Origin of the Moon, Impactor Theory

Caitlin Ahrens

NASA Goddard Space Flight Center, Greenbelt, MD, USA

Definition

The Giant Impactor Theory is currently the leading theory for the origin of Earth's Moon. Scenarios for lunar origin should be able to explain the Moon's physical and dynamical properties, including chemical composition (volatile and stable isotope ratios), and the Moon's initial thermal state. Recent work on the evolution of the proto-lunar disk give insight to the Moon's early thermal and chemical history. Samples from the Moon, impact-generated disk processes, and impact parameters have yielded some unique geophysical challenges toward this theory.

Impactor Theory

In the past few decades, the Giant Impact theory has arisen as the leading formation explanation of the Moon (Hartmann and Davis 1975; Cameron and Ward 1976; Wood 1986). The current iteration of the theory involves the collision between two planetary embryos during the late stages of planetary accretion (Canup and Asphaug 2001; Canup 2004). Numerous numerical simulations have identified a specific impact scenario, referred to as the *canonical impact* (Canup 2014), where a proto-Earth experiences a collision with a Mars-sized protoplanet (Theia, discussed below; Canup 2004). This in turn would place more than a lunar mass of silicate-rich material into orbit (Asphaug 2014; Barr 2016).

Much of the material in the proto-disk originates from the referred impactor (Canup 2004). In the few hundred to thousand years after the fateful impact, the material within the disk and silicate vapor atmosphere of the proto-Earth cool and condense (Ward 2012; Charnoz and Michaut 2015). This post-impact material spreads viscously beyond the Earth's Roche limit, where gravitational tidal forces are small enough to allow accretion, clumping and accreting into the Moon (Ida et al. 1997; Kokubo et al. 2000; Asphaug 2014; Barr 2016).

Impact conditions influence the composition, mass, and phase of orbiting material, all important factors for ultimately determining the current state of the Moon (Canup et al. 2001; Nakajima and Stevenson 2014). From several works (Ward 2012; Salmon and Canup 2012; Barr 2016), disk simulations have included enough more complex physical processes to provide a self-consistent image of the Moon originating from impact-induced debris.

Evidence for Planet-Scale Collision

Such a geophysical problem of large oblique impactors is particularly challenging with a lack of experimental or observations to validate such a numerical model of this size. While laboratory-scale experiments could construct scaling relationships, these experiments are not capable of simulating the collision of self-gravitational dominant objects (Barr 2016). At present, the planets within the Solar System are not arranged in orbits in which collision is an issue, so observations of such collisions in other systems are possible, but still years away (Miller-Ricci et al. 2009).

Any (successful) scenario to model the Moon's origins must satisfy these observations, as stated by Barr (2016): (i) a planet of roughly Earth's mass $M_E = 5.98 \times 10^{24}$ kg, and a Moon of Mass $M_L = 7.35 \times 10^{22}$ g; (ii) an angular momentum of $L_{EM} = 3.5 \times 10^{34}$ kg m^2 s^{-1} within the system; (iii) a Moon with ~8% iron (by mass); (iv) stable isotope ratios similar to the terrestrial mantle, including a degree of volatile depletion; and (v) a magma ocean initially 200 – 300 km thick. Of note is that the Moon is anomalously large relative to the size of the Earth and contains the bulk of the total angular momentum of the system (MacDonald 1966). The size ratio in turn gives the mass of the Moon ~1% the mass of the Earth, whereas outer planetary satellites (combined) are ~10^{-4} times the mass of their parent planets. Just based on the mass and angular momentum alone that the Moon formed by a process different that those that formed the satellites of the outer planets (Jupiter, Saturn, Uranus, Neptune). The density of the Moon ($\rho = 3.34$ g cm^{-3}) is lower than the density of the Earth (4.4 g cm^{-3}), indicating a lower iron content (Barr 2016). The Earth is about 33% iron by mass, whereas the bulk lunar iron content is ~8 – 10% (Jones and Delano 1989; Lucey et al. 1995; Jones and Palme 2000). From Williams et al. (2014), spacecraft data shows that the Moon has a small iron core (~325 km radius). The mass fraction of iron in the final Moon is usually assumed to be equal to the mass fraction of iron in the disk (Canup 2004; Barr 2016).

All simulations assume that the proto-Earth and impactor behave as strength-less fluids, although strength can affect the outcomes of oblique impacting (Schultz 1996; Stickle and Schultz 2011). Melting and vaporization will be localized, leading to heterogeneity in the post-impact distribution of thermal energy in the impact-generated disk (Rudge et al. 2010; Barr 2016). Therefore, the predicted lunar mass depends on the ratio between the disk angular momentum per unit mass and the angular momentum per unit mass of a circular orbit with a semimajor axis equal to the Roche limit around the Earth (Ida et al. 1997).

However, gravitational forces are thought to control the outcome of an impact (Melosh 1989), which also dominate over material strength (Asphaug et al. 2015; Barr 2016). On the other hand, the strength of the mantles and cores of the impactor and target could also affect the distribution of energy, and consequently the way the impactor and target are heated (and deformed) during the impact event (Grady 1980; Schultz and Gault 1990; Ramesh et al. 2014; Barr 2016).

Salmon and Canup (2012) numerical simulations show that the Moon is assembled in three phases: (i) in the first few years post-Giant Impact, material outside the Roche limit accretes into a lunar "parent body" initially half the mass of the present Moon state; (ii) growth stalls for a few tens to hundred years as the inner disk spreads outward due to viscosity; and (iii) the viscous disk spreads beyond the Roche limit, spawning small bodies (moonlets) that accrete onto the parent body in the next few hundred years (Figure 1). Time-dependent models show that even longer timescales (10^4 – 10^5 years) may also be possible (Charoz and Michaut 2015).

Johnson et al. (2012) also observes that circumstellar debris disks show possible direct signatures of giant impacts. The relevance of such a major planetary-scale collisions boils down to: (i) the combined Earth-Moon angular momentum; (ii) the globally melted geology of the Moon; (iii) lack of lunar volatiles; (iv) lack of a lunar core; and (v) the late-stage timing of formation.

Theia

Dynamical modelers (e.g., Benz et al. 1989; Cameron and Benz 1991; Ida et al. 1997; Canup and Asphaug 2001) along with increasingly advanced simulations, led to the acceptance of the Giant Impact scenario, where a Mars-sized terrestrial object accreted by the proto-Earth (Figure 2). This object is labeled ***Theia***. Concurrently, the isotopic geochemistry of lunar rock samples was leading to the conclusion that the Moon was too Earth-like to include substantial fractions of Theia, unless Theia is supposed to be indistinguishable from Earth (or at least proto-Earth) (Wiechert et al. 2001; Melosh 2009). The lunar petrology, geology, gravity, and corresponding remote sensing all indicate regional (or global) melting of the Moon (Warren 1985; Shearer et al. 2006).

The standard giant-impact model (in a graze-and-merge collision with an impact angle of $\theta \approx 45°$) has Theia in a captured orbit and swings by a second time for a high-angular-momentum merger (Asphaug 2014). Therefore the Moon is mostly made out of Theia in the model.

A planet-scale collision would provide the enthalpy for global melting by the mantle of Theia, assuming it is already statically loaded to tens of gigapascals (Asphaug et al. 2006) and further loaded by shock (Stevenson 1987), exploding upon decompression into an orbiting torus of melt and vapor. However, the production of a melted-vaporized disk is a good indication for a rather dry Moon scenario, but does not satisfy all the details (Asphaug 2014) such that the dependency on temperature, pressure and droplet size requires lunar accretion to be slower than vapor diffusion (Asphaug 2014). Conversely, (Hauri et al. 2011) states that coagulation must occur before volatile loss was complete to account for the water (and retention of sodium and potassium) retained by the Moon. Such an impact from a Mars-sized object like Theia would be consistent with a terrestrial magma ocean (Elkins-Tanton 2012) and further explain the loss of the primordial atmosphere (Ahrens 1993; Genda and Abe 2005). Post-impact processes where a massive silicate atmosphere radiating at ~2,500 K would certainly dissipate out in a few thousand years. Sharp and Draper (2013) consider that a giant impact might have even ejected a primordial ocean.

Isotope Constraints

Overall, models imparting upon late-stage planetary formation from giant collisions is strongly supported by geochemical heterogeneity (Wetherill 1985; Raymond et al. 2006; Rudge et al. 2010). Numerous analyses of lunar samples show that stable isotopic ratios are nearly identical to those in the Earth's mantle (Zhang et al. 2012; Armytage et al. 2012). For example, oxygen isotope ratios in Apollo samples are consistent with mass-dependent fractionation with Earth (Epstein and Taylor 1970; Wiechert et al. 2001), and lunar anorthosites have the same $^{53}Cr/^{52}Cr$ ratios as the terrestrial mantle (Lugmair and Shukolyukov 1998). Such geochemical analyses provide clues to the bulk chemistries of the Moon, especially for the amount of volatile materials retained in the accretion-stage of the Moon's formation. Several studies (Taylor et al. 2006; Taylor and Wieczorek 2014) note that the Moon is depleted in certain volatiles relative to the Earth and CI chondrites. Such depletion of volatiles provides constraints on temperatures in the protolunar disk, for example low-Ti basalts indicate a depletion of elements (e.g., K, Na, Zn) with solar condensation temperatures < 1,100 K (Wolf and Anders 1980). However, water volatiles are still an ongoing investigation, especially with the detection of ~40 ppm have been detected in melt inclusions of lunar volcanic glasses (Saal et al. 2008; Chen et al. 2015). Water contents in apatites on

Earth and the Moon are however similar, providing further insight as to some water retention during lunar accretion (Boyce et al. 2010; Greenwood et al. 2011; Boyce et al. 2014).

Isotopic ratio similarities between the Moon and Earth could be explained in several ways: (i) a high percentage of the Moon is terrestrial mantle material; (ii) isotopic equilibration during the Great Impact event; or (iii) that the Earth and the impactor (e.g., Theia) formed from isotopically similar material (Wiechert et al. 2001; Barr 2016).

Take for example oxygen isotopes. The three-isotope system ^{16}O, ^{17}O, ^{18}O is enriched on Earth (McKeegan et al. 2011), though lunar rocks are also enriched with the same levels. The original terrestrial embryos (Mars and presumably Theia) are believed to have accreted within a few million years (Dauphas and Pourmand 2011). Differentiated meteoritic parent bodies have strongly (and uniquely) identifiable isotopic traces, presumably because of their creation within exclusive regions of the protoplanetary disk (Asphaug 2014). However, the ^{17}O ratios for the Moon are well within range of measurements for the silicate Earth (Wiechert et al. 2001). Pahlevan and Stevenson (2007) have argued that turbulent diffusion post-impact would homogenize the isotopic composition of the protolunar disk with Earth’s post-impact silicate atmosphere. This process favors the most energetic giant impact scenarios.

References

Ahrens TJ. (1993). Impact erosion of terrestrial planetary atmospheres. Annu. Rev. Earth Planet. Sci. 21:525–55.

Armytage, R. M. G., Georg, R. B., Williams, H. M., & Halliday, A. N. (2012). Silicon isotopes in lunar rocks: Implications for the Moon’s formation and the early history of the Earth. Geochimica et Cosmochimica Acta, 77, 504-514.

Asphaug, E. (2014). Impact origin of the Moon?. Annual Review of Earth and Planetary Sciences, 42, 551-578.

Asphaug E, Agnor CB, Williams Q. (2006). Hit-and-run planetary collisions. Nature 439:155–60.

Asphaug, E., G. Collins, and M. Jutzi (2015), Global scale impacts, in Asteroids IV, edited by P. Michel, F. E. DeMeo, and W. F. Bottke, pp. 661–678, Univ. of Ariz. Press, Tucson.

Barr, A. C. (2016). On the origin of Earth's Moon. Journal of Geophysical Research: Planets, 121(9), 1573-1601.

Benz, W., A. G. W. Cameron, and H. J. Melosh (1989), The origin of the Moon and the single-impact hypothesis III, Icarus, 81, 113–131.

Boyce, J. W., Y. Liu, G. R. Rossman, Y. Guan, J. M. Eiler, E. M. Stolper, and L. A. Taylor (2010), Lunar apatite with terrestrial volatile abundances, Nature, 466, 466–469.

Boyce, J. W., S. M. Tomlinson, and F. M. McCubbin (2014), The lunar apatite paradox, Science, 344, 400–402.

Cameron, A. G. W., and W. Benz (1991), The origin of the Moon and the single-impact hypothesis IV, Icarus, 92, 204–216.

Cameron, A. G. W., and W. R. Ward (1976), The origin of the Moon, in Lunar and Planetary Science Conference Proceedings, vol. 7, edited by R. B. Merrill, pp. 120–122, Lunar and Planetary Science Institute, Houston, TX.

Canup, R. M. (2004a), Simulations of a late lunar forming impact, Icarus, 168, 433–456.

Canup, R. M. (2014), Lunar-forming impacts: Processes and alternatives, Philos. Trans. R. Soc. A, 372, 20130175.

Canup, R. M., and E. Asphaug (2001), Origin of the Moon in a giant impact near the end of the Earth's formation, Nature, 412, 708–712.

Canup, R. M., W. R. Ward, and A. G. W. Cameron (2001), A scaling relationship for satellite-forming impacts, Icarus, 150, 288–296.

Charnoz, S., and C. Michaut (2015), Evolution of the protolunar disk: Dynamics, cooling timescale and implantation of volatiles onto the earth, Icarus, 260, 440–463

Chen, Y., Zhang, Y., Liu, Y., Guan, Y., Eiler, J., & Stolper, E. M. (2015). Water, fluorine, and sulfur concentrations in the lunar mantle. Earth and Planetary Science Letters, 427, 37-46.

Dauphas N, Pourmand A. 2011. Hf-W-Th evidence for rapid growth of Mars and its status as a planetary embryo. Nature 473:489–92.

Elkins-Tanton LT. 2012. Magma oceans in the inner Solar System. Annu. Rev. Earth Planet. Sci. 40:113–39.

Epstein and Taylor 1970; Epstein, S., and H. P. Taylor (1970), 18O/16O, 30Si/28Si, D/H, and 13O/12O studies of lunar rock and minerals, Science, 167, 533–535.

Genda H, Abe Y. 2005. Enhanced atmospheric loss on protoplanets at the giant impact phase in the presence of oceans. Nature 433:842–44

Grady, D. (1980), Shock deformation of brittle solids, J. Geophys. Res., 85(B2), 913–924.

Greenwood, J. P., S. Itoh, N. Sakamoto, P.Warren, L. Taylor, and H. Yurimoto (2011), Nominally hydrousmagmatism on the Moon, Nat. Geosci., 4, 1–4.

Hartmann, W. K., and D. R. Davis (1975), Satellite-sized planetesimals and lunar origin, Icarus, 24, 504–515.

Hauri EH, Weinreich T, Saal AE, Rutherford MC, Van Orman JA. 2011. High pre-eruptive water contents preserved in lunar melt inclusions. Science 333:213–15.

Ida S, Canup RM, Stewart GR. 1997. Lunar accretion from an impact-generated disk. Nature 389:353–57.

Johnson BC, Lisse CM, Chen CH, Melosh HJ, Wyatt MC, et al. 2012. A self-consistent model of the circumstellar debris created by a giant hypervelocity impact in the HD 172555 system. Astrophys. J. 761:45.

Jones, J. H., and J. W. Delano (1989), A three-component model for the bulk composition of the Moon, Geochim. Cosmochim. Acta, 53, 513–527.

Jones, J. H., and H. Palme (2000), Geochemical constraints on the origin of the Earth and Moon, in Origin of the Earth and Moon, edited by R. M. Canup et al., pp. 197–216, Univ. of Ariz. Press, Tucson.

Kokubo, E., S. Ida, and J. Makino (2000), Evolution of a circumterrestrial disk and formation of a single Moon, Icarus, 148, 419–436.

Lucey, P. G., G. J. Taylor, and E. Malaret (1995), Abundance and distribution of iron on the Moon, Science, 268, 1150–1153.

Lugmair, G. W., and A. Shukolyukov (1998), Early solar system timescales according to 53Mn-53Cr systematics, Geochim. Cosmochim. Acta, 62, 2863–2886

MacDonald, G. J. F. (1966), Origin of the Moon: Dynamical considerations, in The Earth-Moon System: Proceedings of an International Conference, January 20–21, vol. 1964, edited by B. Marsden and A. G. W. Cameron, pp. 165–210, Plenum Press, New York.

McKeegan KD, Kallio APA, Heber VS, Jarzebinski G,Mao PH, et al. 2011. The oxygen isotopic composition of the Sun inferred from captured solar wind. Science 332:1528–32.

Melosh, H. J. (1989), Impact Cratering: A Geologic Process, Oxford Univ. Press, New York.

Melosh, H. J. (2009), An isotopic crisis for the giant impact origin of the Moon?, in 72nd Annual Meteoritical Society Meeting, pp. 5104, Meteoritics and Planetary Science Supplement, Nancy, France.

Miller-Ricci, E., M. R. Meyer, S. Seager, and L. Elkins-Tanton (2009), On the emergent spectra of hot protoplanet collision afterglows, Astrophys. J., 704, 770–780.

Nakajima, M., and D. J. Stevenson (2014), Investigation of the initial state of the Moon-forming disk: Bridging SPH simulations and hydrostatic models, Icarus, 233, 259–267

Pahlevan, K., and D. J. Stevenson (2007), Equilibration in the aftermath of the lunar-forming giant impact, Earth Planet. Sci. Lett., 262, 438–449.

Ramesh, K., J. D. Hogan, J. Kimberley, and A. Stickle (2014), A review of mechanisms and models for dynamic failure, strength, and fragmentation, Planet. Space Sci., 107, 10–23.

Raymond SN, Quinn T, Lunine JI. 2006. High-resolution simulations of the final assembly of Earth-like planets. I. Terrestrial accretion and dynamics. Icarus 183:265–82.

Rudge JF, Kleine T, Bourdon B. 2010. Broad bounds on Earth’s accretion and core formation constrained by geochemical models. Nat. Geosci. 3:439–43

Saal, A. E., Hauri, E. H., Cascio, M. L., Van Orman, J. A., Rutherford, M. C., & Cooper, R. F. (2008). Volatile content of lunar volcanic glasses and the presence of water in the Moon’s interior. Nature, 454(7201), 192-195.

Salmon, J., and R. M. Canup (2012), Lunar accretion from a Roche-interior fluid disk, Astrophys. J., 760(1), 83.

Schultz, P. H. (1996), Effect of impact angle on vaporization, J. Geophys. Res., 101(E9), 21,117–21,136.

Schultz, P. H., and D. E. Gault (1990), Prolonged global catastrophes from oblique impacts, in Global Catastrophes in Earth History; An Interdisciplinary Conference on Impacts, Volcanism, and Mass Mortality, edited by J. A. Burns and M. S. Matthews, Geol. Soc. Am. Spec. Pap., 247, 239–261.

Sharp ZD, Draper DS. 2013. The chlorine abundance of Earth: implications for a habitable planet. Earth Planet. Sci. Lett. 369:71–77

Shearer CK, Hess PC, Wieczorek MA, Pritchard ME, Parmentier EM, et al. 2006. Thermal and magmatic evolution of the Moon. Rev. Mineral. Geochem. 60:365–518.

Stevenson DJ. 1987. Origin of the Moon—the collision hypothesis. Annu. Rev. Earth Planet. Sci. 15:271–315.

Stickle, A., and P. Schultz (2011), Exploring the role of shear in oblique impacts: A comparison of experimental and numerical results for planar targets, Int. J. Impact Eng., 38(6), 527–534.

Taylor, G. J., and M. A. Wieczorek (2014), Lunar bulk chemical composition: A post-gravity recovery and interior laboratory reassessment, Phil. Trans. R. Soc. A, 372, 20130242.

Taylor, S. R., G. J. Taylor, and L. A. Taylor (2006), The moon: A Taylor perspective, Geochim. Cosmochim. Acta, 70, 5904–5918.

Ward, W. R. (2012), On the vertical structure of the protolunar disk, Astrophys. J., 744, 140

Warren, P. H. (1985), The magma ocean concept and lunar evolution, Ann. Rev. Earth Planet. Sci., 13, 201–240.

WetherillGW.1985. Occurrence of giant impacts during the growth of the terrestrial planets. Science 228:877–79.

Wiechert U, Halliday AN, Lee D-C, Snyder GA, Taylor LA, Rumble D. 2001. Oxygen isotopes and the Moon-forming giant impact. Science 294:345–48

Williams, J. G., et al. (2014), Lunar interior properties from the grail mission, J. Geophys. Res. Planets, 119(7), 1546–1578

Wolf, R., and E. Anders (1980), Moon and Earth: Compositional differences inferred from siderophiles, volatiles and alkalides in basalts, Geochim. Cosmochim. Acta., 44, 2111–2124.

Wood, J. A. (1986), Moon over Mauna Loa—A review of hypotheses of formation of Earth's Moon, in Origin of the Moon; Proceedings of the Conference, vol. 1984, pp. 17–55, Lunar and Planetary Institute, Kona, Hawaii.

Zhang, J., Dauphas, N., Davis, A. M., Leya, I., & Fedkin, A. (2012). The proto-Earth as a significant source of lunar material. Nature Geoscience, 5(4), 251-255.

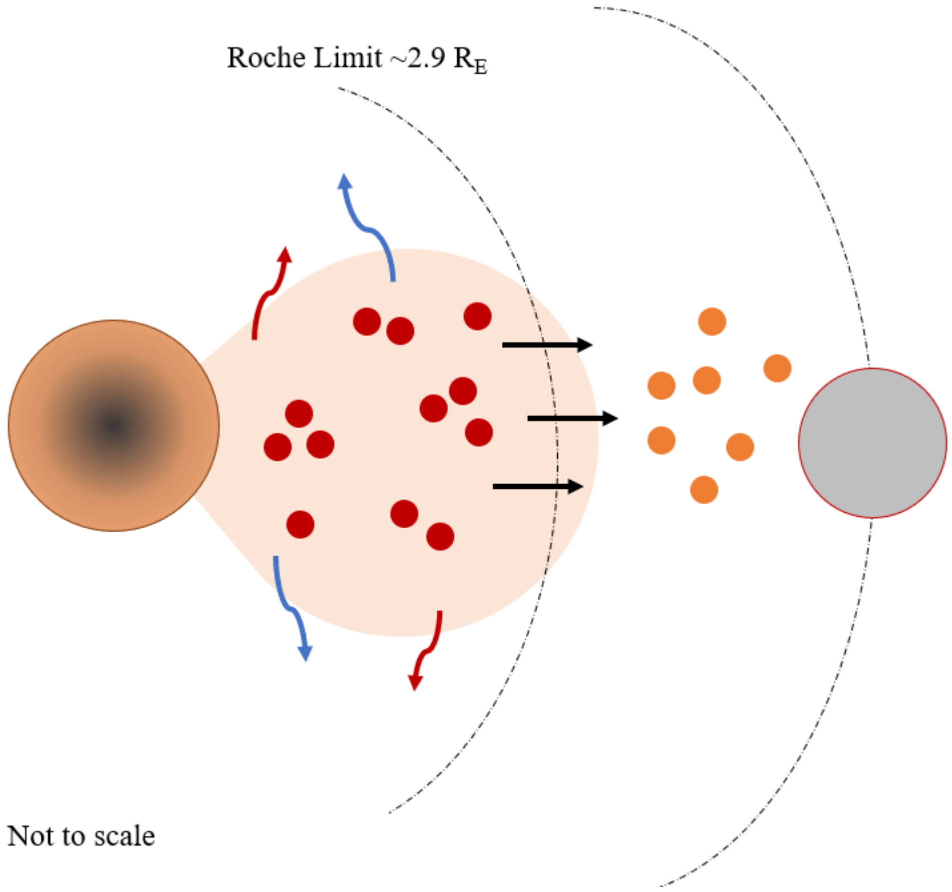

Figure 1: Illustration of the processes at work in the post-impact-generated disk. Inside the Roche limit, tidal forces prevent accretion. The disk contains clumps of silicate melt and condensed vapor (red circles) and vaporized clumps (light red cloud). The disk over time cools from radiation (red arrows) and loses water (blue arrows). The disk spreads diffusively, toward the proto-Earth, and beyond the Roche limit, where the disk fragments into clumps (orange circles) to accrete onto the Moon (grey circle). Adapted from Barr (2016).

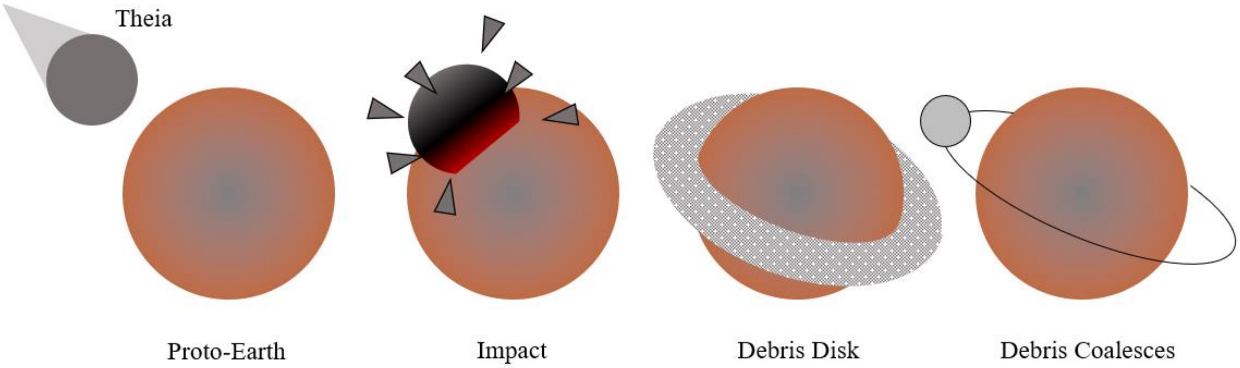

Figure 2: Probable sequence of events involving the collision of Theia, a Mars-sized object, creating a diffuse debris disk around a proto-Earth and eventually coalescing into a natural satellite.