An Alternative EARTH

Warren B. Hamilton, *Department of Geophysics, Colorado School of Mines, Golden, Colorado 80401, USA, whamilto@mines.edu*

ABSTRACT

The standard Earth of geodynamics and geochemistry is rationalized from assumptions that the mantle is compositionally inverted—still-unfractionated lower mantle beneath volatile-depleted upper mantle—and that material circulates easily from bottom to top. Multidisciplinary data better fit a less-volatile and less-radioactive planet wherein depleted lower mantle, fractionated early and irreversibly, is decoupled from upper mantle plus crust that evolve and circulate separately. Early Archean fractionation produced global(?) felsic crust and refractory upper mantle. Later Archean granite-and-greenstone upper crust formed atop this ancient crust, which remained hot and weak; distinct continents and oceans did not exist, and upper mantle was much hotter than now. Plate tectonics began ca. 2.0 Ga when continents could stand above oceans and oceanic lithosphere could cool to subduction-enabling density and thickness. Upper mantle has since become more fertile and new increments of continental crust more mafic as continental crust has been progressively diminished by recycling into cooling mantle. Plate circulation is driven by subduction, which is enabled by density inversion produced by seawater cooling from the top of oceanic lithosphere, is self-organized, and is confined to upper mantle. The matching rates of hinge rollback and of advance of fronts of overriding plates are keys to dynamics. Slabs sinking broadside from retreating hinges drive both subducting and overriding plates and force seafloor spreading in both shrinking and expanding oceans. An Antarctica-fixed framework depicts prediction-confirming "absolute" plate motions that make kinematic sense, whereas hotspot and no-net-rotation frames do not. Plumes from deep mantle, subduction into deep mantle, and bottom-up convective drive do not exist.

INTRODUCTION

The conventional model (e.g., Turcotte and Schubert, 2002) of Earth's evolution and dynamics postulates that most of the mantle is little fractionated, major differentiation continues, and continental crust has grown progressively throughout geologic time; through-the-mantle convection operates, lithosphere plates are moved by bottom-driven currents, and plumes rise from basal mantle to surface; and plate tectonics operated in early Precambrian time. All of these conjectures likely are false. They descend from speculation by Urey (1951) and other pioneers, reasonable then but not now, that Earth accreted slowly and at low temperature from fertile chondritic and carbonaceous-chondritic materials, heated gradually by radioactive decay and core segregation, and is still fractionating.

The notion of a cold, volatile-rich, young planet has long since been disproved, but its corollary of unfractionated lower mantle survives as the basic assumption from which the standard model is rationalized. This assumption and its derivatives are incompatible with major constraints. This essay summarizes, and continues and diverges from, a long paper (Hamilton, 2002) wherein many topics only noted here were discussed and documented. My sparse citations here are mostly of representative current papers from among hundreds of sources incorporated.

EVOLVING EARTH

Uniform tectonics should not be assumed over the 4.4 b.y. span recorded by dated Earth minerals. Crust and mantle processes changed greatly as Earth cooled by loss of original heat and by decreasing production of radiogenic heat.

Accretion and Fractionation

Earth accreted hot, fast, and violently. Distribution of ^{182}Hf (half-life 9 m.y.) and its daughter ^{182}W indicates that Earth likely reached two-thirds of its present size within ~10 m.y. of first condensation of solid grains in the solar nebula, ca. 4.567 Ga (Jacobsen, 2003). Orbital considerations require that accretion was dominated by large bolides formed near Earth's distance from the Sun but that much material came from elsewhere sunward of Jupiter; the bulk of atmospheric, hydrospheric, and carbon components may have arrived late, mostly from the outer part of the asteroid belt (e.g., Alexander et al., 2001; Wetherill, 1994). Pre-aggregation material was zoned heliocentrically by evaporation—and condensation beyond "snow lines"—of volatile elements and compounds. This is shown by irregular zonation of asteroids linked to meteorite types by reflectance spectroscopy. Outer asteroids are dominated by ices and organic compounds (no meteorites), medial ones by organic materials and variably hydrated silicates (carbonaceous chondrites), and inner ones by silicates and metal (most meteorites, including chondrites; Alexander et al., 2001; Britt and Lebofsky, 1999). The heliocentric trend obviously continues through Mars and Earth, but to what extent? Modern calculations (e.g., Allègre et al., 2001) of Earth's composition accept volatile fractionation relative to meteorites but assume a relatively low-temperature cutoff and a still-unfractionated lower mantle and may much overestimate Earth's volatile elements.

Accreting Earth was repeatedly partly melted and devolatilized by great impacts, and the likely Moon-forming mega-impact (before 4.45 Ga) would have wholly melted and further devolatilized the protoplanet (Zhang, 2002). Irreversible final fractionation of deep Earth mantle is made likely by intrinsic density, very low thermal expansivity, and very high melting temperature of perovskite, the probable dominant mineral, at lower-mantle pressure (Anderson, 2002a). Barriers, uncrossable by material and hence layering convection, may occur near depths of 2000 km (Badro et al., 2003) and 1000 km (Anderson, 2002a) in the lower mantle and 660 km at its top, but I discuss only

the last of these because upper-mantle circulation likely is closed above it.

A bulk Earth low in K, hence volatile elements, and fractionation of K and the incompatible elements U and Th into upper mantle and crust (and hence layered circulation), are indicated also by thermodynamics. Earth's heat loss, now largely of radiogenic heat, is much overstated in the standard model. The conventional value, ~44 TW, is derived by substituting for measured heat flows on young oceanic crust vastly higher values calculated with the assumptions that heat flow must fit a square-root-of-age function and that oceanic lithosphere has constant conductivity. As conductivity in fact decreases greatly with temperature, actual heat loss is ~31 TW, close to the measured value, and thermodynamic and mineral-physics data require that nearly all radioactivity be above 660 km (Hofmeister and Criss, 2003).

Evolving Crust and Lithospheric Mantle

Continental crust became more depleted in mobile elements with decreasing age, whereas continental lithospheric mantle, and likely the entire upper mantle, became more enriched in them. These observed trends are opposite to those required by the standard model. Archean crust commonly is more felsic in bulk composition than younger crust and lacks the latter's typical thick underplated norite and mafic granulite (e.g., Durrheim and Mooney, 1994). Both xenoliths (Griffin et al., 2003b) and the inverse correlation of crustal age with mantle contribution to continental heatflow (Artemieva and Mooney, 2001) prove the opposite trend in continental lithospheric mantle, which becomes vastly richer in low-melting compounds with decreasing crustal age. Oceanic upper mantle is less well constrained but resembles Phanerozoic, not Archean, subcontinental lithosphere (Griffin et al., 2003b). Magnesian Archean subcontinental peridotite is lower in density, even though much colder, than asthenosphere and rides buoyantly with cratons.

Upper mantle has cooled during geologic time. Archean lithospheric peridotite has retrograde-metamorphic mineralogy and metasomatism of sampling-kimberlite ages, but texture and bulk composition of some of it fits initial (Paleoarchean?) crystallization from liquidus melt ~400 °C hotter than present near-solidus asthenosphere (Herzberg, 1993). (I use the nonstandardized term Paleoarchean to mean 4.4–3.5 Ga; Mesoarchean, 3.5–3.0 Ga; and Neoarchean, 3.0–2.5 Ga.) Ultramafic lava (komatiite), whose liquidus composition indicates a mantle residuum of olivine + orthopyroxene, is common in Mesoarchean and Neoarchean volcanic sections and requires upper mantle ~200 °C hotter than modern asthenosphere.

If the 660 km discontinuity is uncrossable by material, as I argue subsequently, then the inverse secular compositional trends of crust and upper mantle reflect thorough initial separation followed by progressive mixing back into upper mantle of most early crust and of sediments derived from it, and successively less complete separation of new crust from that increasingly enriched mantle. Volume of continental crust has thus decreased greatly through time. The common contrary claim that crust has grown through time derives mostly from isotopic calculations based on false assumptions of constant compositions of crust and mantle.

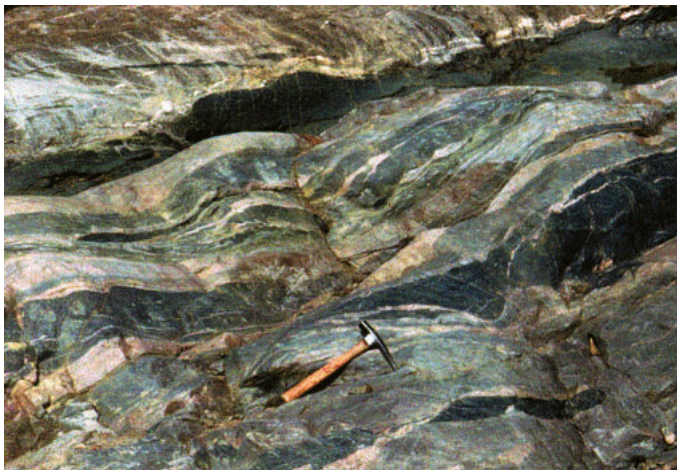

Figure 1. Paleoarchean heterolithic and polycyclic migmatite (tonalite, granodiorite, pegmatite, and mafic rocks), Acasta Lake, Slave craton, northwest Canada. Nearly concordant ion-microprobe U-Pb ages, by S.A. Bowring, of zircons from this outcrop scatter from 3.96 to 3.5 Ga. Only felsic gneisses are known as basement beneath Archean supracrustal rocks, including ultramafic volcanics.

Paleoarchean Continental Crust

Archean mafic and ultramafic volcanic rocks are known to overlie only ancient felsic basement. Pre–3.5 Ga rocks in cratons are polycyclic felsic migmatites and gneisses dominated by hydrous (biotite > hornblende) tonalite, trondhjemite, and granodiorite (TTG) and containing abundant small to huge enclaves of ultramafic, mafic, and anorthositic rocks (Fig. 1; Bleeker, 2002). Concordant ion-probe U-Pb age determinations of cores of zircons in these gneisses reach 4.2 Ga (and of clastic zircons in derivative quartzite, 4.4 Ga) in SW Australia, 4.1 Ga in NW Canada, 4.0 Ga in Wyoming, 3.9 Ga in central Canada, Labrador, and SW Greenland, 3.8 Ga in China, 3.7 Ga in NW Australia and South Africa, and 3.6 Ga in West Africa, as reported by various authors. Most zircon ages in these migmatites are, however, much younger, with wide variations between nearby samples, and typically scatter along or close below U-Pb concordia down toward or to the 3.3–2.6 Ga ages of intrusion of granites into supracrustal sequences. The ancient gneisses record a billion years of repeated partial melt or at least of temperatures above that required for formation of new zircon, and remobilized melts must have been hydrous and low-temperature where inherited zircons are preserved (cf. Miller et al., 2003). Isotopes of Pb (Kamber et al., 2003), Hf (Scherer et al., 2001), and Nd in the gneisses accord with separation from mantle of protoliths well before 4 Ga, followed by complex recycling, rather than with prolonged incremental growth from the mantle.

Earth's felsic components may have been mostly fractionated by 4.4 Ga into "continental" crust, most of which has since been returned to upper mantle. Hydrous TTG melt can form at low pressure by fractionation of hydrous basaltic melt, by partial melting of hydrated mafic metamorphic rocks, or by recycling. No Archean mantle or basal crust is exposed but seismic velocities suggest that felsic and intermediate migma-

Figure 2. Archean fuchsitic quartzite, Lake Winnipeg, Superior craton, Manitoba, part of thin section deposited on basement gneiss (zircons ca. 3.0 Ga) and overlain by thick pillow basalt (ca. 2.7 Ga). Mapping and dates by J.A. Percival and associates. Coin diameter: 3 cm.

Figure 3. Landsat satellite image of Mesoarchean granite-and-greenstone assemblage, central Pilbara craton, NW Australia. Domiform crustal-melt granitic batholiths (light; 3.5–2.9 Ga zircons) rose during and after deposition of 3.5–3.2 Ga ultramafic, mafic, and felsic volcanic rocks and sediments (multicolor stripes) preserved in open to tight and complex synclines. Another greenstone cycle, Neoarchean mafic and felsic volcanic rocks and iron formation, ca. 2.8–2.6 Ga (brown), overlies older units in S and in small syncline in N. Geology mapped by Geological Survey of Western Australia geologists. Area 175 km wide. Image supplied by Australian Centre for Remote Sensing, Australian Surveying and Land Information Group. Colors assigned to spectral bands: blue to visible green, green to near-IR, red to mid-IR.

tites extend to the base of the crust, so major complementary materials are now in the mantle and a dominant origin from metavolcanic rocks is precluded. Direct fractionation, followed by recycling, is inferred. Paleoarchean felsic crust may have been kept hot from below, from within by concentrated radioactivity, and from above by dense greenhouse atmosphere. Crust was too hot and weak to stand high as continents, its tracts moved differentially atop rapidly convecting upper mantle, and it flowed internally.

Regional temporal and isotopic variations in ancient gneisses (Bleeker, 2002; Griffin et al., 2003a) may relate to both impacts and upper-mantle motions. Huge bolides hit the Moon as late as 3.9 Ga, which more likely dates the tail end of main accretion than a late bombardment. Lunar impact history transfers with increased intensity to Earth, where bolides produced great transient excavations into upper mantle, forming giant magma lakes by impact and decompression melting, and redistributed huge tracts of crust. Paleoarchean anorthosites are low-pressure calcic fractionates, now churned into ancient migmatites, and perhaps they and the voluminous ultramafic and mafic enclaves and unidentified hot, dry protoliths for some of the ancient gneisses differentiated in impact lakes.

Mesoarchean and Neoarchean Tectonics

Volcanic and sedimentary "greenstone belt" assemblages were deposited after ca. 3.5 Ga (possibly 3.8 Ga) in some cratons-to-be and ca. 3.0 Ga in others on TTG basement. Only felsic rocks are known from outcrop or seismology to underlie mafic volcanic rocks, and no supracrustal rocks are proved ensimatic or deep-ocean. Where depositional contacts with basement are preserved, strata commonly begin, not with mafic volcanic rocks, but with thin micaceous quartzite (Fig. 2), derived from basement that was above sea level nearby, and/or with thin chert and banded iron formation (Bleeker, 2002). Above this, lower parts of supracrustal successions are typified by plains of kilometers-thick mafic and subordinate ultramafic submarine lavas, and upper parts by felsic volcanic rocks, often intercalated in mafic rocks, or by repeated cycles of the contrasted types. Basalt associations poorly match modern oceanic basalts in major- and trace-element compositions, and much Archean basalt records voluminous assimilation of felsic basement in more-mafic melt (Green et al., 2000). Early supracrustals are shown to be subregional stratal sheets where concordant ion-probe zircon U/Pb ages are tied to mapped units (e.g., Ayer et al., 2002). Single thick and rapidly erupted ultramafic flow complexes are exposed with strike lengths of hundreds of kilometers and may have reached volumes of 10,000 km^3 (Hill et al., 2001).

Diapiric batholiths rose concurrent with deposition, fed felsic volcanics and porphyries, and formed domiform granites

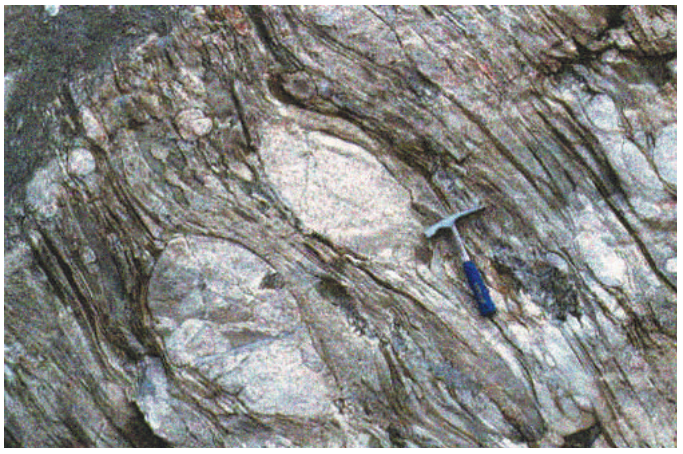

Figure 4. Neoarchean conglomerate. Cobbles of leucogranite and pebbles of felsite, eroded from diapiric granite, with downdip stretching due to continuing rise of granite and sinking of greenstone syncline. Vertical surface near Wawa, Superior craton, Ontario.

separated by synclines of contact-metamorphosed and dip-dragged supracrustals (Figs. 3, 4). Little deformation preceded diapirism, and strain and metamorphism are low, distant from batholiths. Greenstone belts are networks of synclines, mostly upright, sunk between, and crowded aside by the batholiths and are not volcanic chains (Bleeker, 2002). Late clastic sediments, eroded mostly from batholiths and silicic volcanoes that rose as islands, blanketed older units (Fig. 4).

Deformation of Archean crust defines a quasi-floating tectonic style. Modest regional syndiapiric shortening and orthogonal extension oriented batholiths, and strike-slip faults disrupted them, in some regions severely, but this deformation did not produce thrust blocks, strongly overturned structures, regional highlands, or major variations in crustal level of exposure—features common only after ca. 2 Ga. Upper-crustal granite-and-greenstone assemblages are sharply decoupled (e.g., Goleby et al., 2002) from undulating TTG gneisses of the middle and lower crust. Where best studied (e.g., Moser et al., 1996), the deep gneisses show pervasive vertical flattening, horizontal extension parallel to granite-and-greenstone elongation, and long-lasting near-magmatic temperatures.

Earth had a hydrosphere continuously after 3.5 Ga, or perhaps 3.8 Ga, but what was its earlier history? Superdense atmosphere and weak young Sun are complicating factors. Transient oceans and a hot early ocean under a greenhouse atmosphere are possibilities. The oldest proved waterlaid rocks, 3.5 Ga, are 0.9 Ga younger than the oldest zircons in basement. Although supracrustal rocks in SW Greenland are widely regarded as partly older than the oldest zircons, 3.7 or 3.8 Ga, in polycyclic migmatites in contact with them, those migmatites might instead be basement remobilized during the later severe plutonism and deformation. The supracrustals may be Neoarchean, as are those in palinspastically contiguous Labrador. An undocumented claim has been made recently for 3.8 Ga high-grade supracrustals in NW Quebec.

Voluminous mafic and ultramafic melts reached the surface only after ancient felsic crust cooled (in part because of condensation of hydrosphere?) below the density of such melts. Early low islands and shallow basins were covered by thickening shallow-water lava plains. Dense supracrustal assemblages insulated and foundered into basement that rose, variably remelted, in diapiric batholiths of regionally uniform areal density and, commonly, terminal-emplacement age (e.g., Collins and Van Kranendonk, 1998; cf. Griffin et al., 2003a; Fig. 3). Inherited zircons are common. Heat conducted from upper mantle, and internal radiogenic heating 3× greater than now, facilitated mobilization. Concentration of radioactivity high in the crust in fractionating batholiths led to cooling and cratonization. These structural and magmatic assemblages have no Phanerozoic analogs.

Plate tectonics did not operate within preserved Archean crust. Supracrustal assemblages display none of the many stratigraphic, structural, and petrologic indicators of rifting, trailing-edge stratal wedges, oceanic separation, subduction, and collision that characterize plate-tectonic interactions of post–2.0 Ga continents. Vague geochemical analogies of Archean igneous rocks with modern rocks are widely cited to support conjecture that diverse plate-tectonic settings are recorded by intercalated supracrustal lithologies and that layer-parallel successions were assembled from sheets, often thinner than 100 m, thrust hundreds of kilometers from unknown directions. I see no credible evidence for such conjecture. Rocks and associations at issue are quite different from proposed analogs—e.g., Archean tonalites and felsites are far more felsic and have steeper rare earth element patterns than the Phanerozoic arc rocks to which they are wrongly likened.

Proterozoic Tectonics

Changes through Proterozoic time reflect continued mantle cooling. The little preserved very early Proterozoic (ca. 2.5–2.0 Ga) crust contains komatiitic basalts and suggestions of granite-and-greenstone tectonics, which otherwise is restricted to the 3.5–2.5 Ga Archean. The latest Archean and earliest Proterozoic of South Africa and Western Australia include thick supracrustal assemblages of greenstone-belt type, deformed slightly by rejuvenation of diapiric granites in underlying granite-and-greenstone crust (Fig. 3).

Abundant stratigraphic, structural, and petrologic evidence for rifting, rotation, and collision of crustal masses requires initiation of plate tectonics ca. 2.0 Ga. This presumably accompanied cooling of upper mantle through a threshold of thickening and stiffening of oceanic

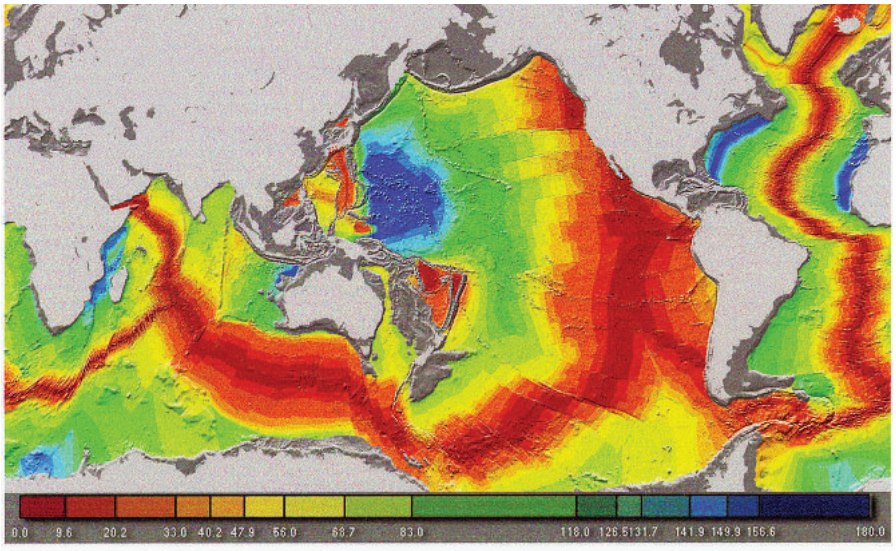

Figure 5. Ages of oceanic crust. Pacific Ocean shrinks because subduction hinges roll back into it, yet spreads rapidly because broadside-sinking slabs push upper mantle back under it. Pacific-plate spreading geometry and rates continued smoothly through the 45 Ma age (orange bands) of the Emperor-Hawaii elbow, disproving speculation that those chains formed above fixed plume. Unpublished 2003 compilation provided by R.D. Müller; pre–80 Ma patterns poorly constrained.

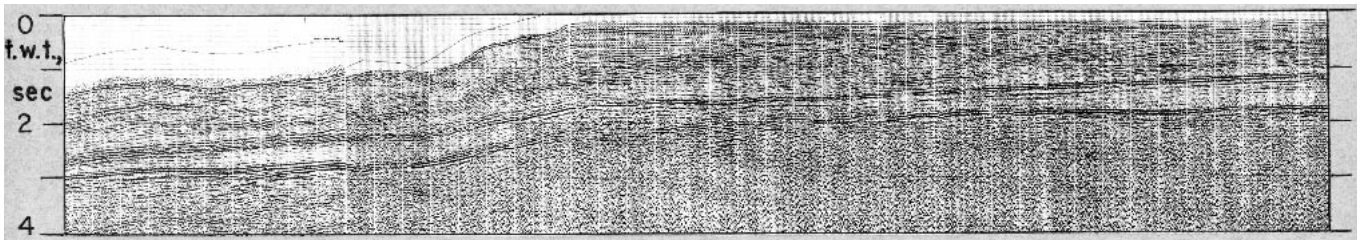

Figure 6. Seismic-reflection profile of part of Sumatra fore-arc basin. Section 35 km long; 1:1 for V = 3 km/s, so VE on seafloor 2:1; Sumatra coast ~30 km to right (NE), fore-arc ridge, near Enggano island, 60 km to left (SW); trench, 170 km. Subducting slab is ~40 km beneath SW end of section. Thin front of overriding plate has not been crumpled since deposition of lower Miocene limestone (semi-transparent) at base of section, so velocities of hinge rollback and arc advance have been equal. Profile from White Shield Indonesia.

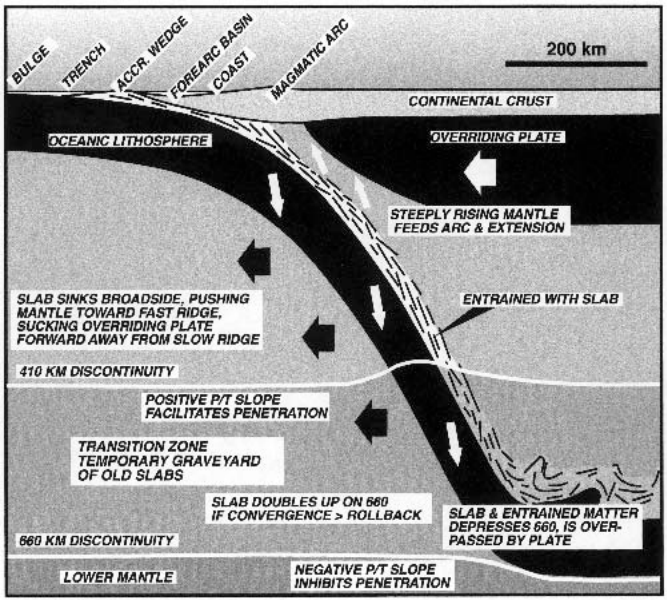

Figure 7. Subduction drive of plate tectonics, illustrated for a common type of continent-margin system. Hinge rolls back so slab sinks broadside, pushing sub-plate upper mantle back under shrinking ocean and forcing rapid spreading, and sucking overriding plate forward. Slab plus entrained material is plated down on uncrossable 660 km discontinuity, overridden by continent, and recycled into slowly enlarging ocean offscreen to right. Shallow part scaled to modern Sumatra and Cretaceous California; low-velocity mantle wedge deduced from high-resolution tomography in Japan; kinematic inferences from this paper. Drafted by Dietrich Roeder, Murnau Geodynamics, Inc.

lithosphere, and shallowing of the gabbro-eclogite transformation, which enabled the subduction that drives plate tectonics. Evolving mantle (Griffin et al., 2003b) and orogenic assemblages show that upper mantle has continued to cool since 2.0 Ga. Giant Proterozoic radially injected gabbroic dike swarms require more decoupling between crust and mantle than now operates. High-pressure low-temperature metamorphic rocks are unknown in subduction complexes older than late Proterozoic, and rare in those, but are widespread in late Phanerozoic ones.

DYNAMICS OF MODERN EARTH

Plate Kinematics

Earth's outer shell is now broken into plates that variously converge, diverge, and move past one another. About one-sixth of the area of the plates undergoes obvious internal deformation. Present relative motions of internally rigid parts of plates are constrained primarily by seafloor-spreading geometry, which integrates motions over millions of years, and by geodesy, mostly satellite, which defines motions over a decade or so and also quantifies internal deformation. The two data sets mostly are broadly compatible (Kreemer et al., 2003).

What drives plates? Most published models and explanations overlook known plate characteristics and behavior and instead elaborate false assumptions.

Ridge Spreading

Ridges form where near-solidus asthenosphere wells into gaps between diverging plates of oceanic lithosphere. Oceanic plates commonly are topped by 5 km or so of basaltic, gabbroic, and cumulate rocks erupted at spreading centers and

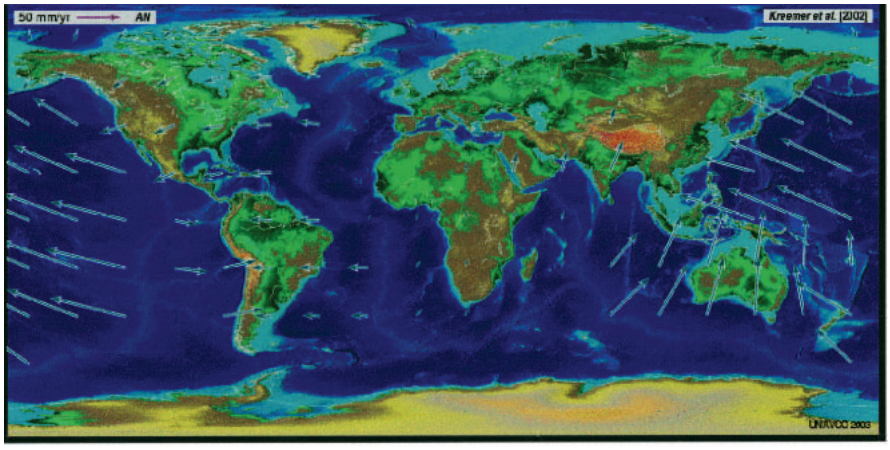

Figure 8. Relative-motion vectors of plates in Antarctica-fixed framework. Predictions that subduction, due to top-down cooling, drives plates, that ridges are passive responses to that drive, and that lower mantle does not participate in plate motions are satisfied in this framework when allowance is made for back-arc spreading and plate-internal deformation. Data from geodesy-dominated model by Kreemer et al. (2003). Image provided by L.H. Estey, University NAVSTAR Consortium; see http://jules.unavco.ucar.edu/Voyager/ILP_GSRM.

consist otherwise of lithospheric mantle formed by top-down cooling of asthenosphere. Slow-spreading ridges are almost-amagmatic extensional complexes. Ridges migrate and change shape and length with time (note ridges around Antarctica and Africa: Fig. 5) in response to motions of bounding plates—ridges are not fixed divergences—and thus tap fresh asthenosphere. The fastest-spreading ridge, south-tropical East Pacific Rise, is crossed by detailed geophysical transects. Bathymetry, volcanism, P and S wave tomography, Rayleigh-wave phase velocities, shear-wave splitting, and electrical resistivity all show partial melting in high upper mantle to extend much farther west than east of the spreading axis (Toomey et al., 2002). I see a westward-migrating passive-ridge gap that progressively depressurizes new, fertile upper mantle, from which melt filters upward toward the advancing tent-shaped gap—a stationary ridge would quickly exhaust available melt—whereas Toomey et al. postulated lateral injection from a far-distant plume.

Subduction

The key to plate dynamics is rollback of hinges into subducting plates as oceanic lithosphere, made dense by top-down cooling of asthenosphere, falls away. Hinges have large radii and are under, not in front of, overriding plates (Fig. 7). Ubiquitous cartoons notwithstanding, hinges are not abrupt downturns at steep-sided trenches, which do not exist. Bounding hinges roll back into the Pacific so it shrinks, despite its rapid internal spreading (Fig. 5). Other features requiring rollback include advance of small hinges into large plates, arc-arc collisions, and increasing arc curvatures with time. Rollback shows that slabs sink more steeply than their inclined seismic zones, which define positions, not trajectories. Accretionary wedges are overridden by plate bases inclined gently arcward (Fig. 7). Inner trench slopes in arcs lacking sediment supply for wedges have diverse lithologies, as dredged, consistent with imbrication, and these trenches also lack steep sides.

Fronts of overriding plates migrate with retreating hinges, and shapes of fronts and hinges tend to change together to maintain accordance. Where present, fore-arc basins atop thin leading edges of overriding plates (Figs. 6, 7) and island arcs show that those edges were little shortened or deformed during periods recorded by basin sedimentation, and require velocities of advance of fronts of overriding plates to have equaled those of rollback. The Cretaceous and Paleogene fore-arc basin of California records 100 m.y. of such equality. Plate geometry requires that convergence is faster than rollback in some places (e.g., Nazca-Chile) and equal to it in others (e.g., Antarctica–south Chile and Caribbean-Atlantic).

Where intra-arc or back-arc spreading occurs, as is common, the front of an overriding plate moves faster than the back. Crustal thickening due to arc magmatism and underplating often has been mistaken for shortening. Where intra-arc shortening does occur, it usually is accomplished from the bottom, not the front, of an overriding plate. Plate fronts are crumpled primarily when arcs collide.

Seismicity and high-resolution tomography show that the only directly overlying material with which a slab is in thrust contact is the accretionary wedge and fore arc. Seismicity under arc and back arc is within a slab, not at its top, and chilled mantle directly above a slab is coupled to it. Hot low-seismic-velocity, high-seismic-attenuation material characterizes the medial part of the mantle wedge between slab and overriding lithosphere and rises toward the magmatic arc (Fig. 7). This near-zero-strength zone, not the top of the slab, marks the main pull-apart plate boundary from which slab and coupled material sink. Hot material sucked up into the wedge above a sinking slab feeds the magmatic arc, the position of which is determined kinematically.

Old subducting slabs (lithosphere older than ca. 60 Ma when it began to sink) are shown by earthquakes and by high-amplitude tomographic anomalies to be plated down on the 660 km discontinuity, which they depress by 20 km or so and may not penetrate (e.g., Pankow and Lay, 2002) and are defined at least as far as 1000 km (possibly 3000 km) horizontally from the base of an inclined slab. Slabs downplated on 660 are overpassed by plates; they do not represent (as most modelers have assumed) lateral injection of slabs. Young slabs, or small-basin slabs, sink only to middle upper-mantle levels, into which they are mixed as they are passed over by arcs and plates.

Upper-Mantle Circulation

660 km Discontinuity

Circulation may be closed above the sharp discontinuity at a depth of 660 ± 20 km, which separates upper from lower mantle and is the most profound seismic break within the mantle. The 660 represents a velocity step and a strong decorrelation of seismic waves, and velocity patterns of all sizes are almost randomly mismatched across it (Gu et al., 2001). The 660 km discontinuity is at the base of the transition zone, wherein seismic velocity increases rapidly downward from the upper-boundary velocity step at 410 ± 15 km. That the 410 represents a density-phase change with a positive P/T slope and the 660 one with a negative slope, is indicated by their weak inverse depth variations as functions of transition-zone velocities, and by mineral physics and experimental petrology (Lebedev et al., 2003). The 410 phase change facilitates sinking of cold material and rising of hot, whereas the 660 change inhibits, and may prohibit, penetration in either direction. Subducted slabs could sink through 660 only after being heated and thereby losing thermal negative buoyancy. Further crossability arguments hang on specific assumptions of P/T slope and of bulk and mineral compositions. If the 660 transformation indeed is uncrossable by material, then circulation above and below is layered, and the boundary must now be compositional as well as a phase change.

That much of the lower mantle is far below its solidus T and thus has very high effective viscosity that precludes rapid circulation is indicated by temperature constraints, experimental petrology, and little-varying V_p/V_s. The very low thermal expansivity at lower-mantle pressure requires that factors other than temperature (grain size, anisotropy, mineralogy including mixed-phase gradients, bulk composition) control much variation in velocity (Hofmeister and Criss, 2003).

No Deep Subduction?

Some seismic tomographers depict low-amplitude positive-velocity anomalies, inferred to mark subducted slabs, deep into lower mantle. Penetration would be most expected in East Asian and western Pacific systems that involve the oldest, densest, and thickest subducting slabs (Fig. 5), yet is not seen there with high-resolution tomography (Grand, 2002; Pankow and Lay, 2002). Interpretations of deep subduction elsewhere are shaky.

Tomographic theory requires waves crossing all mantle volumes from many directions—but the extremely irregular distribution of earthquakes and seismographs results in no sampling of vast volumes and of sampling of much else primarily by bundles of minimally crossing waves. Most earthquakes occur near tops of subducting slabs, and high slab velocities not incorporated in models are smeared downward and gain directional bias from slab anisotropy. Purported deep slabs are defined mostly by waves, subparallel to them, from such subduction quakes. Effects of mapping upper-mantle energy into lower mantle are magnified because of contrasted amplitudes of shallow and deep anomalies. Many published models are misleading: colors are saturated for lower-mantle anomalies one-tenth or one-fifth the amplitude of upper-mantle ones, huge unsampled volumes are assigned average values or populated with spherical-harmonic artifacts, illusory continuity is generated by severe smoothing and sharpening, cross sections are placed where subduction interpretations look most plausible. Models account for only a few tenths of traveltime variance, and have not been tested by attempting to squeeze solutions back into upper mantle. Nonsubduction alternatives receive little evaluation. (Zhou et al., 1990, provided an exception.)

The visually most impressive example of purported deep subduction, an anomaly inclined eastward under the Caribbean region through the entire mantle, is depicted by several tomographers (Grand, 2002; most use the same database). It is widely cited as conclusive evidence for rapid whole-mantle convection regardless of other considerations. The anomaly has the problems noted above, and it lacks the along-strike continuity required by its slab interpretation. Sections drawn through the same models, north and south of published sections placed where anomalies plausibly depict subduction, would show disconnected blobs and irregular shapes, despite severe smoothing, in erratic positions and "wrong" and multiple directions. More detailed tomography by a different method requires depression, but not penetration, of the 660 between 5° and 30°S (Collier and Helffrich, 2001).

Plumes Do Not Exist

A huge literature (e.g., Condie, 2001) speculates that narrow plumes of hot, fertile material rise from the base of the mantle to upper mantle and crust. The notion descends from conjecture (rejected earlier) that the mantle is compositionally inverted, and also from speculation that the general SSE, then ESE, age progression in Emperor and Hawaiian seamounts and islands was produced by a stationary plume squirting up through the moving Pacific plate. There, in its only rigorously testable locale, the fixed-plume concept is falsified. That the Pacific plate did not change direction 60°, as required by the conjecture, at ca. 45 Ma (the age of the Emperor-Hawaiian elbow) is proved by paleomagnetism and by continuity of spreading geometry and rates (as defined by spreading angles and transform directions) through that age interval around the entire east and south margins of the plate (Fig. 5). Hawaiian heat flow, tomography, and positive correlation between volume of volcanism and rate of propagation further falsify plume predictions. Plume advocates respond to such refutations by making their conjectures ever more complex, unique to each example, and untestable.

All other geochemical, kinematic, and tectonic plume rationales derive from the misinterpretation that Hawaii stands atop a plume. Circular rationales assign diverse magmatic and tectonic provinces to plumes, but I here discuss only oceanic basalts. Ocean-island basalts (OIB) tend to be more fertile than most mid-ocean ridge basalts (MORB), and this contrast is widely assumed to reflect derivation of OIB from enriched lower-mantle plumes and of MORB from depleted upper mantle. The false logic has been that ridges are fed from upper mantle, islands are not ridges, and therefore islands are not fed from upper mantle. As MORB in fact varies widely and largely overlaps OIB, a narrowly selected subset of high-Mg, low-alkali ridge rocks (N-MORB, often mislabeled "MORB") has then been assumed to be the "normal" product of asthenosphere melting, whereas voluminous interspersed and intercalated, or distant, ridge basalts higher in incompatible elements and lower in compatibles are assigned, with circular reasoning, to plumes, often as lateral injections of thousands of kilometers. Plume proponents have not addressed the OIB composition of large seamounts, which are abundant and mostly randomly distributed (J.H. Natland and E.L. Winterer, 2003, written commun.).

Pro-plume conjectures assume wrongly that magmas are transported from melting sites to surface without modification, thereby preserving elemental and isotopic ratios of sources. Thermodynamics and phase petrology, including cotectic compositions of basalts and presence of underlying cumulates, prove that instead melts are profoundly modified before eruption (O'Hara and Herzberg, 2002). Simplistic geochemical source modeling is further invalidated by the occurrence of incompatible elements in mantle rocks primarily in intergranular films of secondary minerals rather than in major-mineral phases. The contrasts that do exist between many ridge and island basalts and within ridge basalts mostly are required by contrasts in polybaric exothermic crystallization and endothermic assimilation by which melts are forced to evolve while rising through asthenospheric and lithospheric columns differing in thermal gradients and composition.

No plumes extending downward into lower mantle have been detected by credible tomography. Purported-plume displays are flawed in methodology, artifacts, and presentation, most purported hotspots do not overlie abnormally hot upper mantle, most tomographically inferred hot regions are not overlain by purported hotspots, and there is no indication of plumes in upper-mantle geophysics (Anderson, 2000).

Oceanic asthenosphere is everywhere near solidus temperature so local excess

heat is not required for island volcanism. Access of melt to surface, as by propagating rifts, is needed—crackspots, not hotspots. Island alignments, in this view, reflect regional stresses, perpetuation of directions once established, and lithosphere properties. The Emperor and Hawaiian chains avoid huge basalt plateaus, and the elbow between them occurs near a fracture zone across which lithosphere age changes by 10 or 15 m.y.

Heterogeneity

Recycling by subduction of oceanic lithosphere and continental material has made upper mantle increasingly heterogeneous and more fertile. Old slabs and entrained material sink into the transition zone and await lateral and upward recycling, whereas young slabs mix into upper mantle. Temperature, age, and isotopic ratios of sunken slabs vary regionally. Contrasts within and between MORBs and other oceanic basalts presumably partly reflect upper-mantle heterogeneity and also varying degrees of initial melting.

Driving Mechanism of Plate Tectonics

Top-down Subduction Drive

Cooling from above by seawater produces oceanic lithosphere denser than asthenosphere from which it forms, and this density inversion is righted by gravitationally driven subduction that controls plate motions (cf. Anderson, 2002b). Plate behavior accords with self-organized subduction.

Mechanism of Plate Tectonics

Subduction is the primary drive for plate motions, and its effects are explicable by circulation closed above 660 km (Fig. 7). Slabs and entrained material sink broadside from retreating hinges so mantle beneath slabs is pushed backward, and overriding plates or island arcs are sucked forward. Old slabs sink to 660 km and are overpassed by overriding plates and thus transferred out of the Pacific for recycling into enlarging oceans or back-arc basins. Pacific shrinkage is due to removal of lithosphere, but the volume of upper mantle—everything from 660 km to the base of lithosphere—pushed back under the ocean by slabs is much greater than that of departing lithosphere, so Pacific spreading is much faster than that of passive oceans. The volumes of lithosphere transferred from shrinking oceans, of enlargement of passive oceans, and of advance of overriding plates and arcs are approximately equal. Where convergence is faster than rollback and advance, sunken slabs are doubled up and transfer is accelerated. Overpassed slabs are heated from above as well as from below and so partly insulate continents from lower-mantle heat. Downplated slabs are recycled from the transition zone when new slabs crowd them out, or into spreading oceans, and become magma sources as they rise into basalt-melting P/T fields. Hot material sucked forward by slabs provides heat and substance for back-arc spreading and arc magmatism.

No Bottom-Up Drive

Observed plate features and interactions—including migration and changing shapes and lengths of ridges, rollback and other behavior of hinges, matching of velocities of rollback and advance—are incompatible with popular models wherein plates are colliding rafts dragged from beneath by currents flowing from hot upwellings to cool downwellings. The few bottom-drive models that add awareness that subduction is enabled by top-down cooling (Bercovici, 2003) also cannot explain the observations.

Framework and Test

The subduction-drive mechanism predicts plate motions relative to sluggish lower mantle. Relative plate motions can be transposed to any reference frame, so there should be a frame wherein hinges roll back, directions and rates of plate rotations relate systematically to subduction, and ridges migrate and spread at rates compatible with subduction drive and slab transfer. These predictions are not met in the popular hotspot and no-net-rotation frameworks. The stationary-hotspot frame is tied to Hawaii and is invalidated by disproof of a plume fixed there. The no-net-rotation framework minimizes motions of ridges and randomizes motions of subduction hinges so that many roll forward—impossible with gravity drive. The lack of rational kinematics in these popular frameworks fosters acceptance of mysterious bottom drives.

Subduction-drive predictions are met when plate motions are transposed to the Antarctica-fixed framework (Fig. 8). Antarctica is ringed by spreading ridges (Fig. 5) and should be nearly stationary, as is suggested also by its Cenozoic paleomagnetism. (It moves rapidly in hotspot and no-net-rotation frames.) In fixed-Antarctica frame, all hinges roll back when back-arc spreading and internal deformation of plates are integrated. Ridges migrate to tap fresh asthenosphere. South America is sucked toward its west-bounding subduction systems. North America moves toward subduction boundaries and obliquely retreating Pacific plate. Africa, almost ringed by ridges, is rotating very slowly toward minor subduction systems in the northeast instead of moving rapidly east as in hotspot and no-net-rotation frames. Stable Eurasia, away from its deforming southern and eastern regions, is almost stationary in accord with its lack of subduction boundaries instead of moving rapidly. South and East Asian deformation clearly relates to subduction and rollback, although the continuing northward motion of India is puzzling. Widely varying spreading rates accord with transfer of overpassed oceanic lithosphere from shrinking to enlarging oceans as the basic control, complicated by factors such as the rapid sliding of narrow eastern Pacific plates down their inclined bases toward South and Central American hinges. Lithosphere has a net westward drift, presumably due to Earth-spin influence.

OVERVIEW

The Earth described here differs profoundly from that accepted as dogma in most textbooks and research papers. Crust and upper mantle have formed a mostly closed system throughout geologic time, and their dramatic temporal changes are responses to cooling. The changing processes define a Punctuated Gradualism and not Uniformitarianism. Major stages in Earth evolution:

(1) 4.567–ca. 4.4 Ga. Hot accretion and major irreversible mantle fractionation. Giant bolides continue to ca. 3.9 Ga.

(2) 4.4–3.5 Ga. Era of nearly global felsic crust, too hot and mobile to stand as continents.

(3) 3.5–2.0 Ga. Granite-and-greenstone era. Permanent hydrosphere. Old crust cooled to density permitting mafic melts to reach surface. Diapiric batholiths mobilized from underlying old crust.
(4) 2.0 Ga–continuing. Plate tectonic era. Distinct continents and oceans. Top-down cooling of oceanic lithosphere enables subduction that drives plates, forces spreading, and mixes continental as well as oceanic crust into upper mantle.

ACKNOWLEDGMENTS

This essay wanders far from my primary expertise. Discussions with W. Bleeker, T.M. Boyd, W.R. Dickinson, G.R. Foulger, S.P. Grand, A.M. Hofmeister, J.H. Natland, M.J. O'Hara, S.Y. O'Reilly, J. Ritsema, E.L. Winterer, and particularly D.L. Anderson have contributed importantly to concepts advocated here. Presentation was much improved as a result of manuscript reviews by Foulger, K.A. Howard (who invited me to write it), R.V. Ingersoll, G.J.H. McCall, W.D. Mooney, and T. Skulski.

REFERENCES CITED

Alexander, C.M.O., Boss, A.P., and Carlson, R.W., 2001, The early evolution of the inner solar system—a meteoritic perspective: Science, v. 293, p. 64–68.

Allègre, C., Manhès, G., and Levin, É., 2001, Chemical composition of the Earth and the volatility control on planetary genetics: Earth and Planetary Science Letters, v. 185, p. 49–69.

Anderson, D.L., 2000, The thermal state of the upper mantle; no role for mantle plumes: Geophysical Research Letters, v. 27, p. 3623–3626.

Anderson, D.L., 2002a, The case for irreversible chemical stratification of the mantle: International Geology Review, v. 44, p. 97–116.

Anderson, D.L., 2002b, Plate tectonics as a far-from-equilibrium self-organized system: American Geophysical Union Geodynamics Series, v. 30, p. 411–425.

Artemieva, I.M., and Mooney, W.D., 2001, Thermal thickness and evolution of Precambrian lithosphere—A global study: Journal of Geophysical Research, v. 106, p. 16,387–16,414.

Ayer, J., Amelin, Y., Corfu, F., Kamo, S., Ketchum, J., Kwok, K., and Trowell, N., 2002, Evolution of the southern Abitibi greenstone belt based on U-Pb geochronology; autochthonous volcanic construction followed by plutonism, regional deformation, and sedimentation: Precambrian Research, v. 115, p. 63–95.

Badro, J., Fiquet, G., Guyot, F., Rueff, J.-P., Struzhkin, V.V., Vanko, G., and Monaco, G., 2003, Iron partitioning in Earth's mantle—Toward a deep lower mantle discontinuity: Science, v. 300, p. 789–791.

Bercovici, D., 2003, The generation of plate tectonics from mantle convection: Earth and Planetary Science Letters, v. 205, p. 107–121.

Bleeker, W., 2002, Archaean tectonics—a review, with illustrations from the Slave craton: Geological Society of London Special Publication 199, p. 151–181.

Britt, D.T., and Lebofsky, L.A., 1999, Asteroids, in Weissman, P.R., et al., eds., Encyclopedia of the Solar System, San Diego, Academic Press, p. 585–605.

Collier, J.D., and Helffrich, G.R., 2001, The thermal influence of the subducting slab beneath South America from 410 and 660 km discontinuity observations: Geophysical Journal International, v. 147, p. 319–329.

Collins, W.J., and Van Kranendonk, M.J., 1998, Partial convective overturn of Archaean crust in the east Pilbara craton, Western Australia—driving mechanisms and tectonic implications: Journal of Structural Geology, v. 20, p. 1405–1414.

Condie, K.C., 2001, Mantle plumes and their record in Earth history: Cambridge, Cambridge University Press, 306 p.

Durrheim, R.J., and Mooney, W.D., 1994, Evolution of the Precambrian lithosphere—seismological and geochemical constraints: Journal of Geophysical Research, v. 99, p. 15,359–15,374.

Goleby, B.R., Korsch, R.J., Fomin, T., Bell, B., Nicoll, M.G., Drummond, B.J., and Owen, A.J., 2002, Preliminary 3-D geological model of the Kalgoorlie region, Yilgarn Craton, Western Australia, based on deep seismic-reflection and potential-field data: Australian Journal of Earth Sciences, v. 49, p. 917–933.

Grand, S.P., 2002, Mantle shear-wave tomography and the fate of subducted slabs: Royal Society of London Philosophical Transactions, v. A-360, p. 2475–2491.

Green, M.G., Sylvester, P.J., and Buick, R., 2000, Growth and recycling of early Archaean continental crust—geochemical evidence from the Coonterunah and Warrawoona Groups, Pilbara Craton, Australia: Tectonophysics, v. 322, p. 69–88.

Griffin, W.L., Belousova, E.A., Shee, S.R., Pearson, N.J., and O'Reilly, S.Y., 2003a, Crustal evolution in the northern Yilgarn Craton—U-Pb and Hf-isotope evidence from detrital zircons: Precambrian Research (in press).

Griffin, W.L., O'Reilly, S.Y., Abe, N., Aulbach, S., Davies, R.M., Pearson, N.J., Doyle, B.J., and Kivi, K., 2003b, The origin and evolution of Archean lithospheric mantle: Precambrian Research, doi:10.1016/S0301-9268(03)00180-3 (in press).

Gu, Y.J., Dziewonski, A.M., Su, W., and Ekström, G., 2001, Models of the mantle shear velocity and discontinuities in the pattern of lateral heterogeneities: Journal of Geophysical Research, v. 106, p. 11,169–11,199.

Hamilton, W.B., 2002, The closed upper-mantle circulation of plate tectonics: American Geophysical Union Geodynamics Series, v. 30, p. 359–410.

Herzberg, C., 1993, Lithosphere peridotites of the Kaapvaal craton: Earth and Planetary Science Letters, v. 120, p. 13–29.

Hill, R.E.T., Barnes, S.J., and Dowling, S.E., 2001, Komatiites of the Norseman-Wiluna greenstone belt, Western Australia—A field guide: Geological Survey of Western Australia Record 2001/10, 71 p.

Hofmeister, A.M., and Criss, R.E., 2003, Evidence for layered mantle convection-implications for lower mantle plumes, abstract in The Hotspot Handbook, Proceedings of Penrose Conference Plume IV; Beyond the Plume Hypothesis: Hveragerdi, Iceland, August 2003, 3 p.

Jacobsen, S.B., 2003, How old is planet Earth?: Science, v. 300, p. 1513–1514.

Kamber, B.S., Collerson, K.D., Moorbath, S., and Whitehouse, M.J., 2003, Inheritance of early Archaean Pb-isotope variability from long-lived Hadean protocrust: Contributions Mineralogy and Petrology, v. 145, p. 25–46.

Kreemer, C., Holt, W.E., and Haines, A.J., 2003, An integrated global model of present-day plate motions and plate boundary deformation: Geophysical Journal International, v. 154, p. 8–34.

Lebedev, S., Chevrot, S., and van der Hilst, R.D., 2003, Correlation between the shear-speed structure and thickness of the mantle transition zone: Physics of the Earth and Planetary Interiors, v. 136, p. 25–40.

Miller, C.F., McDowell, S.M., and Mapes, R.W., 2003, Hot and cold granites? Implications of zircon saturation temperatures and preservation of inheritance: Geology, v. 31, p. 529–532.

Moser, D.E., Heaman, L.M., Krogh, T.E., and Hanes, J.A., 1996, Intracrustal extension of an Archean orogen revealed using single-grain U-Pb zircon geochronology: Tectonics, v. 15, p. 1093–1109.

O'Hara, M.J., and Herzberg, C., 2002, Interpretation of trace element and isotope features of basalts—relevance of field relations, petrology, major element data, phase equilibria, and magma chamber modeling in basalt petrogenesis: Geochimica et Cosmochimica Acta, v. 66, p. 2167–2191.

Pankow, K.L., and Lay, T., 2002, Modeling S wave amplitude patterns for events in the Kurile slab using three-dimensional Gaussian beams: Journal of Geophysical Research, v. 107, no. B8, paper ESE 6, 22 p.

Scherer, E., Münker, C., and Mezger, K., 2001, Calibration of the lutetium-hafnium clock: Science, v. 293, p. 683–687.

Toomey, D.R., Wilcock, W.S.D., Conder, J.A., Forsyth, D.W., Blundy, J.D., Parmentier-E.M., and Hammond, W.C., 2002, Asymmetric mantle dynamics in the MELT region of the East Pacific Rise: Earth and Planetary Science Letters, v. 200, p. 287–295.

Turcotte, D.L., and Schubert, G., 2002, Geodynamics: Cambridge, Cambridge University Press, 456 p.

Urey, H.C., 1951, The origin and development of the earth and other terrestrial planets: Geochimica et Cosmochimica Acta, v. 1, p. 209–277.

Wetherill, G.W., 1994, Provenance of the terrestrial planets: Geochimica et Cosmochimica Acta, v. 58, p. 4513–4520.

Zhang, Y., 2002, The age and accretion of the Earth: Earth-Science Reviews, v. 59, p. 235–263.

Zhou, H.-w., Anderson, D.L., and Clayton, R.W., 1990, Modeling of residual spheres for subduction zone earthquakes. 1. Apparent slab penetration signatures in the NW Pacific caused by deep diffuse mantle anomalies: Journal of Geophysical Research, v. 95, p. 6799–6827.

Manuscript submitted March 12, 2003; accepted August 11, 2003.