, formation, and preservation of these thick continental roots are still debated. Eastern Canada has a geological record stretching back over more than 3 billion years, providing an excellent opportunity to study the Earth’s early history and the formation and evolution of continents. To investigate the deep structure beneath this region, we use earthquake energy recorded at seismic stations across eastern Canada to build an image of the continental interior. Our results show that the ancient continental root is made up of at least two layers, each with a different formation mechanism. The upper layer preserves the signature of continental deformation older than 2.5 billion years, whereas the lower layer properties suggest later downward growth of the root above a slow-moving mantle. The results support the idea that the thick continental roots formed episodically and across multiple geological periods.

1. Introduction

Cratons are the geologically distinct Archean (>2.5 Ga) cores of the continents that have survived multiple Wilson cycles. They are typically underlain by thick (<300 km), seismically fast [e.g., Schaeffer and Lebedev, 2013] lithospheric roots (keels) that may be highly depleted in Ca, Al, Fe, and Mg [Griffin et al., 2003]. Some keels have been found to comprise fast wave speed, chemically depleted upper layers, underlain by a cold but compositionally less distinct thermal boundary layer [Yuan and Romanowicz, 2010].
The deep structure of cratons worldwide has been analyzed using a range of seismic imaging methods and through kimberlite-borne mantle xenoliths. Broadband seismology studies detected multiple intralithospheric seismic discontinuities [e.g., Romanowicz, 2009; Abt et al., 2010; Wirth and Long, 2014] that may represent the bottom of the chemically depleted layers and younger cratonic roots. Mantle xenoliths in some cratons (e.g., the Slave Craton in NW Canada and the North Atlantic Craton in Greenland [Griffin et al., 1999; Kopylova and Russell, 2000; Bizzarro and Stevenson, 2003]) record compositional stratification, although this is not evident in other cratons such as the Siberian or the Kaapvaal Cratons [e.g., Boyd et al., 1997]. The existence of multiple anisotropic layers has also been reported, with a change from lithospheric fabrics to present-day mantle flow beneath several stable continental regions, e.g., North America [Darbyshire and Lebedev, 2009; Yuan and Romanowicz, 2010], South Africa [e.g., Adam and Lebedev, 2012], northern Europe [e.g., Plomerová and Babuska, 2010], and Australia [e.g., Fichtner et al., 2010].

The tectonic processes that led to keel development remain uncertain. A range of contrasting generation mechanisms has been proposed, including mantle plume melt extraction [e.g., Arndt et al., 2009], catastrophic delamination events [Bédard, 2006], and multiple island arc accretion episodes [e.g., Miller and Eaton, 2010]. The origin of the intracratonic layers and discontinuities has also been attributed to metasomatism [e.g., Savage and Silver, 2008], partial melt [e.g., Thybo, 2006], elastically accommodated grain boundary sliding [Karato, 2012], decrease in Mg content with depth [e.g., Yuan and Romanowicz, 2010], and stacked subduction of noneclogitized oceanic crust [e.g., Bostock et al., 2010]. Key to discriminating these hypotheses and to understanding the nature and relative timing of cratonization processes is an improved detailed knowledge of the lithospheric structure beneath Archean continental regions and its subsequent deformation history.

Eastern Canada is home to one of the largest and best preserved Archean cratons in the world, the Superior Craton, thought to have been assembled by 2.5 Ga [Hoffman, 1989]. Its edges experienced collisional orogenies and Andean-style subduction throughout the Proterozoic [e.g., Rivers, 1997] and witnessed the closure of the Rheic and Iapetus Oceans during the Phanerozoic [van Staal and Barr, 2012]. Eastern Canada thus provides an excellent opportunity to gain fundamental insights about the types of tectonic processes that formed the cratonic mantle and subsequently modified it.

To investigate the geological structures beneath eastern Canada, we use earthquake data to construct a tomographic image of the subsurface in period domain. Using the dispersion properties of surface waves, we estimate the variation of isotropic and anisotropic components of the Rayleigh wave phase velocity [Deschamps et al., 2008], between 20 and 220 s period, broadly sensitive to depths of 30–350 km. Variations in phase velocity are used to map out the extent of the cratonic core beneath the study region. Measurements of seismic anisotropy allow us to observe fabrics from episodes of tectonic deformation that affected various parts of the cratonic keel. Our results provide evidence for the multistage mechanisms that participated in the formation and evolution through space and time of a layered cratonic mantle.

2. Tectonic Setting

The geology of Eastern Canada documents tectonic events dating back to the Archean eon (>2.5 Ga). It includes one of the largest cratons in the world, the Superior Craton (Figure 1), which was part of the Laurentia paleocontinent. The Superior Craton is composed of several microcratonic blocks, whose tectonic assemblage was completed by ~2.5 Ga [Hoffman, 1989]. The dominant direction of Archean surface tectonic boundaries changes from E-W in the south and west of the province to N-S in the northeast (Figure 1).

The edges of the Superior Craton experienced multiple supercontinent and Wilson cycles, including the Paleoproterozoic Trans-Hudson orogeny to the north and west and the New Quebec orogeny to the northeast, respectively (Figure 1). The latter is part of the Eastern Churchill Province [Wardle et al., 1990], comprising an Archean core trapped between two Paleoproterozoic N-S trending orogenic belts, the New Quebec and Torngat orogens.

A long-lived Andean-style paleosubduction system is thought to have dominated the eastern cratonic margins of the Superior between ~1.7 Ga and ~1.2 Ga [Rivers, 1997]. Between 1090 Ma and 980 Ma, the Superior Craton was involved in a Himalayan-scale collision with Amazonia [Li et al., 2008] as part of one of the most important supercontinent cycles in the Proterozoic. The Grenville Province is the remnant of this Proterozoic orogeny and comprises rocks that were originally part of the Archean Superior Craton [Rivers, 2009].
Southeast of the Grenville Province, the Appalachian orogenic belt records the closure of the Iapetus and Rheic Oceans between 480 Ma and 260 Ma [van Staal and Barr, 2012].

Between 160 Ma and 80 Ma, the North American plate drifted over the Great Meteor hot spot, with surface geological signatures such as kimberlite pipes in the Superior Craton and a track of igneous intrusions in the Grenville Province [Sleep, 1990].

3. Method and Data

3.1. Rayleigh Wave Dispersion

Rayleigh waves travel with phase velocities that are frequency dependent and are a powerful tool to study upper mantle structure. By cross-correlating surface waveforms simultaneously recorded at pairs of seismic stations, it is possible to estimate the path-averaged Rayleigh wave phase velocity between them as a function of wave period [Meier et al., 2004].

We processed seismic records from 3218 teleseismic earthquakes with $M_w > 5.5$ between 2009 and 2014 that were recorded at 48 permanent and temporary broadband seismic stations in eastern Canada, Greenland, southern Baffin Island, and the northeast U.S. (Figure 1). Data from stations outside of the region of interest widens the aperture of the network, improving the resolution of our model at the edges. Of 1128 possible station pairs, ~500 recorded earthquakes along great-circle paths within ±5° of the interstation back azimuth and were thus usable for analysis.

The recorded displacement is approximated with a Jordan-Wenzel-Kramer-Brillouin (JWKB) response to a moment tensor source on an isotropic, spherically symmetric Earth model [Snieder and Nolet, 1987]. The vertical displacement seismograms recorded at pairs of seismic stations are shown in Figure 2, along with their amplitude as a function of time and frequency, to emphasize the frequency range for which fundamental mode Rayleigh waves can be used in the two-station method analysis. Strong scattering influences the 0.6–0.1 Hz frequency range (10–16 s period). The cross-correlation function is plotted in Figure 2e along with its frequency-time representation (Figure 2f). A frequency-dependent Gaussian bandpass filter was applied to enhance the signal-to-noise ratio and minimize noise and interferences with scattered waves. A subsequent time domain frequency-dependent Gaussian filter, centered around the peak amplitude time, was applied to attenuate the side lobe effects arising from fundamental mode interference with higher modes. The cross-correlation function is then transferred to the frequency domain, and its complex phase is used
Figure 2. The interactive cross-correlation procedure for constructing a dispersion curve. The vertical displacement seismograms of an earthquake as recorded at stations (a) ACCN and (c) SUNO. (b,d) The corresponding spectrograms of the recorded seismograms. (e) The time series cross correlation of the two plotted seismograms. (f) The corresponding spectrogram of the cross-correlation function. (g) Rayleigh wave phase velocity dispersion curves. Red lines are the array of curves resulting from the calculation of phase velocity from the phase of the cross-correlation function. Black dashed line is a theoretical dispersion curve calculated from the ak135 model. (h) The dispersion curves interactively selected from previous cross correlations of recorded seismic energy of earthquakes coming from the direction of station ACCN (black lines) or SUNO (green lines).

to estimate the frequency-dependent phase velocity, \( c(\omega) \), where \( \omega \) is the angular frequency (see Meier et al. [2004] and Sato [1955] for a comprehensive description of the method and related equations):

\[
    c(\omega) = \frac{\omega(\Delta_1 - \Delta_2)}{\arctan\left(\frac{\mathbb{R}\{\phi(\omega)\}}{\mathbb{I}\{\phi(\omega)\}}\right) + 2n\pi},
\]

where \( \Delta_1 \) and \( \Delta_2 \) are the epicentral distances of the two stations and \( \phi \) is the cross-correlation function. The \( 2n\pi \) term in the denominator emerges from the intrinsic ambiguity of the arctan function. The frequency-dependent phase velocity solution is thus nonunique, comprising a discrete set of dispersion curves, one for each value of \( n \) (Figure 2g). The most plausible value of \( n \) or, equivalently, the most likely dispersion curve, is recovered by visually inspecting the solution space and manually selecting the portion of the dispersion curve most similar to a theoretical global reference curve (Figure 2g). Cross-correlation procedures suffer from cycle-skipping problems, which may arise if the phase difference between the two signals is larger than half a period of the dominant wavelet. To minimize cycle skipping, we perform the computation of the phase starting from the longest period. Incorrect phase unwrapping will result in a discontinuity in the dispersion curves, which is easily detectable in the manual selection step. The final dispersion curve is the average of the selected solutions for all recorded earthquakes at the two seismic stations. Seismic diffraction and fundamental mode Rayleigh wave interference with higher modes can introduce irregularities and roughness in the measured curve [e.g., Legendre et al., 2014]. However, these effects are frequency dependent and are smoothed in the manual selection and averaging steps (Figures 2g and 2h). The resulting dispersion curve is a path-averaged phase velocity variation of the Rayleigh waves sampling the structures between the two stations. Using this method, we estimated 285 interstation dispersion curves (Figure 1) for the period range 20–220 s from \( \sim 180,000 \) seismograms. The number of curves accurately retrieved for each period varied between 189 and 276.
We minimize potential bias from off-great-circle propagation [Pedersen, 2006] by combining dispersion measurements from multiple earthquakes at different epicentral distances and reciprocal paths for each station pair (see Appendix A for more information).

3.2. Seismic Tomography

Each dispersion curve yields an average phase velocity along the interstation path it was estimated for, as a function of period. The total phase velocity anomaly $\delta \bar{C}$ can be written as the integral of all the local anomalies $(\delta C(\varphi, \theta))$ at each grid knot of latitude $\varphi$, longitude $\theta$, along the corresponding path $i$:

$$
\delta \bar{C}(\varphi, \theta) = \int_\delta \int_\varphi K_i(\varphi, \theta) \delta C(\varphi, \theta) d\theta d\varphi, \tag{2}
$$

where index $i$ refers to the path for which the anomaly is computed and the kernels $K_i(\varphi, \theta)$ define the sensitivity area of the Rayleigh waveforms, regulating the weight of each grid point along the path [Lebedev and van der Hilst, 2008]. The dependence on $\omega$ is implied.

For each period, the local phase velocity anomaly at each grid knot, $\delta C(\varphi, \theta)$, can be decomposed into isotropic and anisotropic contributions, following a Fourier series expansion, in terms of the azimuth, $\psi$, of the wave number vector [Smith and Dahlen, 1973; Deschamps et al., 2008].

$$
\delta C(\psi) = \delta C_{iso} + A_{2\psi} \cos(2\psi) + B_{2\psi} \sin(2\psi) + A_{4\psi} \cos(4\psi) + B_{4\psi} \sin(4\psi), \tag{3}
$$

where the latitude $(\varphi)$, longitude $(\theta)$, and period $(2\pi/\omega)$ dependence is assumed. The $A$ and $B$ terms can further be combined to estimate the amplitude ($\Lambda$) and the angle of the fast $S$ wave polarization ($\Theta$), for both $2\psi$ and $4\psi$ contributions (see Appendix B). These are useful to visualize the anisotropy patterns across the model.

Equation (2) is then discretized, and a linear system of equations (one equation for each path, Appendix B) is built [Deschamps et al., 2008]. The system of equations is then solved independently at each frequency, using an LSQR algorithm (least squares with QR factorization [Paige and Saunders, 1982]), which does not provide a covariance matrix that can be used to quantify uncertainties. However, the fit to data may be expressed as the variance reduction of the model at each frequency/period, and resolution and leakage tests evaluate further the inversion’s ability to recover seismic structure (section 3.2.2).

3.2.1. Model Parametrization, Regularization, and Data Fit

The model was first parametrized with a triangular grid [Wang and Dahlen, 1995] with an approximately constant interknot spacing of 40 km. The spacing was chosen to be sufficiently small for an accurate integration of the Rayleigh wave sensitivity kernels, $K(\varphi, \theta)$ [e.g., Lebedev and van der Hilst, 2008]. A larger triangular model grid was then constructed, with a 200 km interknot spacing similar to the average interstation distance across the network.

The LSQR method [Paige and Saunders, 1982] used to solve the linear system of equations requires two types of damping. Gradient damping penalizes the first derivatives of the phase velocity anomaly distribution. Double derivative damping, or “smoothing,” moderates the second spatial derivatives, spreading out local anomaly peaks in the final velocity distribution [Darbyshire and Lebedev, 2009].

After visually inspecting different model outputs, for different damping and smoothing values, we chose the following combination for the isotropic, $2\psi$ and $4\psi$ velocity contributions, respectively: 0.05, 0.1, and 0.1 for damping and 0.5, 0.7, and 0.7 for smoothing, at all periods (see supporting information for more information).

3.2.2. Resolution and Leakage Tests

To estimate the resolution of the phase velocity distribution, we built pseudocheckerboard models (Figure 3), calculated synthetic dispersion data, and inverted the resulting dispersion curves to recover the synthetic model. The pseudocheckerboard tests are constructed by setting the velocity anomaly at model grid points to ±2.4% and interpolating the values on the dense integration grid. Since the model grid used in this kind of tomography is based on a triangular geometry, instead of the classic Cartesian framework [e.g., Nolet, 1993; Pawlak et al., 2011; Boyce et al., 2016], the anomaly shapes in Figure 3 are in fact the closest to a “standard” checkerboard. This type of synthetic test was successfully used in previous models [e.g., Deschamps et al., 2008; Darbyshire and Lebedev, 2009; Legendre et al., 2014]. The width of the negative anomalies that we test for ranges from 200 to 400 km. The ~200 km wide anomalies were recovered only in the central and southwest parts of the model (Figure 3). The ~300 km anomalies were relatively well recovered throughout the model space (within a ±1% error), except in the extreme north and northeast of the region. The ~400 km anomaly
Figure 3. Resolution tests for isotropic low velocity checkerboard anomalies at period 100 s. The peak amplitude of the anomalies is ±2.4%. (top row) The input model and (bottom row) the recovered model.

features were well recovered except in the extreme northeast, southeast, and southwest (Figure 3). We also note a ∼0.5% increase in the recovered anomaly amplitude south of Hudson Bay, at the westernmost edge of our model (Figure 3).

Imperfect amplitude recovery may be due to poor resolution at the model edges, sparse interstation path coverage, and low degree of path crosscutting, resulting in directional sampling bias. For this reason, we manually shadowed the edges of the model and the areas that we consider unreliable by simultaneously visualizing resolution and leakage tests and the path distribution at each period. We thus focused our interpretations on features that are within the resolution capability of the area. Smoothing was also found to affect resolution by decreasing the peak amplitude and spatially spreading out anomaly features in the distribution [e.g., Darbyshire and Lebedev, 2009]. The number of paths (189–276), directions, and degree of crosscutting vary slightly in the 30–220 s period range (Figure S1 in the supporting information). A more detailed discussion of path coverage and resolution at short periods can be found in the supporting information [Meier et al., 2004; Deschamps et al., 2008]. Resolution tests performed at longer periods provided similar solutions as for shorter periods (e.g., 100 s, Figure 3), demonstrating that the change in path coverage/crosscutting, the main control on the resolution, does not impact significantly enough to justify an increase in smoothing/damping.

Numerical leakage between the inverted isotropic and anisotropic velocity components can introduce false velocity anomalies during the inversion procedure. To explore the impact of such numerical artifacts, we performed three separate resolution tests, in which we inverted each component individually: (1) purely isotropic (to investigate leakage into anisotropy), (2) purely anisotropic (to investigate leakage into isotropy), and (3) a combination that includes rotated anisotropy. Figure 4 shows the recovery of these synthetic components and the kind of numerical artifacts that appeared on the other components, at 100 s and 180 s periods.

In all three cases, the magnitudes of falsely induced isotropic or anisotropic anomalies were similar for the two periods. In the purely isotropic Earth model case, a mean anisotropy anomaly magnitude of 0.3% was
Figure 4. Three types of leakage tests performed at periods of 100 s and 180 s. The isotropic leakage test comprises the synthetic purely isotropic Earth model inputs, as recovered by the initial full inversion, and the corresponding output models at periods (a, b) 100 s and (c, d) 180 s, respectively. The anisotropic leakage test comprises the synthetic purely anisotropic Earth model inputs, as recovered by the initial full inversion, and the corresponding output models at periods of (e, f) 100 s and (g, h) 180 s, respectively. The third test comprises the isotropic Earth models, as recovered by the initial full inversion, with anisotropy orientations rotated by 90°, and the corresponding output models at periods of (i, j) 100 s and (k, l) 180 s, respectively.
Figure 5. Rayleigh wave isotropic phase velocity anomaly and azimuthal anisotropy maps of eastern Canada at periods ranging from 20 to 110 s. Vref is the regional reference velocity in km/s for each period. Phase velocity results are plotted within ±50 km from any ray path. Shadowed areas represent model edges that we deemed unreliable after simultaneously considering leakage, resolution tests, and path coverage at each period. Bottom right diagram shows the Rayleigh wave depth sensitivity kernels [Lebedev and van der Hilst, 2008].

leaked into the $2\psi$ component, with some model edge regions exhibiting false anisotropy anomalies up to 1% (northwestern area). Isotropic velocity patterns were well recovered with slightly reduced amplitudes (~0.3%) in parts of the Superior Craton at 100 s and increased values (+0.3%) at 180 s. In the purely anisotropic Earth model case (either $2\psi$ or $4\psi$), there was little leakage into the isotropic component (mean isotropic phase
Figure 6. Rayleigh wave isotropic phase velocity anomaly and azimuthal anisotropy maps of eastern Canada at periods ranging from 120 to 220 s. Vref is the regional reference velocity in km/s for each period. Phase velocity results are plotted within ±50 km from any ray path. Shadowed areas represent model edges that we deemed unreliable after simultaneously considering leakage, resolution tests, and path coverage at each period. Bottom right diagram shows the Rayleigh wave depth sensitivity kernels [Lebedev and van der Hilst, 2008].

velocity leakage of 0.08%). The azimuthal anisotropy amplitudes and directions were well recovered, with an average anomaly reduction of ~0.2% and a mean azimuthal deviation from the input values of ~15°, mainly resulting from differences at the model edges. In the third test we inverted for an Earth model that comprises a combination of both isotropic and anisotropic components at 100 s and 180 s period. The anisotropy
Figure 7. (a) The variation of anisotropy magnitude with period at selected locations in the Superior Craton; UL, upper lithosphere; LL, lower lithosphere; TZ, transitional zone. (b) Standard deviation of phase velocity distribution at selected locations in the Superior Craton. (c) Sample locations for the selected regions in the Superior Craton. (d) Mean Rayleigh wave phase velocity dispersion at selected locations in the Superior Craton (bold dark grey line), mean regional phase velocity from our model (thin grey line), and synthetic dispersion curve calculated from the iasp91 global reference model (light grey line).

directions were rotated by 90°, to test for bias in the azimuth of the estimated $S$ wave polarization. Figure 4 shows that the rotated anisotropy directions were well recovered, with high azimuthal deviations (>60°) occurring in the northeast and the southernmost areas of the model. Together, these resolution tests provide further assurance that the inversion method as applied to our data set is able to generate credible tomographic models without significantly distorting each component or inducing significant numerical artifacts that could bias the interpretation.

4. Results

Figures 5 and 6 show the distribution of anisotropic Rayleigh wave phase velocities across eastern Canada for periods 20–220 s, with respect to the model average for each period. A first-order feature that dominates the model at periods $<140$ s is the strong difference between the seismically fast western section, which comprises provinces older than 2.5 Ga and the slow eastern section, comprising younger provinces.

The Superior Craton is seismically fast at periods $\leq 140$ s, compared to adjacent provinces, and exhibits multiple distinct zones of anisotropy. Phase velocity anomalies range between $\sim 0.5\%$ near the Proterozoic border between the Superior Craton and the Grenville Province and $\sim 2.5\%$ locally to the SE of Hudson Bay at periods sampling mostly crustal and uppermost mantle depths (Figure 5). The local high-velocity maxima relocates toward the center of the province at intermediate periods (60–100 s). The 20 s map shows consistently similar isotropic velocity features with maps at periods of 35 s, and similar anisotropy magnitude and directions in the Grenville Province. Azimuthal anisotropy magnitude ranges between 0.5% and 1% in the 20–100 s period interval. Anisotropy directions at short periods (20–40 s) are not distributed in a coherent fashion, although an interesting pattern becomes increasingly evident in the 40–110 s period interval (peak depth sensitivity between 50 and 150 km, Figure 5): fast velocity axes gradually rotate from predominantly E-W in the south to N-S in the northern Superior section.

In the 110–140 s period interval, surface waves mostly sensitive to 150–200 km depth range sample upper mantle material that appears to be faster than the regional average ($dC/C \approx 1\%$) and more homogenous...
(the mean standard deviation of absolute phase velocity values is <25 m/s). The anisotropy directions are less coherent (Figure 6), and the mean anisotropy magnitude decreases by ~0.5% (Figure 7a), indicating weakly anisotropic material beneath the province. The observed sudden drop in anisotropy anomaly strength is higher than twice the mean standard deviation in the distribution of anisotropy anomaly values estimated in the 110–140 s period interval in the Superior Province (~0.1%, Figure 7a), making it a robust feature of our model.

The 140–180 s period range reveals a highly heterogeneous ($dC/C \approx 3\%$) and anisotropic (~1.5%) mantle zone. At periods >180 s, the anisotropy increases significantly to 4% and strong heterogeneities are present ($dC/C \approx -1.5\%$ to 2.5%).

The Grenville Province is characterized by the transition from the positive seismic anomaly that dominates the Archean Superior Craton to its west to the negative anomaly that underlies the Appalachian Province to its southeast (Figure 5). Wave speed anomaly amplitudes vary between −1.5% near the Appalachian Front and the Great Lakes and ≤1% close to the Archean Superior boundary. Anisotropy amplitudes range between 0.2% and 1% and span a wide range of directions. Proterozoic provinces are ~2% slower than the cratonic mantle yet faster than the global average (Figure 5).

The Appalachian Province is underlain by consistently low isotropic phase velocities ($dC/C \approx -1.5\%$ to −3.5%) with respect to the regional average.

5. Discussion

5.1. The Heterogeneous Mantle Beneath Precambrian Terranes

5.1.1. Comparison With Previous Studies

Most global [e.g., Nettles and Dziewonski, 2008; Lebedev and van der Hilst, 2008; Ritsema et al., 2011; Debayle and Ricard, 2012; Schaeffer and Lebedev, 2013] and continental scale seismic tomography studies [e.g., van der Lee and Frederiksen, 2005; Bedle and van der Lee, 2009; Yuan et al., 2011; Schaeffer and Lebedev, 2014] generally image a wide high-velocity anomaly beneath Precambrian North America, similar to our model (Figures 5 and 6). The large-scale velocity decrease beneath the Proterozoic Grenville lithosphere in our model (Figures 5 and 6) is also present in some global models [e.g., Shapiro and Ritzwoller, 2002; Nettles and Dziewonski, 2008; Ritsema et al., 2011]. In contrast, some continental scale models [e.g., van der Lee and Frederiksen, 2005; Bedle and van der Lee, 2009; Yuan et al., 2011; Schaeffer and Lebedev, 2014] image uniformly fast wave speeds beneath both Archean and Proterozoic terranes. A number of regional models of southeast Canada [Chen and Li, 2012; Rondenay et al., 2000a; Akta and Eaton, 2006] reported a high-velocity anomaly corridor dipping away from the Superior Craton beneath the Grenville Province at 50–150 km depth.

Beneath the Eastern Churchill Province, the decrease in velocity in our model correlates with the models of Nettles and Dziewonski [2008], van der Lee and Frederiksen [2005], Bedle and van der Lee [2009], and Yuan et al. [2011], while high-velocity anomalies are detected by Schaeffer and Lebedev [2014], Lebedev and van der Hilst [2008], and Debayle and Ricard [2012]. Previous tomographic models that included anisotropy in the inversion [Nettles and Dziewonski, 2008; Yuan et al., 2011] correlate better with our model.

5.1.2. Seismic Heterogeneities Within the Superior Craton

The high seismic velocity generally imaged beneath cratonic regions is commonly thought to be caused by extensive chemical depletion in heavy elements such as Fe, Al, or Ca [e.g., Jordan, 1988; Griffin et al., 2003]. Purely compositional variations ranging from highly depleted dunite, generally associated with Archean mantle, to typical modern mantle peridotites result in 1–2% velocity variations [e.g., Goes et al., 2000; Lee, 2003; James et al., 2004]. A higher degree of depletion in heavy metals is generally associated with lower density and higher velocity [e.g., Griffin et al., 2009]. Some studies, however, suggest that the effects of melt depletion may be too subtle ($\delta V_s < 0.5\%$) to be resolved seismically [e.g., Afonso et al., 2008; Schutt and Lesher, 2006]. The phase velocity anomalies beneath the Superior Province range between 0.5% and 2.5% at shorter periods (20–90 s) and −1% to 2.5% at longer periods (≥150 s), indicating strong lateral heterogeneities within the cratonic mantle. Studies of xenoliths from the Renard kimberlite field in central Quebec showed a 200 km thick lithosphere composed mainly of herzolite and harzburgite [{Hunt} et al., 2012], although more depleted harzburgite may be present, as well as eclogitic portions [{Hunt} et al., 2010], which can result in increased seismic velocity and seismically distinct domains [e.g., Worthington et al., 2013]. Variations in chemical
composition of xenoliths, indicative of mantle heterogeneities, have been reported in other Precambrian regions [e.g., Siberian craton, Snyder et al., 1995, North China Craton, Tang et al., 2013, and Slave Craton, Kopylova and Caro, 2004].

A 100 K temperature anomaly may cause a velocity change of 0.5–2% [e.g., Hieronymus and Goes, 2010]. Within the stable continental interiors, large temperature variations are unlikely, however, since the thermal equilibration times (~2 Ga for lithospheric thickness of 200–270 km [Michaut et al., 2007]) are similar to the time elapsed since the last tectonic event (~2.5 Ga in the Superior Craton). The most recent tectonic event that affected parts of the southern Superior Craton is the passage of the North American plate over the Great Meteor hot spot track at ~190–110 Ma ago (Figure 1). The hot spot may have refertilized the mantle [e.g., Frederiksen et al., 2013; Boyce et al., 2016] in the southern section of the study region, consistent with the negative velocity anomalies we observe at all periods in our model (Figure 5) also detected in some regional [e.g., Frederiksen et al., 2007; Villenarie et al., 2012; Boyce et al., 2016] and continental scale models [e.g., French et al., 2013]. The thermal impact associated with a mantle plume head can be as high as ~500°C, which will rapidly diffuse and decay to ~80°C at 200 km over 120 Ma [Eaton and Frederiksen, 2007]. A major thermal contribution to the lateral velocity variations within the upper lithosphere is thus unlikely. Seismic heterogeneities at periods sampling deeper mantle material may result from local lithospheric thickness variations. At the same period and depth range of propagation, Rayleigh waves may encounter low-velocity asthenospheric material, due to the local topographic relief of the cratonic root, resulting in strong lateral phase velocity heterogeneities at longer periods (≥150 s).

5.1.3. Velocity Decrease Beneath Proterozoic and Younger Provinces
Within the Precambrian regions, the gradual decrease in velocity (≤3%) is a large-scale feature of our model that persists to ~140 s period and broadly correlates with the Proterozoic terranes that surround the Superior Craton. The decrease in velocity from Archean to Proterozoic mantle may reflect the transition from a melt-depleted cratonic lithosphere to a more fertile modern mantle beneath the Proterozoic. Interestingly, this gradual change in mantle seismic properties with age mirrors a change in crustal properties: the post-Archean crust was recently found to be thicker, more heterogeneous, and more compositionally evolved [Petrescu et al., 2016]. A recent body wave tomographic study of the upper mantle [Boyce et al., 2016] detected a sharp vertical change in P and S wave velocity beneath the Grenville Front (Figure 1), which they interpreted as a metasomatic front, likely related to the pre-Grenvillian subduction system altering the composition of the cratonic margin.

A thermal effect linked with lithospheric thickness may also contribute to the phase velocity decrease, although it cannot be unequivocally discriminated from compositional effects. The conductive geothermal gradient of a thick lithosphere, generally imaged beneath cratons, can be associated with a decrease in $V_s$ from 4.7 km/s in the upper lithosphere to 4.65 km/s at ~180 km depth, without considering anelastic attenuation [James et al., 2004]. Superior Craton lithospheric thickness was previously inferred to be >220 km [e.g., Yuan and Romanowicz, 2010; Schaeffer and Lebedev, 2014; Porritt et al., 2015], which would induce colder geotherms and consequently higher velocities. A thinner lithosphere that likely underlies the Proterozoic Grenville Province will have a steeper geothermal gradient, resulting in higher temperatures and progressively lower seismic velocities. The eastern continental margin of the Laurentian continent (present-day Grenville Province) is also thought to have been involved in a ~500 Ma long Andean-style paleosubduction system [Rivers, 1997; Rondenay et al., 2000a]; the Grenville Province was dominated by a hot long-duration plateau, similar to Tibet, that subsequently collapsed ~1 Ga ago [Rivers, 2009]. These tectonic events likely left an imprint on the cratonic mantle, altering its composition and lithospheric thickness and consequently redefining the geothermal gradient. We thus conclude that the joint thermochemical effect of both metasomatic refertilization [e.g., Boyce et al., 2016] of the Archean Laurentian mantle partially underlying the Grenville Province and a progressively thinner lithosphere can explain the ~3% velocity decrease we detect beneath the Proterozoic regions.

The ~5% velocity contrast between the Precambrian (i.e., the Superior, Eastern Churchill, and Grenville Provinces) and Phanerozoic terranes may arise from a thermal contrast: the Phanerozoic Appalachian Province may have experienced lithospheric delamination [e.g., Levin and Park, 2000; Yuan and Levin, 2014], which would have been accompanied by upwelling of hot asthenospheric material. Additionally, a shallower lithosphere underlain by hydrated asthenospheric melts beneath the Appalachians may also be contributing to the observed decrease in seismic velocity [e.g., Rychert et al., 2005].
5.2. Causes of Upper Mantle Anisotropy

The main cause of anisotropy in the upper mantle is usually associated with olivine, the most abundant and highly anisotropic mineral of the mantle. Strain results in development of crystallographic preferred orientation (CPO) of olivine crystals, leading to a axis alignment in the direction of flow [e.g., Zhang and Karato, 1995; Bystricky et al., 2000; Tommasi et al., 2000]. When a shear wave traverses such a medium, the faster-traveling component will travel parallel to the a axes of the crystal [e.g., Silver and Chan, 1991]. Anisotropic fabrics can form due to asthenospheric flow [e.g., Vinnik et al., 1989; Fouch et al., 2000; Sleep et al., 2002], the preferential alignment of fluid or melt [e.g., Blackman and Kendall, 1997], preexisting anisotropy frozen into the lithosphere [e.g., Silver and Chan, 1988; Vauchez and Nicolas, 1991; Bastow et al., 2007], or any combination thereof.

5.3. Frozen-in Archean Tectonic Deformation in the Upper Lithosphere

Whether or not tectonic processes in the Archean were capable of imparting a seismic anisotropic fabric on the lithosphere is debated [e.g., Fouch et al., 2004], but younger Precambrian fabrics are recognized worldwide. For example, the thick lithosphere of northern Canada retains a strong fossil lithospheric fabric of the ∼1.8 Ga Trans-Hudson Orogen [Bastow et al., 2011; Darbyshire et al., 2013]. Previous measurements of anisotropy are concentrated in the southern sections of our study area and come from magnetotelluric [e.g., Mareschal et al., 1995] and SKS measurements [e.g., Darbyshire et al., 2015; Gilligan et al., 2016], which have low vertical resolution and may often record a vertically averaged signal from the lithosphere and the convecting mantle.

At periods <110 s (peak sensitivity at ∼120 km depth), anisotropy patterns within the Superior Craton transition from predominantly E-W in the south to N-S in the north (Figure 5), mimicking the orientation of the Archean surface boundaries (Figure 8). SKS splitting directions in the southern Superior are also parallel to the tectonic fabrics [Sénéchal et al., 1996], although Darbyshire et al. [2015] found this correlation to be partial. Since the peak depth sensitivity of the maximum period at which this anisotropic signal is detected is at ∼120 km depth, the source must reside in the lithosphere. Short scale lateral variation of SKS measurements also suggests that the origin of the anisotropic signal lies mostly in the 50–150 km depth range [Sénéchal et al., 1996].

The striking agreement in direction between surface geological features, our calculated anisotropy in the 40–110 s period range (Figure 5), and previous SKS measurements (Figure 8) suggests that the anisotropic signal must be due to deeply rooted lithospheric deformation zones. This idea is also supported by a systematic obliquity between electric and seismic anisotropy identified in the southern Superior Craton [Mareschal et al., 1995; Sénéchal et al., 1996], thought to arise from the alignment of a-type olivine minerals in noncoaxial shear deformation zones in the lithosphere [Ji et al., 1996]. The surface geology indicates that tectonic assembly of the Superior Craton was completed by 2.5 Ga [Hoffman, 1989], implying that the detected lithospheric deformation is of Archean age. The persistence of fossilized Archean tectonic deformation at mantle depths...
imply that the crust and lithospheric mantle deformed coherently and that the constituent microcratonic blocks must have been already thick (≤120 km) and strong enough to transmit stress across lithospheric depth scales, at the time of Archean assemblage.

### 5.4. Archean Mantle Beneath Surrounding Terranes

Phase velocity anomalies of ~1%, indicative of cratonic material, partially extend beneath the Grenville Province. Anisotropic fabrics are oblique to the general NE-SW Grenville orogenic trend and parallel Archean deformation zones in the Superior. This observation is most evident at periods ≤70 s, between 46° and 52°N (Figure 5), where the dominant E-W anisotropy direction persists up to 200 km away from the Archean-Proterozoic boundary. In the same area, SKS splitting measurements at a seismic array crossing the Archean-Proterozoic boundary showed no variation in the fast axis direction [Sénéchal et al., 1996; Ji et al., 1996]. Active seismic profiling of the Archean to Proterozoic crust detected Archean lower crust underlining the Grenville Province as far as ~250 km away from the Archean-Proterozoic boundary [White et al., 2000]. Magnetotelluric studies detected dipping resistors to 120 km depth beneath the Grenville Province, interpreted as refertilized Archean mantle [e.g., Adetunji et al., 2014]. On the other hand, a recent body wave tomographic study of the upper mantle [Boyce et al., 2016] detected a sharp vertical change in P and S wave velocity beneath the Grenville Front (Figure 1). The joint thermochemical effect of metasomatism or a progressively thinner and warmer lithosphere, as discussed previously, may influence the slight reduction in seismic velocity beneath the circumcratonic regions. However, the persistence of anisotropic fabrics that parallel Archean surface deformation suggests that cratonic mantle extends beneath surrounding regions and maintains frozen-in Archean fabrics despite possible metasomatic reworking.

### 5.5. Anisotropy Decrease in the Lower Cratonic Lithosphere

We believe that the nature of the weakly anisotropic zone sampled in the 110–140 s period interval (peak sensitivity at 150–200 km depth) is lithospheric. Previous seismic [e.g., Yuan and Romanowicz, 2010; Schaeffer and Lebedev, 2014] and geochemical studies [Hunt et al., 2012] inferred the lithospheric thickness in the area to be >220 km which is ~50 km deeper than the peak sensitivity at 140 s (Figure 6). Also, phase velocities are ~1% faster than the regional (~4.3 km/s, Figure 6) and the global average (~4.2 km/s, Figure 7d), indicating that the material is either colder or more compositionally depleted, akin to cratonic regions. Competing effects of multiple anisotropic layers with perpendicular fast directions may cause an almost null cumulative anisotrophic signature, which may be observed in SKS studies [e.g., Girardin and Farra, 1998]. Similarly, the relatively broad range of depth sensitivity of Rayleigh waves at periods between 110 and 140 s (Figure 6) may include weak signals from both the upper lithosphere and the underlying mantle, with contrasting anisotropic signatures, resulting in an apparently isotropic transition zone. However, if this was true, the observed splitting of SKS waves, which traveled through both anisotropic layers with hypothetically perpendicular anisotropy directions, should be almost null, contrary to previous measurements in the area [Darbyshire et al., 2015; Sénéchal et al., 1996]. Also, the plate motion vectors (Figure 8), commonly considered to represent upper mantle flow [e.g., Yuan and Romanowicz, 2010], are oblique to the anisotropy directions estimated in the upper lithosphere (20–140 s, Figure 5). We thus consider it unlikely that the interference between the anisotropic signatures of the upper lithosphere and the underlying flowing asthenosphere is sufficient to cause the observed drop in the absolute anisotropy magnitude in a period interval sampling a ~100 km depth range.

#### 5.5.1. Hypotheses of Formation and Age of the Lower Lithosphere

Theories explaining the reduction in anisotropy within the upper mantle include a change in the creeping mechanism from dislocation to diffusion creep [Karato, 1992] or the olivine glide system [Mainprice et al., 2005], both of which would imply a younger formation age for the lower lithosphere than for the upper lithosphere. Forward modeling of seismic anisotropy generated by olivine CPO developed in the presence of mantle flow suggests that a sudden switch in flow direction generates a depth-averaged response of the mineralogical CPO development, resulting in a weakly anisotropic basal layer at the transition between different strain orientations [Tommasi, 1998]. However, in a global study of seismic anisotropy, Debayle and Ricard [2013] found that for plate velocities of <4 cm/yr (current absolute plate motion (APM) magnitude in eastern Canada is ~20 mm/yr), the basal drag from plate-asthenosphere interaction is likely too weak to develop flow-aligned anisotropy. The absence of significant seismic anisotropy in the lower lithosphere may therefore suggest a low relative velocity between the growing Superior Craton lithosphere and the convecting mantle beneath during Precambrian times. The formation of the lower lithosphere may have been sufficiently protracted that no single, coherent mantle flow or basal drag anisotropic fabric was preserved within it.
Alternative explanations for the presence of weakly anisotropic layers include metasomatism by water-rich fluids or melts [e.g., Baptiste et al., 2012]. Annealed microstructures of coarse-grained peridotites extracted from ~150 km depth from the Kaapvaal Craton also indicate that dislocation creep, recording early Archean deformation, was followed by effective annealing, due to a protracted period of cooling and metasomatism [Baptiste et al., 2012]. Melt percolation derived from a volatile-rich subcontinental mantle may have reworked a preexisting layer from beneath, possibly erasing its original seismic signature. However, previous studies of mantle seismic anisotropy in Spain [Vauchez et al., 2005] and Polynesia [Tommasi et al., 2004] suggested that fossil tectonic fabrics, in the form of preexisting olivine crystal preferred orientations, inherited from older deformation, can survive partial melting or melt percolation. Baptiste et al. [2012] also speculated that metasomatism would occur episodically, which would induce lateral heterogeneities within affected areas. The Superior Craton is pervaded by multiple dyke swarms that record magmatic pulses at 2.5, 2.2, and 2.1 Ga, linked with episodes of extension at the paleocratonic margins [Ernst and Bleeker, 2010] or the arrival of mantle plume heads [Ernst and Bleeker, 2001], providing possible metasomatic sources. Geochemical data in the Slave Craton [Griffin et al., 2004] suggest the presence of a second uniform Archean mantle layer, with the top located between 140 and 160 km depth. Griffin et al. [2004] interpreted this layer as subcreted mantle plume material at ca. 3.2 Ga. However, the radial spread of plume material from one or more sources or the presence of pervasive dyke swarms in the lower lithosphere may be expected to cause discernable seismic anisotropy patterns [Yuan and Romanowicz, 2010; Snyder and Bruneton, 2007] and would decrease seismic velocities relative to normal lithospheric mantle [Tommasi et al., 2004], contrary to our results.

We speculate the timing of formation of the lower lithosphere to be Proterozoic, although our data alone cannot constrain this. While most geochemical evidence indicates that most melt extraction occurred in the Archean [Carlson et al., 2005; Pearson, 1999], some xenoliths record Proterozoic mantle addition to the cratonic root [e.g., Pearson et al., 1995; Smart et al., 2014], supporting the idea that root formation and stabilization may be a younger phenomenon than the bulk depleted craton age. The wide extent of the almost isotropic and relatively homogenous lower lithosphere likely fits the hypothesis in which the base of the lithosphere grew by progressive cooling after the tectonic assemblage of the upper lithosphere was completed [e.g., Bédard and Harris, 2014]. The recognition of at least two temporally distinct lithospheric zones in our model endorses the idea that cratonic keel growth is not due to one particular dominant tectonic process and, most importantly, is not exclusively an Archean phenomenon. Previous seismic studies postulated a Paleoproterozoic origin for the lower layer beneath Hudson Bay and surrounding terranes [e.g., Darbyshire et al., 2013; Porritt et al., 2015], supporting episodic formation of cratonic roots in the Precambrian.

5.6. Lowermost Lithosphere — Transition to the Convective Asthenosphere

While we do not attempt to constrain the lithosphere-asthenosphere boundary due to the broad sensitivity range of Rayleigh waves at longer periods, we note that pronounced seismic heterogeneities develop at periods ≥140 s and the anisotropy magnitude significantly increases at 190 s (~4%, Figure 6). Geodynamic models predict the existence of a thin transitional layer (≤50 km thick) at the base of the chemically depleted cratonic core, which could absorb some of the shear deformation [e.g., Cooper et al., 2004]. The heterogeneous anisotropic material sampled between 140 and 190 s may be related to the geodynamically predicted transitional layer at the base of the keel.

At periods ≥190 s, with peak depth sensitivity at ~270 km depth (Figure 6) high phase velocity anisotropy and heterogeneity develops. The data fit at these periods decreases by ~30% compared to shorter periods (Figure C1), which calls for caution when interpreting these features. Since we independently invert for phase velocity anomalies at each period, the consistency of large-scale features between periods suggests that these features are not random numerical artifacts but may indicate real material properties. Also, Figure 2 shows that velocity dispersion is clearly visible at frequencies as low as ~0.003 Hz (~333 s period). In accordance with our results, strong anisotropy is expected in the 250–400 km depth range beneath cratons, due to asthenospheric flow [Gung et al., 2003]. Mantle convection and locally complex flow around the cratonic root may thus be responsible for the anisotropy increase and locally strong heterogeneities at periods ≥190 s (Figure 6), corroborating the view that subcratonic mantle flow is complex [e.g., Fouch et al., 2000].

5.7. Comparison With Other Cratons

Seismic anisotropy has been studied in other cratons using a range of methods such as SKS splitting [e.g., Silver et al., 2001; Snyder and Bruneton, 2007], Love and Rayleigh wave anisotropy [e.g., Adam and Lebedev, 2012; Yuan et al., 2014], and P and S receiver functions [e.g., Abt et al., 2010; Miller and Eaton, 2010; Sodoudi et al., 2013], each reporting slightly different results and interpretations, depending on the imaging method and
resolution. A comparison of the seismic anisotropy between different stable continental interiors yields a few similarities (see, e.g., Fouch and Rondenay [2006] for a review of anisotropy beneath cratons), although a unified anisotropic model of craton structure and growth seems increasingly unlikely. Some cratons are found to be anisotropic from top to bottom and may retain the signature of underplated lithosphere [e.g., Slave, Bostock et al., 2010], while others may have been affected by nearby plume magmatism [e.g., Tanzania, Sebai et al., 2006] or other types of thermochemical erosion (e.g., delamination beneath the North China Craton [Xu, 2001; Gao et al., 2009]), which erased their initial anisotropic signatures or contributed to their growth or destruction in different ways. Most studies, however, agree that on average, anisotropy measurements display a correlation with surface tectonic trends within the upper ~150 km of the lithosphere, related to initial craton amalgamation, similar to our findings (Figure 8): the Baltic shield [e.g., Pedersen, 2006], the South African cratons [e.g., Silver et al., 2004; Fouch et al., 2004; Adam and Lebedev, 2012; Freibourger et al., 2001], western Australian cratons [e.g., Girardin and Farra, 1998; Fichtner et al., 2010], the North China Craton [e.g., Bai et al., 2010], and the Slave Craton [e.g., Snyder et al., 2003].

Most studies report a switch from dominant fossil anisotropy in the upper lithosphere to anisotropy related to present-day asthenospheric flow [e.g., Deschamps et al., 2008]. The transition from layers containing frozen-in Archean deformation signatures to subcratonic flow may, however, be more complex. Mechanically weak basal layers such as deformable boundary layers that absorb some of the shear stress at the base of the keel have been geodynamically inferred [e.g., Cooper et al., 2004] and reported beneath some cratons [e.g., Kaban et al., 2015]. Also, lithospheric regions exhibiting a reduction in their azimuthal anisotropy similar to our observations (Figure 7) were previously reported in the 100–200 km depth range in a 3-D anisotropic model of North America [Yuan et al., 2011] and between 70 and 150 km depth in Australia from full waveform tomography [Fichtner et al., 2010]. Forward modeling of 5 receiver functions in the Kaapvaal Craton revealed an anisotropic layer (3% anisotropy) between 40 and 90 km depth, underlain by a dominantly isotropic metamasomatized lower lithosphere [Sodoudi et al., 2013]. Thus, depending on the tectonic history of each craton, cratonic roots may comprise chemical and metamasomatic layers [e.g., Sodoudi et al., 2013], sometimes separated by intralithospheric discontinuities [e.g., Wirth and Long, 2014].

6. Conclusion

To investigate the structure of the lithosphere beneath previously unstudied parts of the Canadian Shield, we analyzed surface wave dispersion data from teleseismic earthquakes. We inverted Rayleigh wave dispersion data for isotropic phase velocity heterogeneity and azimuthal anisotropy variations at different frequencies. A first-order feature of our model is the clear distinction in the isotropic phase velocity between the fast upper mantle underlying Archean terranes and that beneath Proterozoic belts, between 20 and 140 s. We interpret this contrast as a compositional change between relict depleted cratonic lithosphere and metasomatized Archean mantle extended beneath Proterozoic lithosphere.

The Superior Craton lithosphere is anisotropically complex. The upper lithosphere, sampled at periods of 40–100 s, is seismically fast and weakly heterogeneous, with anisotropic fabrics that parallel surface tectonic boundaries of Archean age. This correlation suggests that cratonic blocks must have been strong enough to support plate-scale collision at ~2.5 Ga. Cratonic upper lithosphere partially underlies the surrounding Proterozoic belts and preserves the signature of Archean deformation despite subsequent mantle refertilization. The lower lithosphere, sampled at periods of 110–140 s, is seismically fast, more homogenous, and weakly anisotropic. We speculate that it formed in Proterozoic times, subsequent to the accretion of the overlying Archean terranes. The weak anisotropy detected at lower lithospheric depths could result from the interference between present-day anisotropy in the flowing asthenosphere below and the fossil anisotropy in the upper lithosphere or the gradual transition between different strain orientations as a result of changing convection patterns at the time of lithospheric growth. Nevertheless, we find it more likely that the decrease represents the weakly anisotropic fabric of the material, which may point toward regionally sluggish mantle convection and/or slow absolute plate motion at the time of formation. Finally, the root of the cratonic mantle may be imaged between periods of 140 and 190 s and contains highly heterogeneous and anisotropic seismic fabrics. Significant topography may be present at the base of the craton, and the transition to the flowing mantle is therefore likely to be complex. Eastern Canada’s heterogeneous and anisotropically complex lithosphere documents the development of cratonic keels through Archean and Proterozoic times.
Appendix A: Measurement Bias in the Two-Station Method

Individual dispersion curves measured between pairs of seismic stations describe the average medium properties between the two stations, which can be affected by the presence of strong lateral heterogeneity. Measuring the dispersion curves from multiple events from opposite directions enables the detection of such heterogeneities and improves the measurement quality [Legendre et al., 2014]. Laterally heterogeneous media also generate seismic wave diffraction, introducing nonplane wave energy. Surface wave diffraction arises from an inherently complex relationship between waves and heterogeneities [Pedersen, 2006] and is not taken into account by the method used here. This phenomenon is manifested as irregularities in the dispersion curves (Figure 2). We used 10–250 measurements for each station pair (Figure 1) with seismic waves coming from opposite directions (Figure 2). These were sufficient to identify and exclude outlier dispersion curves, reducing the effects of heterogeneity and diffraction [Meier et al., 2004; Lebedev et al., 2006], and resulting in smooth dispersion curves (Figure S2 in the supporting information).

Bias may arise from angular deviations of the surface wave propagation paths from their predicted great-circle arc [Pedersen, 2006], with deviations of $\leq 20^\circ$ possible [e.g., Cotte et al., 2000; Foster et al., 2014]. At short periods ($\sim 20$ s) great-circle arc deviations may be as high as $15^\circ$ [Pedersen et al., 2015]. Foster et al. [2014] showed that the standard deviation of the distribution of great-circle arc deviations decreases from 0.4 at short periods ($\sim 20$ s) to less than 0.1 at longer periods ($\sim 100$ s). Pedersen [2006] estimated that phase velocities can be $\leq 0.5\%$ different from the structural velocities for a deviation of $10^\circ$ and $2-2.5\%$ for a deviation of $20^\circ$. Meier et al. [2004] calculated that a deviation of $15^\circ$ amounts to a phase velocity increase by a factor of 1.035. To reduce the effect of great-circle path deviations, we imposed a $\pm 5^\circ$ limit between the station pair azimuth and the earthquake azimuth, which should introduce an error $<0.5\%$ [Pedersen, 2006]. Also, averaging over hundreds of dispersion curves obtained from earthquake waves coming from opposite directions should significantly reduce the impact on phase velocity from wrongly predicted arrival angles for a few events.

Lateral scattering strongly influences fundamental mode Rayleigh waves at periods $\leq 16$ s [Meier et al., 2004]. The JWKB approximation we use is insufficient to describe the complex wave propagation at such short periods. This is reflected in the incoherent pattern observed in the spectrograms above $\sim 0.07$ Hz (Figure 2). For this reason we do not use short-period data.

Appendix B: Inversion Setup

For each earthquake ray path connecting pairs of seismic stations, equation (2) is discretized. A linear system of equations is thus built and independently solved for at each period [Deschamps et al., 2008]:

$$d = Gm.$$  
(B1a)

$$d^T = (\delta C_1 \ldots \delta C_N),$$  
(B1b)

$$m^T = (\delta C_{iso,1} \ldots \delta C_{iso,M} A_{2\psi,1} \ldots A_{2\psi,M} B_{2\psi,1} \ldots B_{2\psi,M} A_{4\psi,1} \ldots A_{4\psi,M} B_{4\psi,1} \ldots B_{4\psi,M}),$$  
(B1c)

where $N$ is the number of paths obtained at each period and $M$ is the number of grid knots.

The $A$ and $B$ terms are then combined to estimate the amplitude ($\Lambda$) and the angle of the fast $S$ wave polarization ($\Theta$), for both $2\psi$ and $4\psi$ contributions. These are useful to visualize the anisotropy patterns across the model (Figures 5, 6).

$$\Lambda = \sqrt{A^2 + B^2},$$  
(B2a)

$$\Theta = \frac{1}{2} \arctan \left( \frac{B}{A} \right).$$  
(B2b)

Appendix C: The Importance of Anisotropy

Previous studies noted the importance of accounting for anisotropy in inverting and interpreting seismic data [e.g., Hess, 1964; Forsyth, 1973; Deschamps et al., 2008; Darbyshire and Lebedev, 2009]. To assess how much the solution deteriorates if we neglect anisotropy, we calculated the variance reduction of the inverted models at all periods with (1) a purely isotropic inversion, (2) an inversion that includes isotropic heterogeneity and $2\psi$ anisotropic variation, and (3) the full inversion which allows for isotropic variations as well as $2\psi$ and $4\psi$ anisotropy. Figure C1 shows that the data fit decreases up to $25\%$ when a purely isotropic solution is forced...
on the inversion versus a full solution. The phase velocity patterns are qualitatively similar if we only invert for the isotropic component (Figure S6 in the supporting information). However, the magnitude of the isotropic $V_s$ anomaly estimated by an isotropic-only inversion may increase or decrease up to 2% in areas where the anisotropy estimated by the full inversion is $\geq 2\%$ or at the model edges, where path coverage and crosscutting is poor. The velocity heterogeneity can thus be overestimated, leading to biased interpretations.

Anisotropic olivine crystals in the mantle are thought to have only a small $4\psi$ signal [Montagner and Natof, 1986], but the silica-rich nature of the continental crust [e.g., Christensen, 1996] means that it may have different anisotropic properties, rendering both the $2\psi$ and $4\psi$ anisotropy terms important [e.g., Polat et al., 2012]. Calculations of anisotropy for mineralogical and petrological upper mantle models indicate that Rayleigh waves have weak $4\psi$ anisotropy dependence, as opposed to Love waves [Montagner and Natof, 1986; Montagner and Anderson, 1989]. Spatial coherence between $2\psi$ and $4\psi$ terms in a study of the Hellenic subduction zone found weak directional changes in the $2\psi$ terms, if the $4\psi$ component is not accounted for in the inversion [Endrun et al., 2011]. They also found a correlation between higher $4\psi$ anomalies and regions with poorer ray path coverage. Our results, however, do not show such a correlation (see the supporting information), indicating that $4\psi$ terms are not simply an indication of fitted noise or a numerical strategy to improve data fit in less resolved model regions. The roughness of this component, defined as the average difference between the value at each grid knot and the nearest neighbor values, is also found to increase at longer periods (Figure C1b). Interpreting the $4\psi$ anisotropy from a physical point of view is not straightforward: we speculate tentatively that it may be a combination of Rayleigh-Love wave coupling effects or more complex mineralogical anisotropy systems developed locally. We therefore follow the practice of previous studies that calculate both $2\psi$ and $4\psi$ terms [e.g., Deschamps et al., 2008; Darbyshire and Lebedev, 2009; Fry et al., 2010; Polat et al., 2012] but focus our interpretation solely on the isotropic and the $2\psi$ anomalies, in the context of upper mantle structure. 

**References**


