ANNUAL REVIEWS

Annual Review of Astronomy and Astrophysics

Transneptunian Space

Brett Gladman[1] and Kathryn Volk[2]

[1]Department of Physics and Astronomy, University of British Columbia, Vancouver, V6T 1Z1, Canada; email: gladman@astro.ubc.ca

[2]Lunar and Planetary Laboratory, The University of Arizona, Tucson, Arizona 85721, USA

Annu. Rev. Astron. Astrophys. 2021. 59:203–46

The *Annual Review of Astronomy and Astrophysics* is online at astro.annualreviews.org

https://doi.org/10.1146/annurev-astro-120920-010005

Keywords

transneptunian objects, small Solar System bodies, planetary system formation, Kuiper Belt, protoplanetary disks, circumstellar disks

Abstract

We provide a nonspecialist overview of the current state of understanding of the structure and origin of our Solar System's transneptunian region (often called the Kuiper Belt), highlighting perspectives on planetesimal formation, planet migration, and the contextual relationship with protoplanetary disks. We review the dynamical features of the transneptunian populations and their associated differences in physical properties. We describe aspects of our knowledge that have advanced in the past two decades and then move on to current issues of research interest (which thus still have unclear resolution).

- The current transneptunian population consists of both implanted and primordial objects.
- The primordial (aka cold) population is a largely unaltered remnant of the population that formed in situ.
- The reason for the primordial cold population's current outer edge is unexplained.
- The large semimajor-axis population now dynamically detached from Neptune is critical for understanding the Solar System's history.
- Observational constraints on the number and orbits of distant objects remain poor.

Contents

1. INTRODUCTION

The outer regions of our Solar System (beyond Neptune) have gone from a state of mostly theoretical speculation more than 30 years ago to one whose details are emerging. At the time of the last Annual Reviews article on this broad topic (Luu & Jewitt 2002), only the coarsest elements of structure of the transneptunian region were known. There were ~400 observed objects in the region then, many of which had poorly known orbits, compared with >2,000 today, ~1,300 of which have well-determined orbits. A thorough literature review of the past 20 years is impossible; this review is targeted at both beginning graduate students in the field and nonexpert astronomers and planetary scientists in other research domains. We often refer the reader to other recent reviews for more detailed treatments of many topics and, thus, are providing a high-level review of reviews. We preferentially cite foundational papers, other reviews, and some of the most recent results connected to future directions. Our goal is to give an introductory, big picture overview with enough references that the interested reader can then enter the vast literature on any subtopic of their choosing.

We aim to cover current observational and theoretical knowledge of transneptunian objects (TNOs hereafter), focusing on understanding the various populations and ideas about their origins. (See the sidebar titled We Will Guide the Reader Through Five Themes.) We note that the transneptunian populations have often been referred to as the Kuiper Belt or the Edgeworth–Kuiper Belt (for historical reasons; reviewed by Fernández 2020); the use of the term transneptunian has become more prevalent in the past decade, and we use it throughout. Sections 2 and 3 provide brief overviews of the observational and dynamical concepts required to understand the state of TNO science. Section 4 describes our knowledge of the subpopulations of TNOs, highlighting physical and orbital aspects of these populations that can be linked to early planetesimal formation in the Solar System as well as the era of giant planet migration. Section 5 discusses current active areas of research into how TNOs constrain the Solar System's dynamical history

WE WILL GUIDE THE READER THROUGH FIVE THEMES

- Basic observational aspects for outer Solar System small bodies.
- Basic dynamical processes that affect the orbits of small bodies.
- What knowledge about the transneptunian region do nonexperts need to understand?
- What are current topics of hot research (with still unclear answers)?
- What major topics are likely amenable to exploration in the next decade?

and the planetesimal formation process. Section 6 provides some context with recent advances in studies of protoplanetary disks around other stars and discusses the future of TNO science.

2. OBSERVATIONAL ASPECTS

This review focuses on the structure and cosmogonic implications of the transneptunian populations. However, it is important to understand how these are detected to appreciate potential dangers in interpretations drawn from the available biased TNO sample.

Almost all known objects beyond Neptune are discovered by detecting their motion against the background stars and galaxies, using multiple exposures to see a moving point source in reflected solar light. This is extremely challenging because of the $1/d^4$ nature of reflected light, making a $D = 100$ km diameter object roughly twenty-fifth magnitude at heliocentric distance $d = 40$ AU and thirty-fifth magnitude at 400 AU (i.e., fainter than known galaxies at the edge of the visible Universe). It is thus no surprise that the statistical exploration of the inner edges of transneptunian space had to wait for CCD detectors mounted on ≥4-m telescopes, and that deeper surveys steadily continue to uncover new aspects of this region.

Intrinsic TNO angular motions for near-circular orbits of semimajor axis a are only $360\ \mathrm{deg}/a_{\mathrm{AU}}^{3/2} \simeq 0.6\ (40/a_{\mathrm{AU}})^{3/2}$ arcsec h^{-1} (as seen from the Sun); this is usually dwarfed by the opposing retrograde (i.e., parallactic) motion induced by Earth's movement. At opposition, a TNO's distance is $d \gg 1$ AU, and this reflex rate is $\simeq 150/d_{\mathrm{AU}}$ arcsec h^{-1}. With exposures spaced far enough apart in time (allowing movement of at least the seeing FWHM), searches use this position change to detect TNOs. The dominance of the reflex component when detecting motion near opposition gives an immediate estimate of the object's distance, though it is imperfect because orbital eccentricity, e, changes the intrinsic component. For moderate e ($\lesssim 0.3$), the determination of d and orbital inclination i converges quickly. However, because it is difficult to constrain the line-of-sight velocity component using projected sky motion, the orbit's a, e combination converges slowly, and it usually takes at least a year before better-than-order-unity confidence can be placed in a and thus e (see, e.g., Jones et al. 2010 for a full discussion of this). Historically, this has resulted in preferential recovery of low-e orbits and the biasing of the detected catalog against large-a orbits.

Large-a, large-e TNOs are usually discovered at distances only slightly above their perihelion distance $q \equiv a(1 - e)$, and their initial orbit fits often have an assumed low-e orbit with $a \simeq d$; without careful tracking via prompt follow-up observations, these objects are easily lost because their predicted and actual sky positions diverge significantly over time. Note that the reflex component's dominance in opposition surveys means there is no bias against detecting the motion of TNOs in the search field with large orbital inclination (even heliocentric orbits going backward with $i \simeq 180$ deg), because they still move to lower right ascension on the sky, i.e., the same direction as direct TNOs at opposition. Such $i \simeq 180$ deg TNOs would be thought to be

closer initially, but if tracked within the first few months the d, and thus i, would be ascertained. Kavelaars et al. (2008) gives a thorough discussion of all types of observing biases, especially the unusual concept of ephemeris bias (the preferential nonrecovery of objects with unusual orbits). An orbit of sufficient quality to do a dynamical classification requires plentiful observations over at least a year's baseline for low-e TNOs. For large-e orbits, this takes several years, and even TNOs with more than a decade of observed arc (which have high-precision on-sky ephemerides) often still have a-uncertainties larger than typical resonance widths of $\sim$1 AU (see Section 3).

Outer Solar System surveys have a range of (usually inversely correlated) depths and areal coverages; past surveys (reviewed by Bannister 2020) have detected objects with a rough magnitude range of 14–28. There are few bright objects, as they are intrinsically rare, and few extremely faint TNOs because they are beyond the reach of all but the deepest and narrowest surveys. The vast majority of known TNOs have r-band apparent magnitudes of $m_r = 21$–25. Most detections have occurred close to the Solar System plane; because inclined orbits spend little time near the midplane during their vertical latitude oscillation, there is an increasingly large bias against TNOs as they approach edge-on orbits ($i \rightarrow 90$ deg from above or below). Correcting for this bias (e.g., Petit et al. 2017) indicates that the high-i TNOs are underrepresented in the observed sample.

The IAU (International Astronomical Union) Minor Planet Center (MPC) maintains the catalog of Solar System small bodies. When a discovered TNO is submitted to the MPC, it is uniquely identified with a provisional designation. If and when more observations result in the predicted future uncertainties in the sky position becoming very low, it receives a number; numbered objects are sometimes given names. We provide all designations for specific objects when they are first mentioned in this review to avoid confusion, because in older literature only the provisional designations were available (e.g., the recent *New Horizons* spacecraft flyby target is referred to as 486948, also known as Arrokoth or 2014 MU69; early papers would only have its provisional designation, 2014 MU69).

The inherent faintness of most TNOs means that much of our knowledge of TNO surface properties is based on broadband colors (TNO colors are measured as the magnitude difference between two filters; see, e.g., Schwamb et al. 2018 for a recent color survey). Spectra of moving objects with $m_r > 21$ are challenging to obtain, but there have been spectral studies of the largest and brightest dwarf-planet-sized TNOs (see, e.g., Pinilla-Alonso et al. 2020 for discussion of dwarf planets). These spectral studies reveal the presence of a variety of ices (via absorption bands) on the surfaces of many objects, whereas other objects' spectra are relatively featureless. The interested reader is referred to the reviews by Brown (2012) and Barucci & Merlin (2020), who discuss TNO surface properties in the context of both color and spectral studies.

Because of the large distance range of observed TNOs and the d^{-4} effect on their brightness, planetary astronomers often use the Solar System version of absolute magnitude H, rather than apparent magnitude, as a measurable for TNOs. H is the apparent magnitude (in a specified filter) an object would have in an equilateral triangle Sun–Earth–TNO configuration, but at full phase (note that this viewing geometry is impossible; it is simply the convention for calculating H). In the Sloan r filter at opposition distance d in astronomical units, $H_r \simeq m_r - 2.5 \log_{10} d^4$, which gives an offset $m_r - H_r \simeq 16$ at 40 AU. H_r is related to a spherical effective diameter D, in kilometers, via

$$H_r = m_{\odot,r} - 2.5 \log \left[p_r \left(\frac{D_{km}}{6 \times 10^8} \right)^2 \right] \simeq 9.0 - 2.5 \log \left[\left(\frac{p_r}{0.16} \right) \left(\frac{D}{100 \text{ km}} \right)^2 \right], \qquad 1.$$

where $m_{\odot,r} \simeq -26.9$ is the Sun's apparent magnitude and p_r is the TNO's optical albedo in that filter. For most TNOs, p_r lies in the factor of four range around 0.16 (0.04–0.64), thus introducing a factor of two diameter uncertainty for individual objects with unknown albedos. Notice that

essentially it is the product pD^2 that is observationally determined. Some TNO albedos have been measured by combining thermal and optical observations (reviewed by Stansberry et al. 2008 and Pinilla-Alonso et al. 2020), allowing improved diameter estimates for those TNOs and providing a basis for using albedo assumptions to estimate diameters of others. Direct measurements of the diameters of a small number of TNOs have been made by observing stellar occultations (Section 6.3).

The TNO size distribution is a topic of intense interest as it informs us about the formation and subsequent collisional evolution of these objects (see, e.g., Morbidelli & Nesvorný 2020). Given how few TNOs have direct size measurements, one typically discusses their directly measured H magnitude distribution rather than their size distribution. TNO numbers increase roughly exponentially with magnitude and are often modeled over some restricted magnitude range with a cumulative distribution,

$$N(<H) \propto 10^{\alpha H}, \qquad 2.$$

where α is the distribution's logarithmic slope. In reality, the distribution is more complicated, with different slopes derived over different H ranges; these have been modeled as either broken or rolling power laws or as divots with discontinuous transitions (see Bernstein et al. 2004, Shankman et al. 2013, and Fraser et al. 2014 for examples). For a broken power law H distribution that has a steep ($\alpha > 0.6$) slope for large (low-H) objects that breaks at $D_k = 20$–50 km to a plausible equilibrium collisional cascade of $\alpha \simeq 0.5$ at smaller sizes (Kenyon & Bromley 2020), Gladman et al. (2001) estimated a mass of 0.04–0.1 $M_\oplus$ in the 30–50-AU belt with weak model dependencies; a recent mass measurement derived from Cassini spacecraft radio tracking data is 0.061 $M_\oplus$ (Di Ruscio et al. 2020). Because of the size distribution, the majority of the mass is in objects near D_k; in terms of numbers, there are $\sim 10^5$ TNOs with $D > 100$ km out to 50 AU (Petit et al. 2011).

In summary, almost all TNOs have been discovered in optical surveys with significant biases; orbits with large a and e or with large i are underrepresented in the observed sample due to a mixture of detection biases and ephemeris biases. The extreme faintness of TNOs in reflected light means the sample is restricted to relatively bright (i.e., large) objects for which we generally only know the H magnitude. Photometric colors have been measured for a small subset of TNOs and represent our most abundant information about their surface properties. Sizes have been measured for an even smaller subset of known TNOs, so H-magnitude distributions are used as a proxy for size distributions.

3. DYNAMICAL PRIMER

Although a true introduction is well beyond the scope of this review, we provide the reader with some basic dynamical knowledge and point to other sources to acquire a more detailed understanding. We assume that the reader is familiar with the basics of the two-body problem and understands the definitions of semimajor axis a, eccentricity e, perihelion distance $q = a(1 - e)$, and orbital inclination i. We remind the reader that the inclination depends on the chosen reference plane; the longitude of ascending node, Ω, is the angle between the chosen reference x axis in that plane and the location where the orbital path crosses from beneath to above the x–y plane. An orbit's argument of perihelion, ω, is the angle along the orbital plane from the location of Ω to the location of perihelion.

Objects orbiting the Sun do not perfectly return to the exact same position after a single orbital period because the (nearly) perfect $1/r$ potential field of the Sun is broken by the gravitational effects of all the other massive bodies in the Solar System, causing the orbit to precess. For nonresonant orbits of low e and i dominated by the perturbations from a single interior planet of mass

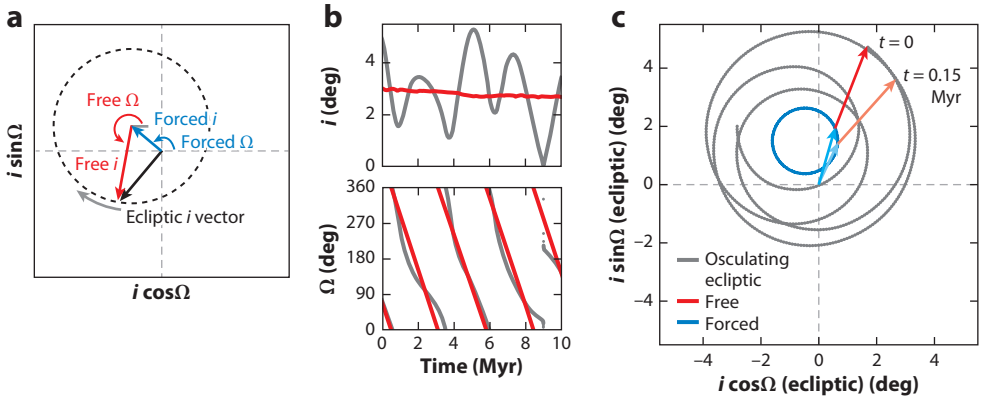

Figure 1

(*a*) Schematic depiction of the free (*red*) and forced (*blue*) inclination vectors. The dashed axes denote the reference plane (e.g., the ecliptic), where the polar distance is i_{ecl}, and the osculating ecliptic node Ω (not labelled) is the polar angle. The positive *x* axis is the reference direction from which all three nodes are measured. For the Sun and a single planet, the forced *i* and Ω values are the planet's orbit in the chosen reference plane. A small body's free *i* and free Ω then regress clockwise around the forced inclination center, following the dotted black circle, which results in oscillations in ecliptic *i* and nonuniform precession of ecliptic Ω. (*b*) Evolution of the osculating ecliptic (*gray*) and free (*red*) *i* and Ω for TNO 2001 QD298, which has a nearly constant barycentric semimajor axis of 42.6 AU. (*c*) The projection of these orbital elements onto the ($i \cos\Omega$, $i \sin\Omega$) inclination vector plane. The free inclination is measured relative to the forced inclination vector (*blue*). The free and forced inclination vectors (and thus also the osculating ecliptic inclination) rotate clockwise over time in the right panel, causing a regression of Ω. The path of the forced inclination vector traces a (*blue*) circle around the Solar System's invariable pole (the total angular momentum vector). An animated version of this figure is available in **Supplemental Video 1**. Abbreviation: TNO, transneptunian object.

Supplemental Material >

m_p and semimajor axis a_p, Ω slowly regresses at a uniform rate,

$$\dot{\Omega} \simeq -\frac{3}{2}\frac{m_p}{M_\odot} n \left(\frac{a_p}{a}\right)^2, \qquad 3.$$

for $a \gg a_p$, $e \ll 1$, and small *i*; here, $n = \sqrt{G M_\odot / a^3} = 2\pi/P$ is the averaged mean motion of an object around its orbit over its orbital period *P*. The orbit's longitude of perihelion $\varpi = \Omega + \omega$ increases at a similar rate; we focus on the *i* and Ω evolution as that is more relevant to later discussion. These kinds of slow (or secular) precessions have periods much longer than the orbital period; Equation 3 shows that this rate for a TNO at twice Neptune's semimajor axis regresses at $\dot{\Omega}/n = 2 \times 10^{-5}$, corresponding to a precession period of 52,000 Neptune orbital periods ($\simeq 9$ Myr).

When a small body is perturbed by a single planet, its orbital plane precesses around the planet's orbital plane, so that its average orbital plane is the planet's. This is most easily demonstrated by examining the evolution of the inclination vector with components ($i \cos\Omega$, $i \sin\Omega$). **Figure 1*a*** shows this simple case: The inclination of the small body measured in a reference frame that is not the perturbing planet's orbital plane (e.g., as is the case when the ecliptic is used as the reference frame) is the sum of a so-called forced *i* and Ω determined by the planet and a free i_{free} and Ω_{free} determined by the small body's initial orbit. The free inclination vector then precesses clockwise around the forced inclination vector at the rate given in Equation 3. The magnitude of *i* in the reference frame then varies over time, whereas the magnitude of i_{free} remains constant. In particular, a TNO's ecliptic *i* and Ω only remain fixed if they match the forced vector (i.e., have

zero $i_{\rm free}$); in all other cases a TNO's ecliptic i varies, but that variation is an artifact of the reference frame choice. The dynamically interesting quantity is $i_{\rm free}$ as it does not depend on the reference frame. Similar analysis can be done for the eccentricity evolution of TNOs, though the forced eccentricities are generally very small. In the real Solar System, TNOs are perturbed significantly by all four giant planets, so the precession behavior is more complex, as shown in **Figure 1*b*,*c***. For this TNO, the forced inclination vector varies over time, largely because Neptune's inclination varies over time as it interacts with the other giant planets, resulting in the complex behavior of the TNO's ecliptic i and Ω; in contrast, $i_{\rm free}$ remains fairly constant and $\Omega_{\rm free}$ regresses smoothly. As a result, dynamical structure related to TNO inclinations is much clearer to interpret using $i_{\rm free}$. As discussed in Section 4.1, inclination cuts are sometimes used to separate TNOs into different categories; using a cut of $i = 4$ deg (a typical value) in **Figure 1** would result in different conclusions for the TNO depending on whether the cut is made in free or ecliptic inclination.

More dramatic orbital evolution can occur when the precession frequency of a TNO approaches one of the fundamental secular frequencies of the Solar System; the mutual gravitational interactions of our four giant planets yield three fundamental inclination frequencies and four eccentricity frequencies (see, e.g., Murray & Dermott 1999). A secular resonance occurs when a TNO's precession for either Ω or ϖ matches one of these frequencies; such resonances result in large increases in the TNO's e or i. For initially low e and i orbits, there are two such secular resonances in the TNO region, the so-called ν_8 eccentricity and ν_{18} inclination secular resonances that occur in the semimajor axis region $a = 40$–42 AU (see, e.g., Chiang & Choi 2008 and Section 4). These are the resonances with the Solar System's slowest secular frequencies; because precession periods increase with increasing a (Equation 3), these are the outermost secular resonances in the Solar System (assuming no other planets exist).

The more widespread TNO resonance type is mean-motion resonances. A mean-motion resonance can occur when the orbital periods of two objects are related by a simple integer ratio, and they appear throughout the Solar System (see, e.g., Gallardo 2006). The most famous resonant TNO is Pluto, in Neptune's 3:2 resonance: Pluto completes two orbits for every three of Neptune, repeating the same relative geometric configuration in physical space, leading to a repeated pattern of gravitational interactions (see, e.g., Cohen & Hubbard 1965, Murray & Dermott 1999). This resonant perturbation leads to stable libration (oscillation) of Pluto's orbit around the exact resonant orbit, in which a TNO would complete two orbits with perihelia located exactly ± 90 deg away from Neptune in longitude. This perihelion confinement relative to Neptune allows objects in the 3:2 resonance to remain stable even if their perihelion distances are below Neptune's semimajor axis. **Figure 2** shows a TNO evolution in another Neptune resonance; in this case the exact 10:3 resonant orbit makes three perihelion passages at ± 60 deg and 180 deg relative to Neptune as Neptune completes ten orbits. In the reference frame that corotates with Neptune, this results in a three-fold symmetry for the resonant orbit. Over the TNO's full resonant cycle, the location of perihelia relative to Neptune in this rotating frame librates back and forth. As this libration occurs, the TNO's semimajor axis and eccentricity also vary (**Figure 2*a*, subpanels *i*,*ii***). The amplitude of the semimajor axis variations depends on the eccentricity and how much the perihelion location librates. At any given eccentricity, there is a maximum allowed variation in semimajor axis within the resonance, and this is referred to as the width of the resonance; widths for Neptune's resonances at typical TNO eccentricities are $\sim$1 AU (see, e.g., Lan & Malhotra 2019). The angular libration in the perihelion location (as seen from the Sun) is related to the oscillation of a resonant angle, ϕ, which can be derived from the orbital element evolution of the TNO and Neptune in numerical integrations. Gladman et al. (2012) provides an accessible introduction to how ϕ is calculated

Secular resonance: when a TNO's orbital precession rate matches a forcing frequency set by the planets, leading to increases in e and/or i

Mean-motion resonance: when two objects' orbital periods are a simple integer ratio and their gravitational interaction results in a repeating, resonant pattern

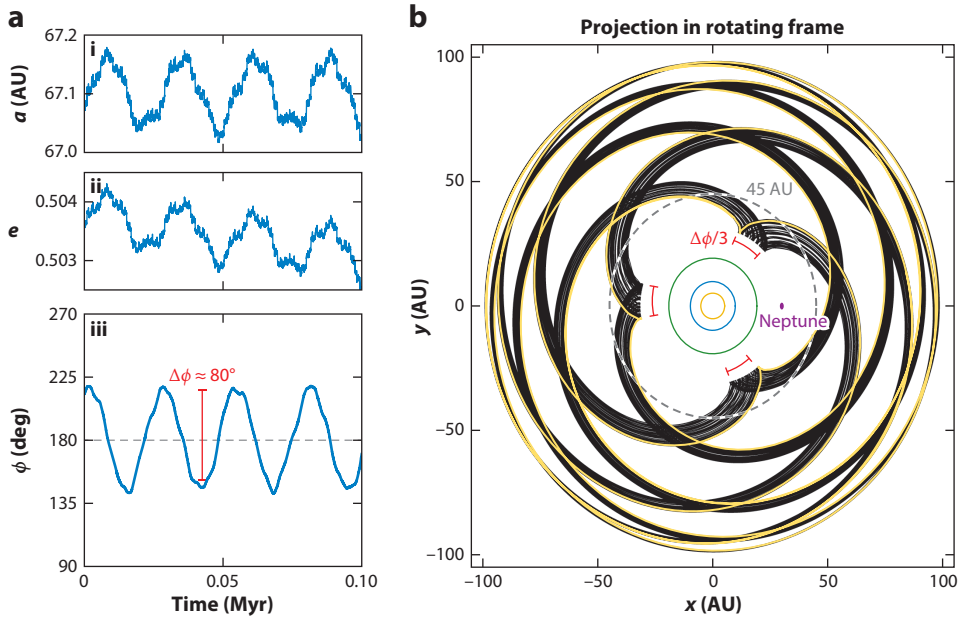

Figure 2

Orbital evolution of TNO 225088 (also known as Gonggong or 2007 OR10) over 10^5 years in Neptune's 10:3 mean-motion resonance. (*a*) The time evolution of a, e, and the resonant angle ϕ; note that all three show the influence of the sinusoidal resonant perturbation. (*b*) Gonggong's position projected on a reference x–y plane that rotates around the Solar System's barycenter at the rate of Neptune's mean motion; Neptune thus remains nearly fixed along the x axis (*purple point*), whereas Gonggong follows the black trajectory (the other giant planet paths are shown, interior to Neptune). Gonggong completes three orbits (an example of three orbits is shown in *gold*) for every ten Neptune orbits, creating the three-fold symmetry in the rotating frame. The locations of Gonggong's perihelia in the rotating frame librate back and forth over the full resonant cycle described by ϕ. The $\Delta\phi \approx 80$ deg libration amplitude corresponds to an angular oscillation of the perihelion location in the rotating frame of $\Delta\phi/3 \approx 27$ deg (both labeled in *red*). We note that the majority of TNO detections occur at distances $\lesssim$45 AU (*dashed gray circle*) due to the flux bias (see Section 2); for resonant TNOs, this results in detection preferentially at specific longitudes relative to Neptune. An animated version of this figure is available in **Supplemental Video 2**. Abbreviation: TNO, transneptunian object.

Supplemental Material >

and other examples of how resonant TNOs are distributed relative to Neptune. The dynamics of different resonances constrain the TNO perihelion locations to varying specific locations relative to Neptune. Combined with the tendency for TNO detection at perihelion in given survey fields in specific directions relative to Neptune, this makes the biases (see Section 2 and **Supplemental Video 2**) in the observed resonant TNO sample particularly complicated.

Kozai effect: inside mean-motion resonances, coupled q and i oscillations associated with the TNO's perihelion angle relative to the reference plane

The dynamics of Neptune's resonances change TNO orbital precession rates. Although non-resonant TNOs typically experience opposite signs for Ω and ϖ, TNOs in mean-motion resonances can experience matching or nearly matching precession rates for these angles, which can lead to fixed or librating values of the argument of perihelion. When ω librates, the variations in e and i can become large and are anticorrelated (meaning that variations in q and i are coupled because a is fixed). This phenomenon has been called the Kozai effect within mean-motion resonances because the resulting dynamics is similar, but not identical, to the secular Kozai resonance described outside of mean-motion resonances (see Malhotra et al. 2016); due to the mean-motion

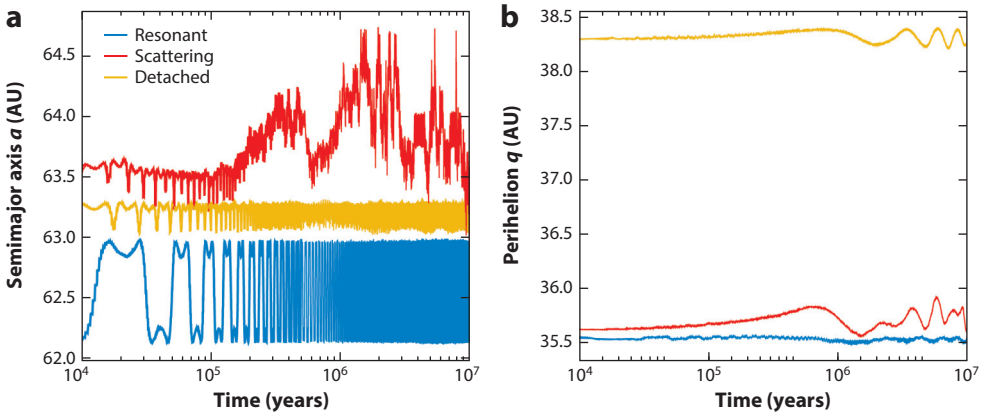

Figure 3

Time evolution of (*a*) semimajor axis and (*b*) perihelion distance for three real TNOs. Note the logarithmic axis used to clarify different variation timescales. The scattering (*red*) and resonant (*blue*) TNOs start at very similar a and q but have very different evolution; the resonant TNO librates around the center of the 3:1 resonance at $a \approx 62.5$ AU, whereas the scattering TNO experiences a random walk in orbital energy due to encounters with Neptune. The detached TNO (*gold*) has a similar a to the other two TNOs, but its larger q means it does not experience strong perturbations from Neptune and remains at a relatively fixed value of a. Abbreviation: TNO, transneptunian object.

resonant dynamics, the Kozai effect within Neptune's resonances can occur at smaller i values than the nonresonant case (see, e.g., Morbidelli et al. 1995, Lawler & Gladman 2013).

TNOs that have low perihelion distances ($q \lesssim 37$–38 AU) and are not phase protected by mean-motion resonances can experience significant gravitational interactions with Neptune that change their specific orbital energy $E = -\frac{1}{2} G M_\odot / a$ (and thus semimajor axis). These scattering events result in a random walk in semimajor axis over time (e.g., Duncan et al. 1987). **Figure 3** demonstrates this scattering evolution, which contrasts with a nearby resonant TNO and a TNO with a q value that is 3 AU larger. The resonant TNO experiences larger periodic changes in a, corresponding to its librations in resonance (as in **Figure 2**). The nonresonant, large-q TNO experiences only minimal changes in a; this kind of orbital evolution is dynamically detached from interactions with the giant planets aside from secular evolution (which cause the small, megayear-period variations in q). Note that the scattering and resonant TNOs in **Figure 3** have very similar initial a and q. The large number of potential Neptune resonances and their complicated phase space mean that numerical integration is the only secure way to determine if a known TNO is in a mean-motion resonance; accurate orbital classification (discussed in Section 4) thus requires integration and cannot be done with simple orbital element cuts.

A dynamical wrinkle important for TNOs is the preference for using barycentric orbital elements rather than heliocentric elements. For inner Solar System orbits, the rapid orbital periods and solar proximity result in the Sun being the natural origin for the heliocentric position and velocity vector, which are then transformed to osculating heliocentric orbital elements. However, using the Sun as the origin makes little sense for a TNO; at large distances they dominantly see the Solar System as a single mass at the barycenter (center of mass) of the planetary system, and the TNO's position and velocity vector should be measured from that location to be converted to barycentric orbital elements (which then describe a roughly constant elliptical orbit). If the barycenter is not used, a TNO's orbital elements will vary on timescales of each giant planet's orbital period due to their gravitational influence on the Sun. If we consider just the most massive

planet, Jupiter, then a near-circular TNO's heliocentric semimajor axis varies with amplitude,

$$\frac{\delta a}{a} = 2\frac{m_{\rm J}}{{\rm M}_\odot}\sqrt{\frac{a}{a_{\rm J}}}, \qquad 4.$$

over Jupiter's 12-year orbital period, and thus the TNO's heliocentric elements depend sensitively on the choice of reference date. Even at 50 AU, a TNO's heliocentric a oscillates by ± 0.3 AU due to Jupiter; this oscillation reaches ± 10 AU by $a \simeq 500$ AU. This amplitude is comparable with or many times larger than typical resonance widths, and thus heliocentric elements cannot be used to judge resonance occupancy; the oscillation can also affect perceived perihelion distances. This variation is entirely a reference frame problem and thus TNO orbital elements should be given in the barycentric frame, as we do throughout this article. (Heliocentric orbits are useful on short timescales to predict TNO ephemerides, so they are still commonly used for that purpose. Care is thus required when tabulating orbits.)

4. THINGS A NONEXPERT NEEDS TO UNDERSTAND THAT ARE KNOWN

We take transneptunian space to encompass TNOs with orbital semimajor axes $a > 30.1$ AU (that of Neptune). This in principle includes the Oort cloud, which merits some discussion but is not a focus of this review. We take the position that dynamical classifications should be based on dynamics in the current environment (Gladman et al. 2008), and thus the Kuiper Belt ends and the inner Oort cloud begins at $a = 2{,}000$ AU, which is where the current galactic environment starts to dominate the dynamics by being capable of altering the perihelion distance q and orbital inclination i on orbital timescales; the outer Oort cloud begins at $a = 10{,}000$ AU, where the tide sphericalizes the structure (see Duncan et al. 1987, Dones et al. 2004). It is possible that these boundaries were different in the early Solar System due to the Sun's birth environment (see Section 5.2). For most of this review we concentrate on the $30.1 < a < 2{,}000$ AU range.

Much of the theoretical and observational history of transneptunian space centered on the issues of comet supply and the nature of Pluto (which remained the only noncometary TNO known for 60 years). Fernández (2020) recently reviewed this history and provides references. Briefly, the existence of a transneptunian population was hypothesized because (*a*) cosmogonic arguments by Leonard, Edgeworth, and Kuiper argued that there was no reason objects would not exist on orbits beyond the orbit of Neptune, as the protoplanetary disk had no reason to end at 30 AU (an edge we return to in Section 6.2); (*b*) the relatively low-inclination short-period, i.e., Jupiter-family, comets needed to come from a flattened source beyond Neptune (i.e., not the Oort cloud); and (*c*) How could Pluto be alone in this region?

The current orbital distributions of transneptunian space are shown in **Figure 4**, which divides TNOs into four main dynamical classes defined in **Table 1**. The current majority consensus is that the TNO population is not just a simple remnant of objects that formed in the protoplanetary disk beyond 30 AU but rather a complicated superposition of some primordial objects immersed in a sea of TNOs that formed elsewhere (often postulated as closer to the Sun) and then transferred to their current orbits during the epoch of giant planet migration. It was realized that orbital inclination served as a very rough proxy to separate these two populations. The relocated objects appear at essentially all semimajor axes and have an inclination distribution that is much hotter dynamically (with the majority having $i > 10$ deg and often $e > 0.2$, leading to larger radial and vertical oscillations over an orbit and hence the term hot). The cold component is confined near the averaged plane of the giant planets (i mostly only a few degrees) and concentrated in $a = 42.5$–47 AU (see Section 4.1). These two populations are mixed radially and vertically, though the cold

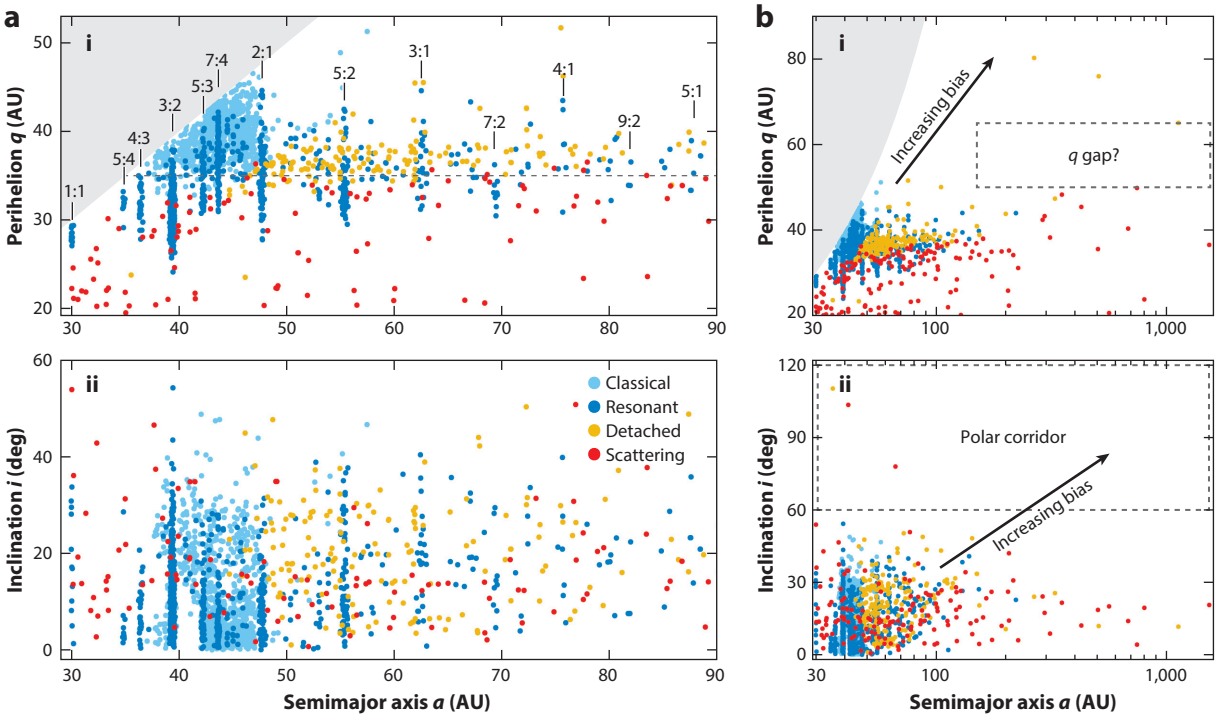

Figure 4

Known TNOs with dynamical classifications and perihelia beyond Uranus ($q > 19.2$ AU), in J2000 barycentric orbital elements; this includes the OSSOS ensemble data (Bannister et al. 2018) and ~1,000 other TNOs with orbits sufficiently good to allow us to determine a reasonable dynamical classification (color coded). The shaded gray regions are nonphysical orbits with $q > a$. (*a*) The $a = 30$–90 AU region, where all but three (see panel *b*, subpanel *ii* and Polar Corridor discussion in Section 6.1.1) known objects with high-precision orbits have $q < 50$ AU and $i < 60$ deg; panel *a*, subpanel *i* identifies selected mean-motion resonances with Neptune. The horizontal $q = 35$ AU line is for comparison with later figures. The obvious 36–42 AU void in the classical *a*/*i* distribution is due to the ν_8 secular resonance. A very high density of classical objects exists at low *i* from 42–47 AU (obscured by many overlapping *blue points*); the inclination distribution in this region is bimodal, with a cold and hot component. (*b*) An expanded view (logarithmic scale) of the sample from $a = 30$–1,600 AU. The trend of detection bias (*arrows*) against large *a*, large *q*, and large *i* are important (Section 2); for example, despite decreasing numbers of known TNOs in large *a* resonances (panel *a*, subpanel *i*), the intrinsic populations of the 2:1, 5:2, and 5:1 resonances are similar to the 3:2 resonance's intrinsic population (Section 4.2.2). A proposed *q* gap (*box* in panel *b*, subpanel *i*) for $a > 150$ AU TNOs is discussed in Section 5.3. Abbreviations: OSSOS, Outer Solar System Origins Survey; TNO, transneptunian object.

population's radial and vertical range is more limited. A similar, perhaps familiar, example of this is the thin disk, thick disk, and halo stellar components of the galaxy; near the galactic midplane, stars could be from any of these components, but at high latitude and large galactocentric distance the fraction from disk components drops off rapidly.

TNOs have a more intricate dynamical structure than the asteroid belt. This is partly because the asteroid belt is bounded by planets on both sides (and low-mass Mars is still gravitationally effective on 4-Gyr timescales due to the small volume and rapid orbital periods), whereas the transneptunian populations are unbounded on the outside (as far as we know with certainty), allowing resonances with Neptune to trap particles at large semimajor axes and preserve them (see Section 4.2.2). However, though secular and mean-motion resonances sculpt the population in both cases, in the asteroid belt almost all orbital substructure is due to asteroid families (see the

Table 1 TNO dynamical categories and interpretations

Dynamical classifications	
Scattering	Definition: TNO whose a Neptune can currently alter significantly on timescales $\ll$1 Gyr; typically $q \lesssim$ 38 AU (e.g., **Figure 3**)
	Interpretation: Likely a decaying remnant of a much larger primordial scattering population (see Section 4.2.1)
Resonant	Definition: TNO in a mean-motion resonance with Neptune (e.g., **Figure 2**)
	Interpretation: The abundant resonant TNOs were likely captured during the epoch of giant planet migration (see Sections 4.2.2 and 5.1)
Detached	Definition: Nonresonant TNO with $a > 47.4$ AU and $e > 0.24$ that is not scattering today (e.g., **Figure 3**)
	Interpretation: The formation mechanism(s) of this population remains an area of active research (see Sections 4.2.4 and 5.2)
Classical	Definition: A TNO that falls into none of the above categories. Divided into subcategories:
	Main Belt: a between the 3:2 and 2:1 resonances ($39.4 < a < 47.7$ AU; e.g., **Figure 5**)
	Interpretation: Observed to have a bimodal i distribution; likely a combination of TNOs that formed in place and others implanted from elsewhere (see Sections 4.1 and 4.2.3)
	Inner Belt: a between Neptune and the 3:2 resonance ($30.1 < a < 39.4$ AU)
	Interpretation: Observed to only have a hot population (see below); either a dynamically excited remnant of the original planetesimal disk or an implanted population
	Outer Belt: a beyond the 2:1 resonance ($a > 47.4$ AU) and $e < 0.24$
	Interpretation: Only a few known. Origin unclear.
Cold versus hot populations	
Cold population	Observed concentration of low-i ($i_{\rm free} < 4$ deg), low-e main-belt TNOs from $a = 42.5$–47.5 AU; likely formed beyond 30 AU and survived in place with only minor e/i perturbations and collisional evolution (see Section 4.1)
Hot populations	TNOs with large e and/or i orbits (existing in all dynamical classes); likely formed at $a < 30$ AU and scattered out to current locations, with the current population a small remnant of the initially scattered population (see Section 4.2)

review by Nesvorný et al. 2015), whereas in the cold TNO main belt there is fine dynamical structure probably created during the late stages of planet formation and migration.

In summary, before descending into greater detail, the cold population's members likely formed very close to their present location and represent a largely unaltered set of primordial planetesimals from ~45 AU. The various hot populations were transported from their formation region to their current locations during giant planet migration. This introduces considerable uncertainty into attempts to infer the protoplanetary disk properties. The chief puzzles (Sections 5 and 6) include the following: (*a*) Where did hot TNOs (and the giant planets) form? (*b*) What was the surface density distribution of primordial condensed solids? (*c*) How were the hot populations emplaced and how did this process not strongly perturb the cold TNOs?

4.1. Cold Population

The dynamically cold population of TNOs is a set of low-eccentricity and low-inclination objects, the understanding of which has steadily become more nuanced. Initial ideas that all TNOs might be a vestigial belt of bodies from a disk starting at ≈35 AU were challenged by the discovery of resonant orbits with Neptune-crossing e, high-i orbits, and the scattering population (reviewed by Davies et al. 2008 and Fernández 2020). The term classical belt denoted the nonresonant, nonscattering orbits, with the main belt (shown in **Figure 5**) being those between Neptune's 3:2 and 2:1 mean-motion resonances at 39.4 and 47.7 AU, respectively. This led to ideas of a

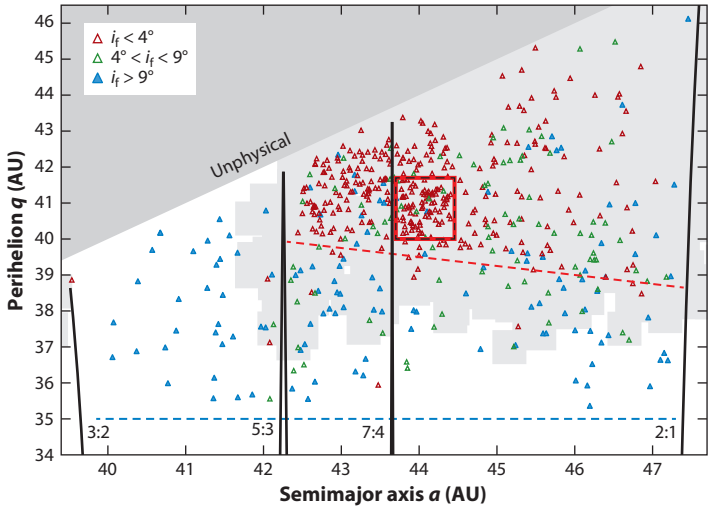

Figure 5

The main belt's detailed structure. We plot the well-characterized OSSOS sample (Bannister et al. 2018), which ensures that the resonant and scattering TNOs are removed to reveal only the classical belt objects. There are no $q < 35$ AU classical TNOs (*blue dashed line*), as they would be scattering. Black curves show approximate resonance boundaries for four major resonances. Dark shading denotes unphysical orbits with $q > a$. Classical TNOs are color coded by their free inclination i_f relative to the local secular forced pole, with cold objects selected with an $i_f < 4$ deg cut (based on Van Laerhoven et al. 2019). We also isolate the $i_f = 4$–9 deg warm population, with a clearly different a, q distribution than the cold and hot. The light gray shading demonstrates a stability map for initially low-i TNOs covering this entire a–q phase space; despite 4 Gyr of stability, the cold population truncates (*red dotted line*) at q values several astronomical units higher. The kernel (*red box*), with many nearly overlapping data points, is briefly addressed in Section 4.1. Abbreviations: OSSOS, Outer Solar System Origins Survey; TNOs, transneptunian objects.

hot–cold mix at all semimajor axes, because the hot population's i distribution extends down to $i = 0$, which precludes perfectly isolating the overlapping populations using solely an inclination cut. It is now understood that, as a population, the cold is only certain to exist between $a = 42.5$ and 47 AU (sometimes called the cold classical Kuiper Belt). The inner classical belt, with $a <$ 39.4 AU, has no cold component even if there are some low-i members (Kavelaars et al. 2009), and nearly all low-i orbits from $a = 39.4$–42.5 AU are removed in a few megayears by the ν_8 secular resonance (Duncan et al. 1995, Chiang & Choi 2008; see Section 3). It's still unclear if the cold population extends beyond the 2:1 resonance (Bannister et al. 2018 lists the few known low-i TNOs beyond the 2:1, but modeling must still be done to confirm they are not the low-i tail of the hot component); if it does, the number per astronomical unit must drop sharply at 48 AU. Note that the outer edge of the classical belt, initially postulated at $a = 50$ AU (Jewitt & Luu 1995, Allen et al. 2001), is more precisely identified with the 2:1 resonance's location for reasons that remain unclear (see Section 6.2.3 and Morbidelli et al. 2008).

Brown (2001) first identified that the classical belt's i distribution has a narrow cold component (of width $\simeq 2$ deg) and a much wider ($\simeq 18$ deg) hot component (hinted at in **Figure 4**). With today's larger TNO sample, it is now necessary to account for planetary perturbations and use free inclinations ($i_{\rm free}$; see Section 3) rather than ecliptic inclinations to separate the cold and hot classical belt objects. Main-belt TNOs with $i_{\rm free} < 4$ deg are almost certainly members of the cold population and those with $i_{\rm free} > 9$ deg belong to the hot population (Van Laerhoven

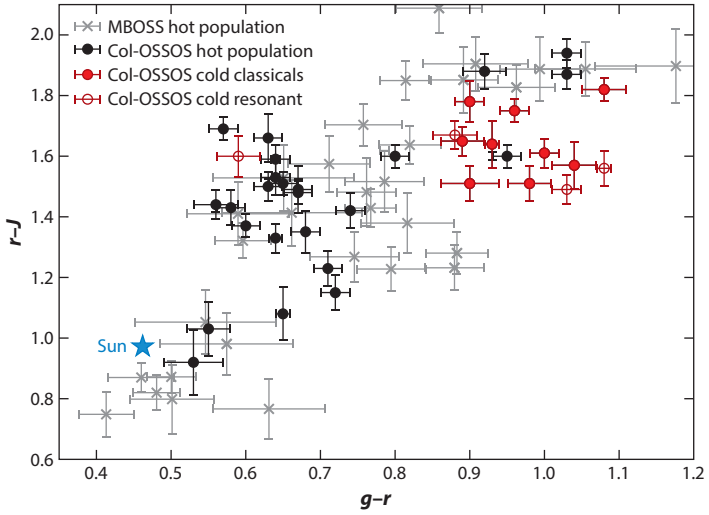

Figure 6

Broadband colors *r–J* versus *g–r* with the cold and hot populations indicated. The Col-OSSOS project's measurements (*circles*; from Schwamb et al. 2019) are shown, where here we defined cold classical (*solid red*) and cold resonant (*open red*) TNOs as those with free inclinations below 4.2 deg; all other objects belong to the hot population (*black* and *gray*). We also plot the highest-precision color measurements (uncertainties <0.1 mag) from the MBOSS database (Hainaut et al. 2012), all of which happen to belong to the hot population (*gray crosses*; we followed the procedure of Schwamb et al. 2019 to convert to Sloan filters); there may be some color selection effects for the MBOSS set. Although both data sets show that the hot population occupies a large range from roughly solar to very red, the cold population is confined to a subset of the color–color space of very red optical *g–r*, with moderate *r–J* colors that are bluer than the hot objects. Abbreviations: Col-OSSOS, Colours of the Outer Solar System Origins Survey; MBOSS, Minor Bodies in the Outer Solar System; TNOs, transneptunian objects.

et al. 2019). These inclination ranges for main-belt TNOs from the Outer Solar System Origins Survey (OSSOS; Bannister et al. 2018) are shown in **Figure 5** along with the TNOs between these free inclination cuts that are sometimes referred to as the warm population. Indications that the cold population is distinct in other ways from the hot population have steadily accumulated. Redder optical colors were noted for low-i (Doressoundiram et al. 2002) and $q > 40$ AU (Tegler & Romanishin 2000) classical TNOs. The H-magnitude distribution of hot TNOs was found to be richer in bright (large) objects (Levison & Stern 2001) and was then realized to differ over larger H ranges (Bernstein et al. 2004, Elliot et al. 2005, Petit et al. 2011, Fraser et al. 2014). Binary TNOs, especially those with comparable sizes, are more frequent in the cold population (reviewed by Noll et al. 2020). Very high-precision colors from reflected optical and near-infrared observations (e.g., Pike et al. 2017, Schwamb et al. 2019) as well as thermal studies (see review by Müller et al. 2020) show the cold objects with $a \simeq 42.5$–47 AU are separable in surface properties from the hot population's members (see **Figure 6**). This is thought to be a preserved signature of the local chemical composition of the solids precipitating out of the nebula at the TNO's formation distance, which likely differed for the cold and hot TNOs (see, e.g., a schematic cartoon of chemical gradations in the early planetesimal disk in Schwamb et al. 2019, their figure 6). Large TNOs experience subsequent surface modifications due to thermal evolution and atmospheric effects, which is amenable to spectral investigation (reviewed by Barucci & Merlin 2020).

The known cold inventory consists mostly of $H_r < 8$ TNOs in the main belt, with an implied debiased population of $\simeq$ 4,000 cold $H_r < 7$ ($D \gtrsim 250$ km) objects (Petit et al. 2011); for perspective, this is about the same as the number of hot TNOs in the main belt for this size range. This surface mass density (projected on the midplane) is far below what was historically thought needed for two-body accretion to form the observed TNOs (Stern 1996), leading to (likely mistaken) ideas that the cold population was collisionally (reviewed by Kenyon et al. 2008) or dynamically (Gladman et al. 2001) depleted by 2–3 orders of magnitude in mass (discussed further below and in Section 5.4).

Survey modeling indicates that there are no cold-population $H_r < 4$ TNOs (roughly, $D >$ 1,000 km for $p_r = 0.16$; Petit et al. 2011). Because of their low-e orbits, these objects would be bright ($m_r < 20$); Schwamb et al. (2014) and Brown et al. (2015) concluded the bright inventory on the sky is nearly complete, especially near the Solar System plane where cold objects reside. (This same flux completeness argument is untrue for $H < 4$ hot objects, which are much fainter when at their larger aphelia; see Section 6.1.) This indicates that, after these largest cold TNOs formed, there was insufficient time given the spatial density of objects for them to increase their mass by accreting other TNOs.

There are now multiple arguments that the cold TNOs were formed in place in a low-surface-density primordial environment and have had little dynamical or physical evolution since. First, the cold classical population is observed to contain many binary TNOs. Measuring the intrinsic binary fraction of TNO populations is difficult owing to the complex biases, but *Hubble Space Telescope* observations confirm that $\approx$20% of cold classicals with $H = 6.15$–8 are binaries and that this fraction is significantly higher than in the hot population; Noll et al. (2020) discuss this measurement and the challenges of observing binary TNOs. The high prevalence and long-term stability of binary pairs argue that the cold binaries could have been neither transported and implanted (Parker & Kavelaars 2010) nor subjected to a very long period of intense bombardment (Parker & Kavelaars 2012). This argues against any significant in situ collisional evolution of the cold population. Second, Lykawka & Mukai (2005) pointed out via numerical integration that cold TNO orbits do not occupy perihelia down to the stability limit at $q \simeq 37$ AU. We independently reproduced this for **Figure 5**, showing which orbits would be occupied today by initially low-i orbits filling the space, which should thus be compared to the $i_{\rm free} < 4$ deg points, with an obvious bound several astronomical units higher (except near the border resonances, which have excited a few cold TNOs to lower q). Thus, either the primordial cold belt was sharply truncated for $q <$ 39 AU and never refilled by subsequent processes (which were, however, somehow able to emplace higher-i orbits there) or, more likely, the q bound shows that the cold belt has only been weakly dynamically heated in e and i after formation and implies little material has been lost. This limit on dynamical excitation has important implications for planet migration models (Section 5.1).

This view has been strengthened by studies of the impact crater record on Pluto and/or Charon (Singer et al. 2019) and the $D \simeq 40$-km cold TNO Arrokoth visited by the *New Horizons* spacecraft (Spencer et al. 2020), believed to be a primordial cold (kernel) TNO (McKinnon et al. 2020). In each case, absolute crater density of the oldest terrains was successfully predicted by models using estimates of the current TNO populations with only mild enhancements over the past 4 Gyr to account for gradual dynamical erosion to today (Greenstreet et al. 2015, 2019). This prediction success, especially in the case of Arrokoth, implies that the size distribution of the cold belt is a preserved relic of its initial state and the cold belt has thus always been of low surface density; we return to this topic in Section 5.4.

In summary, the scenario best explaining current knowledge is that the cold population formed at its present location in an environment of low solid surface density, forming with a maximum size cutoff ($H_r \simeq 4$), rich in binaries, and out of condensed material that produced

THE MAIN-BELT KERNEL

The cold population contains a dense concentration of TNOs spanning $a \simeq 43.8$–44.4 AU and $q \simeq 40$–41.7 AU (see **Figure 5**). Petit et al. (2011) termed this concentration the kernel and quantitatively showed that it contained ~1,000 $D > 200$ km TNOs, about one-quarter of the entire cold-belt population. This structure's origin is puzzling but likely primordial. A collisional family origin is problematic (Chiang et al. 2003), because the parent body would have to be significantly larger than Pluto or Eris and the kernel's velocity dispersion is much less than the escape speed of such a parent body. Petit et al. (2011) posited that cold TNOs swept up into the migrating 2:1 resonance could be released in an $a \simeq 44$-AU clump if Neptune jumped outward (when at $a_N \simeq 28$ AU) by ≃0.5 AU; Nesvorný (2015a) demonstrated this explicitly, motivating the jump as due to scattering inward a fifth hypothetical ice giant planet. This scenario has some issues; as examples, it would generate a $q = 37$–39 AU cold population that is absent (**Figure 5**) and predicts that 2:1 resonant TNOs would outnumber those in the 3:2 (the opposite of reality; Gladman et al. 2012). The kernel is an intriguing primordial feature of the cold population that remains to be fit into a larger cosmogonic scenario.

characteristically red surfaces. This population has suffered little collisional evolution (and thus little modification of its size distribution) down to at least $D \sim 30$ km, and its dynamical excitation in e and i was very limited. In this picture, dynamical structures we see today date to the formation epoch and/or the period of perturbation during the implantation of the hot populations. Any cold population with $a < 35$ AU was destroyed, although a small fraction must still be present because when these objects coupled to Neptune they would have been mixed in with the hot population. The obvious drop in the number per astronomical unit of cold objects at $a = 45$ AU (**Figure 5** and Kavelaars et al. 2009) and the lower envelope in the cold population's q distribution must be a signature of formation and/or planetary migration. Lastly, there is an obvious clump near $a = 44$ and $q = 41$ AU (**Figure 5**, but also long visible in the independent MPC orbital database) called the kernel (see the sidebar titled The Main-Belt Kernel).

4.2. Hot Populations

The dynamically hot populations, with their wide inclination and eccentricity ranges, outnumber the cold population (see, e.g., Petit et al. 2011) and span the full TNO semimajor axis range from Neptune Trojans at 30.1 AU, through the main-belt region, and into the distant scattering and detached populations extending out to many hundreds of astronomical units. **Figure 4** divided TNOs into the different dynamical classes described by Gladman et al. (2008; summarized in **Table 1**), which include the following: the scattering TNOs (Section 4.2.1), whose orbits evolve on relatively short timescales due to gravitational interactions with Neptune; resonant TNOs (Section 4.2.2) that are phase protected from scattering and can sometimes thus survive at $q < 35$ AU; hot classical TNOs (Section 4.2.3), in which the hot population overlaps with cold TNOs in semimajor axis; and the detached TNOs (Section 4.2.4), with an a range similar to the scattering population but with perihelia high enough such that their orbits are much more stable against Neptune perturbations.

As discussed in Section 4.1, the dynamically hot and cold populations show significant physical differences. All known outer Solar System dwarf planets belong to the hot population (see reviews by, e.g., Brown 2008 and Sheppard et al. 2011). The hot population contains relatively few binary objects, and those that are present are not the equal-sized binaries commonly found in the cold population (Noll et al. 2008, 2020). The difference in surface properties between the dynamically hot and cold TNOs (Section 4.1 and **Figure 6**) has been used to support the idea that the hot

TNOs formed at smaller heliocentric distances than the cold population in the protoplanetary disk (e.g., Tegler et al. 2003, Levison et al. 2008), perhaps inside the ice line of particular volatiles that could account for the change in surface colors (e.g., Brown et al. 2011, Schwamb et al. 2019). Although the exact chemistry that leads to the range of TNO surface properties is not well understood (detailed spectral observations for all but the largest and brightest TNOs remain elusive owing to their inherent faintness), the idea that current colors are linked to chemistry in the disk is quite plausible and consistent with the idea that the hot TNOs were scattered outward and implanted onto their current orbits after forming closer in (see Section 4.2.5).

4.2.1. Scattering objects. The existence of the short-lived populations of giant-planet-crossing Centaurs (objects with $5.2 < a < 30.1$ AU) and inner Solar System Jupiter-family comets implied the existence of a flattened influx source population beyond Neptune well before the TNO populations were observationally confirmed. It was realized theoretically that feeding in these objects from $a > 30$ AU would unavoidably simultaneously create a population of TNOs scattered outward by Neptune (Torbett 1989, Duncan & Levison 1997); the first such object was identified early in TNO surveys (Luu et al. 1997). Many observational and dynamical studies quickly followed to confirm the connection among TNOs, Centaurs, and comets (see recent reviews by Dones et al. 2015 and Peixinho et al. 2020).

We define the population of scattering objects (sometimes called the scattered disk) as those that experience significant changes in a on short timescales (10 Myr in the classification scheme of Gladman et al. 2008) due to interactions with Neptune (see Section 3). As seen in **Figure 4**, TNOs with $q \lesssim 37$–38 AU tend to have strong-enough Neptune interactions at their perihelion to belong to today's scattering population, with higher perihelion objects belonging to the more stable detached population. However, this perihelion boundary is approximate; TNOs with very large a have lower-energy orbits that are more easily perturbed, so their semimajor axes can vary significantly on 10-Myr timescales even at larger q (see, e.g., Bannister et al. 2017, Khain et al. 2020). On 4-Gyr timescales, the largest q that exhibits semimajor axis mobility due to the planets rises to $q \approx 50$–60 AU by the $a = 2{,}000$-AU transition to the inner Oort cloud (Bannister et al. 2017). As one approaches this boundary, gigayear-scale q oscillations begin to occur that are the path into the inner Oort cloud (Duncan et al. 1987, Levison et al. 2006).

Figure 4 exhibits a strong drop-off in known scattering TNOs as both a and q increase; some of this is intrinsic and some is due to observing biases, which are strong owing to their large orbits. These objects can only be detected during a tiny fraction of their orbit near perihelion. The intrinsic scattering population fraction is much larger than the observed one; after accounting for observational biases, Lawler et al. (2018b) estimate that there are $\sim 10^5$ scattering TNOs with $D > 100$ km.

Long-term erosion of main-belt and resonant TNOs provides an insufficient influx to explain the observed scattering and Centaur populations (Dones et al. 2015). A better explanation is that today's scattering TNOs are an eroding remnant of an enormous primordial scattering population created as the giant planets dispersed the remaining planetesimal disk (Duncan & Levison 1997; see also the review by Gladman 2005). A small fraction of this massive primordial swarm (gravitationally scattering off the set of massive planets that existed in the early Solar System) appears to have had their perihelia raised to create the decoupled hot populations (Sections 4.2.3 and 4.2.4) through a set of unclear dynamical processes (reviewed by Gomes et al. 2008 and discussed in Section 5.2). Once this epoch of perihelion lifting ended, TNOs with q values low enough to still scatter were steadily depleted because their dynamical lifetimes (i.e., the amount of time they will persist in the scattering population) are shorter than the Solar System's age. Over 4 Gyr, the scattering population would decay to 1% of what it was when the giant planets first reached their

current orbits, with a t^{-2} decay that is slower than exponential owing to resonance sticking (see Duncan et al. 1987, Duncan & Levison 1997, and Section 4.2.2).

4.2.2. Resonant objects. A striking feature in **Figure 4** is the appearance of many pile-ups at specific semimajor axis values (i.e., orbital periods) corresponding to Neptune's mean-motion resonances. A very wide range of Neptune's resonances have observed members, and most of the observed orbits are stably resonant on timescales of the Solar System's age. Although the total population in Neptune's resonances does decay over time owing to chaotic diffusion (see citations to Morbidelli 1997, Tiscareno & Malhotra 2009), the dynamical lifetimes are much longer than they are for the scattering population and so escape is very slow. The prominent resonant populations were most likely captured during the era of giant planet migration (see reviews by Luu & Jewitt 2002, Morbidelli et al. 2008, Malhotra 2019, and further discussion in Section 5.1), although sufficiently populating the distant resonances is problematic (Chiang et al. 2003, Lykawka & Mukai 2007a, Gladman et al. 2012; see the sidebar titled Neptune's Stable Resonances).

Although the 3:2 resonance appears to be the most populated resonance (**Figure 4**), this is mostly due to observational biases. Having the smallest semimajor axis of Neptune's heavily populated resonances, the 3:2's members (called plutinos) are brighter on average over their full orbit (detection is always strongly biased toward perihelion, but larger-a TNOs spend a much larger fraction of their time further away). When observational biases are accounted for, many other TNO resonances have intrinsic populations similar to the 3:2, including the 2:1, 5:2, 3:1, 4:1, and 5:1 resonances, which are each estimated to contain 10^3–10^4 TNOs with $D > 100$ km (Gladman et al. 2012, Pike et al. 2015, Alexandersen et al. 2016, Volk et al. 2016). Many other resonances have too few observed objects (or candidate members whose discovery circumstances or orbits are too uncertain) to produce reliable population estimates. There are also poorer constraints on the resonant nature of many high-a TNOs simply because their semimajor axes are less precise owing to their long orbital periods (see Section 2). The most distant securely determined (see Gladman et al. 2008 for a definition) resonant objects are in Neptune's 9:1 resonance at $a = 130.1$ AU (Volk et al. 2018), but more distant resonances are potentially occupied by known TNOs whose orbital uncertainties are currently too large to be certain about resonant status. For example, applying the Gladman et al. (2008) classification algorithm, we currently find that 148209 (also known as 2000 CR105) and 474640 (also known as 2004 VN112) have 50% and 25% probabilities of being in the 20:1 and 36:1 resonances, respectively, despite relatively plentiful numbers of astrometric measurements.

Due to these observational limitations, for resonances with $a > 60$ AU it is unclear whether the populations are primordially captured or represent temporary resonant members more recently

NEPTUNE'S STABLE RESONANCES

One might be surprised to see so many TNOs trapped in Neptune's resonances given that the Solar System's other prominent stable planetesimal population, the main asteroid belt, is well known to exhibit deficits at most resonant locations. The asteroid belt's Kirkwood gaps occur at the locations of Jupiter's interior mean-motion resonances; significant portions of these resonances are unstable, meaning that resonant asteroids are subject to perturbations that raise e values to large values, leading to their removal from the asteroid belt when they have encounters with planets. In contrast, Neptune's external mean-motion resonances have large, stable libration zones, where TNOs can remain for >4 Gyr. Neptune's resonances are less laced with chaos than Jupiter's because perihelia longitudes in external resonances generally regress, limiting chaotic interactions with secular resonances that aid instability (Morbidelli et al. 1995).

trapped from the scattering population. Early simulations showed that scattering TNOs can stick to Neptune resonances as their semimajor axes wander (Duncan & Levison 1997), with sticking being most common for n:1 and n:2 resonances with Neptune (e.g., Lykawka & Mukai 2007b). Some observed TNOs are now certain to be on such short-lived resonant orbits (e.g., Alexandersen et al. 2016, Bannister et al. 2016, Holman et al. 2018). In simulations, up to 40% of the scattering TNOs can be sticking to resonances at any given time; the resulting expected number of transiently resonant objects in each of Neptune's close-in resonances is significantly lower than the observationally derived population estimates, supporting the idea that most currently resonating TNOs were captured early (Yu et al. 2018). Future observations will better constrain the memberships of additional large-a resonances and confirm that their significant populations are primordial.

Nearly all the prominent resonances exhibit widely dispersed inclination distributions for their members, which is consistent with other hot populations. The 1:1 Neptune coorbitals, the 3:2 resonance on the main belt's inner edge, the 5:2 resonance at $a \simeq 55$ AU, and all the more distant resonances have hot inclination distributions (e.g., Gladman et al. 2012, Parker 2015, Pike et al. 2015, Volk et al. 2016). A few of the resonances near or overlapping the cold TNO population have a mix of hot and locally trapped cold components, reflected in their slightly lower inclination widths compared to the hot population (e.g., the 4:3, 7:4, and 2:1 resonances; Gladman et al. 2012, Chen et al. 2019). The abundance of stable objects in Neptune's resonances, and their large eccentricities and inclinations, provided early clues that the transneptunian populations were likely sculpted by planetary migration (e.g., Malhotra 1995, Chiang & Jordan 2002); we return to this in Sections 4.2.5 and 5.1.

4.2.3. Hot classicals. The known classical TNO population extends from $a \simeq 36$ AU to $\simeq 50$ AU, just past the 2:1 resonance. The inner classical population (see **Table 1**) is entirely dynamically hot (see Section 4.1). Main-belt classicals cover the entire 39–47-AU range but only overlap the cold population beyond 42.5 AU (**Figure 5**). The outer classicals are exterior to the 2:1 and pose a nomenclature issue. Almost all $a > 48$ AU nonresonant and nonscattering objects have large enough e values that they are conceptually part of the detached population (see Section 4.2.4). But the historical usage of classical has been related to the concept of a preexisting belt, possibly weakly locally excited; it is unclear if the few low-e/low-i TNOs beyond the 2:1 are part of a cold outer classical population (Sections 4.1 and 6.2.3) or just the tail in these distributions for the dominant hot population there.

The eccentricity and inclination distribution of the main-belt hot population is largely inconsistent with dynamical excitation of an in situ cold population by giant planet migration (e.g., Hahn & Malhotra 2005). It is difficult to preserve the dynamically cold population's orbits in any scenario able to raise other locally formed TNOs to the level of excitation in the hot population. This, combined with the physical differences observed between the hot and cold populations, has led to consensus that the hot classical population was almost certainly implanted from a different formation location, though this transport and implantation mechanism must also avoid overly exciting the cold population (see Section 5.1).

After correcting for the ν_8's void and removing the cold main-belt component, Petit et al. (2011) showed that the similar i and q distributions (c.f., **Figure 4**) could be modeled with one continuous hot classical population for all TNOs with $a > 35$ AU; one need not think of the hot classical distribution as being split in semimajor axis domains at all. The exception to this is rare $q \gtrsim 43$ AU TNOs (**Figures 4** and **5**), which may result from resonance drop-off (discussed below). One single scattering process that transferred hot TNOs to all semimajor axes, and then raised some of their perihelia to $q \approx 35$–43 AU to strand them away from further scattering, is thus a

viable model. One could thus consider most of the $a > 48$ AU detached population to be just hot classicals, with the probable exceptions of TNOs with much larger i and/or q.

4.2.4. Detached population. The detached TNOs reside beyond the 2:1 resonance, with orbital distribution in a and i similar to the scattering population (see **Figure 4**) but with q values large enough that they lack strong Neptune interactions at their perihelia. Just like the hot classical population, these likely arise from some fraction of the scattering TNOs (at essentially all a values) having their q lifted in the past to create a population that now have their perihelia detached from Neptune. The $a > 48$ AU separation between the hot main-belt classicals and detached objects is somewhat arbitrary; an $e > 0.24$ "line in the sand" was defined to separate the detached TNOs from the outer classical population, motivated by the largest stable classical e values on the other side of the 2:1 resonance (Gladman et al. 2008). The term outer classical belt is (still nebulously) reserved for low-e and low-i TNOs at larger distances (see Section 4.2.3), because if a distant cold belt is eventually discovered (Section 6.2.3), its members will not be referred to as detached because they were never Neptune coupled.

Their larger perihelia make detached TNOs harder to detect compared to scattering TNOs. The large-a and large $q > 37$ AU nature of a detached orbit may not even be recognized without careful tracking over the months and year following discovery (Section 2). Thus, the first detached object was confirmed, and the existence of another dynamical class of TNOs hypothesized (Gladman et al. 2002), later than the other dynamical classes. The initially proposed term extended scattered disk (motivated by the extended perihelion range) is deprecated in favor of detached, as the latter is based on current dynamics rather than beliefs about past dynamical history. Some TNO literature uses the term extreme TNOs, with no generally agreed-upon definition, for a subset of the detached population. Scattering and detached behaviors occur at all observed semimajor axes, so a lower a bound (ranging from 150–250 AU in different papers) for extreme is not currently dynamically well motivated; this a boundary is sometimes said to be where Neptune's resonances stop being important, but we have shown in Section 4.2.2 that observed TNOs at $a > 300$ AU are potentially resonant. In the known Solar System, there is no change in behavior when crossing any a boundaries inside 1,000 AU (see the sidebar titled Nomenclature Matters). The change in dynamics diagnosed by the existing scattering to detached transition results in a fuzzy $q \gtrsim 37$–38 AU bound for detached TNOs with $a < 100$ AU, steadily rising to $q \gtrsim 50$ AU at $a \simeq 1{,}000$ AU. Until at least $a > 1{,}000$ AU (approaching the inner Oort cloud transition where galactic effects start to become important), all detached TNOs currently have essentially the same dynamical behavior (given the known planets) and there is no current justification for another semimajor axis division.

The first several detached discoveries had q values a few astronomical units above the scattering boundary (see **Figure 4**, where the red scattering population transitions into the beige detached). These objects provided the first examples of orbits that could not be generated by the known giant planets on their current orbits and suggested the need for additional masses that sculpted the TNO distribution (Gladman et al. 2002). Then the even higher-q outlier Sedna (also known as 90377 or 2003 VB12) was discovered (Brown et al. 2004). With $q = 76$ AU and $a = 506$ AU, Sedna had both the highest perihelion and largest semimajor axis for a noncometary object at the time and is well beyond the current gravitational reach of Neptune. Later surveys have revealed additional very high-q objects: 2012 VP113 ($a = 262$ AU, $q = 85$ AU; Trujillo & Sheppard 2014), 541132 (2015 TG387; $a \approx 1{,}100$ AU and $q = 65$ AU; Sheppard et al. 2019), and a few other $q \approx 50$ AU TNOs (**Figure 4**). Although the large a values of these objects might suggest the action of an external perturber (Section 5.2), not all high-q TNOs have large semimajor axes; 2004 XR190 (nicknamed Buffy; Allen et al. 2006) has $q = 51$ AU but low $a \simeq 57$ AU and $i = 47$ deg. This and

NOMENCLATURE MATTERS

The names used for grouping TNOs matter because they influence our thinking. Nomenclature based on orbital dynamics, which can be objective, still depends on our knowledge of the gravitational influences. The current scattering–detached TNO split is based on Neptune scattering in integrations with the four giant planets, but if there is a sufficiently massive undetected planet, these integrations are moot: The boundary may change entirely, both because the planet could scatter objects and because its effects could drive TNO perihelia into and out of Neptune's dominance. The current transition between TNOs and inner Oort cloud objects is where galactic effects dominate the orbital evolution. However, there are TNOs near the boundary, and rare deep stochastic (and unknown) stellar encounters (see, e.g., simulations in Sheppard et al. 2019) with the Sun are involved; this must thus be a somewhat arbitrary line in the sand (we use $a = 2{,}000$ AU). But such a boundary is irrelevant in the presence of an unseen giant planet because it may dominate the dynamics for some TNOs and make the inner Oort cloud terminology irrelevant there. Should a planet be discovered, the dynamical classification of large-a objects would need to be redone.

some other high-q TNOs with $a \lesssim 100$ AU are suggested to originate from resonance drop-off mechanisms (see Sections 4.2.5 and 5.2).

This is still a conceptually messy area, as the physical mechanisms that produced the detached population are not established. When thinking about their origins, it may be useful to consider the detached population as having three broad (but ill-defined) subgroups: (*a*) the mildly detached objects at all a, simply an extension of the hot classicals that have q somewhat (1–10 AU) larger than the scattering boundary; (*b*) a set of TNOs, mostly in the range of $a = 48$–100 AU, whose q may have been lifted even higher by resonance interactions (even if resonance interactions have also been suggested as the origin of the mildly detached objects); and (*c*) a set of very high-q TNOs (sometimes called Sednoids) of which only the three clear examples above are currently known.

The detached population has by far the most uncertain intrinsic population estimate. The observed detached TNOs vary in size from the most massive dwarf planet, Eris (also known as 2003 UB313 or 136199; with $a = 68$ AU and $q = 38$ AU), down to 100-km scale members. Due to the difficulty of finding them, and their large (and poorly constrained) ranges of q, a, and i, a population estimate ends up being dominated by the assumptions about those distributions; nominal estimates for the detached population's mass range from comparable with the main belt (Petit et al. 2011, Sheppard & Trujillo 2016) to 100× larger (Brown et al. 2004). Section 5.2 discusses how the detached population is critical to understanding the history of the outer Solar System; further observational constraints are badly needed.

4.2.5. Implantation of the hot populations. It is widely believed that the hot populations were transported from their formation location and implanted onto the orbits we observe today. Although one cannot rule out the possibility that some hot TNOs formed beyond the current main belt (see Section 6.2.3), it is generally assumed that most formed interior to ≈30 AU and were transported outward during the epoch of planetesimal-driven migration in the early Solar System. Fernández & Ip (1984) realized that Uranus and Neptune must migrate outward owing to interactions with the massive early planetesimal population; although they scatter planetesimals outward (allowing hot-population and Oort cloud creation), more angular momentum is gained scattering planetesimals inward to Saturn and Jupiter, causing net outward motion. Malhotra (1993) recognized that this outward migration allowed TNOs, like Pluto, to become trapped in Neptune's mean-motion resonances. The large resonant populations apparent in early observational studies

Planetesimal-driven migration: occurs later during planetary formation when planets exchange angular momentum with a massive planetesimal disk by scattering it

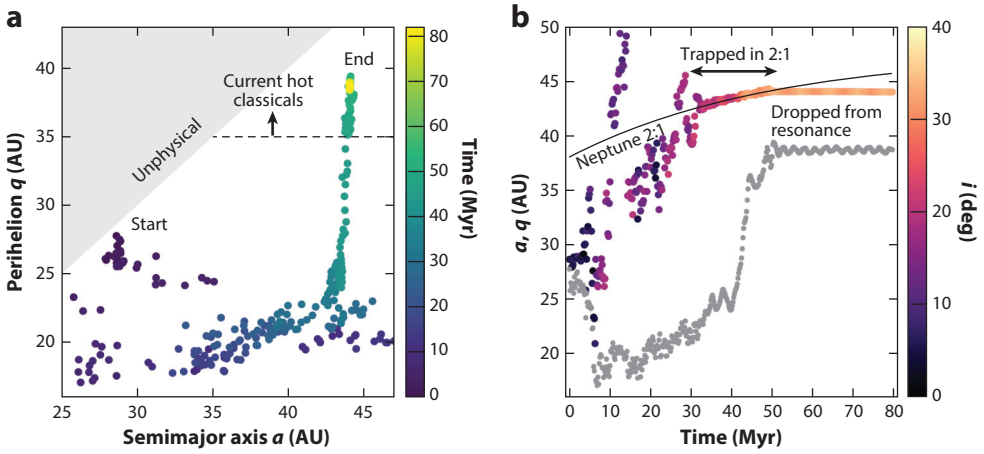

Figure 7

An illustration, drawn from simulations by Volk & Malhotra (2019), of a TNO starting on a nearly circular $i = 0$ deg orbit at 28 AU that is scattered by an outwardly migrating Neptune (in this simulation, Neptune migrates smoothly outward from 24 to 30 AU with a 50-Myr e-folding timescale). (*a*) The TNO's q and a evolution, color coded by time. (*b*) The same evolution versus time, with the particle's inclination color coded on the a evolution and the q evolution shown in gray; the changing location of Neptune's 2:1 resonance is indicated by the black line. Over the first ∼30 Myr, scattering raises the particle's e and i. Trapping in the 2:1 resonance occurs during ∼32–49 Myr, during which time i rises from ∼20 deg to 32 deg and q increases from 25 AU to 38.5 AU. At 50 Myr, the particle drops out of the 2:1 resonance (which continues moving outward, tracking Neptune's migration), leaving the object implanted in the current hot classical region with large i. Note that this particle's final inclination is part of the high-i tail of the inclination distributions produced in these kinds of migration scenarios (i.e., most particles implanted this way have lower inclinations). Abbreviation: TNO, transneptunian object.

(e.g., Jewitt & Luu 1995, Morbidelli et al. 1995, Chiang & Jordan 2002) thus provided strong evidence for outward Neptune migration.

During migration, many giant planet region planetesimals were scattered toward the Oort cloud (see Dones et al. 2015 for a review of Oort cloud formation and dynamics), creating an enormous hot scattering TNO population. These TNOs would initially have perihelia in their giant planet source region ($q < 30$ AU nominally), so their scattering orbits would need to be somewhat circularized and decoupled from Neptune for them to fill the observed a–q phase space of the hot population (see **Figure 4**). Along with decoupling q, the inclinations of most hot TNOs must be raised; scattering an initially flat disk of TNOs off of nearly coplanar giant planets is insufficient to excite the inclination distribution to the observed width. TNO decoupling (and possibly associated inclination increases) by migrating mean-motion resonances was suggested by Gomes (2003) and has been heavily investigated (reviewed by Morbidelli & Nesvorný 2020). We illustrate this mechanism in **Figure 7**, which shows a simulated implantation of a hot main-belt classical TNO. In this example, a TNO is scattered outward from an initially circular and coplanar orbit at 28 AU during the early stages of Neptune's migration, giving the TNO a lower q (higher e) and moderate inclination. Although Neptune is still migrating, the TNO sticks to Neptune's 2:1 resonance and is, for $\simeq$ 17 Myr, carried along with the migrating resonance. While being carried, the resonance dynamics raises the TNO's perihelion and increases its inclination. At some point, q is raised beyond Neptune's reach (e is low), so the TNO drops out of the resonance and is left behind as the 2:1 continues outward migration. This new orbit is stable on 4-Gyr timescales and is in the middle of the hot main belt (compare **Figure 7*a*** with **Figure 5**).

Although this resonant drop-off implantation mechanism almost certainly played a role in detaching some of the hot population, especially in the main belt, it cannot be the only q-lifting mechanism that occurred. Self-consistent models of the outer Solar System's dynamical history have yet to reproduce the detailed q distribution of the implanted hot populations. Migration and/or implantation models consistently struggle to produce the very wide hot-population inclination range. Thus far, the only consensus is that the hot populations were implanted; the details of this implantation remain active areas of research.

In summary, the current best explanation for the hot TNO populations are that they formed interior to 30 AU and were scattered and/or transported onto higher-a orbits during the epoch of giant planet migration. The different formation location compared to the cold population could explain their more neutral surface colors and different size distribution, and the lower binary fraction could result from either those different disk conditions or binaries being disrupted by the scattering process. The current scattering population is the decaying remnant of this more abundant primordial population, and the hot classical and detached populations likely represent members of the primordial scattering population whose orbits became decoupled from Neptune by having their perihelia lifted, allowing them to persist; the mechanism for this decoupling is an active area of research (see Sections 5.1 and 5.2). Most of the current resonant populations were likely also captured into Neptune's resonances during the epoch of planet migration from the scattered hot population, though some of the resonances in the main-belt region likely also have a component captured from the disk of objects that formed beyond 30 AU. Whatever transport and decoupling process produced these hot populations with much larger eccentricities and inclinations, it did not dramatically alter the orbital distribution or surface density of the current cold population from 42.5 AU to 47.5 AU.

5. TOPICS OF EXCITING RESEARCH NOW

We here transition from describing the current state of TNO observations and widely accepted interpretations to discussing some topics of current research. These topics still have multiple competing ideas or have only a framework that needs additional work before consensus will occur.

5.1. Migration Details

Many studies have explored different proposed dynamical histories for the giant planets and the implications of these formation and migration scenarios for the TNO populations. We briefly review key aspects of these models but refer the reader to recent reviews for more details (Morbidelli et al. 2008, Dones et al. 2015, Nesvorný 2018). Models of how Neptune's migration affected the early planetesimal disk have reproduced some of the rich dynamical features of the transneptunian region (Section 4), though models have thus far failed to simultaneously reproduce other important features.

Early migration models envisioned a steady outward movement of Neptune in response to interactions with a large number of relatively small planetesimals. Such smooth migration of Neptune's resonances into a cold TNO disk could explain the prevalence of resonant TNOs, but it overpredicted some resonant populations and did not produce the hot population orbital distribution (e.g., Hahn & Malhotra 2005); more complex migration models were required. Recent models incorporate graininess into Neptune's motion, i.e., small jumps in Neptune's orbit that would have occurred as it scattered dwarf-planet-sized (or larger) objects; how strongly this affects the final TNO population depends on the mass distribution assumed for the scattered objects. Other poorly constrained aspects of planetesimal-driven migration are its timing and total duration, the

possibility of the reshuffling of the order of the giant planets (accompanied by phases of planetary orbital excitation; Thommes et al. 1999), and the possible existence of other ice-giant-scale planets that were subsequently lost (Nesvorný 2011).

The reader might justifiably wonder how the migration is modeled when the planetesimal mass distribution, timing, and even the number of planets involved are poorly constrained. It is too computationally expensive to study suites of models that track all the most massive objects (e.g., from giant planets down to Earth-mass or Pluto-mass scale) and also the evolution of large numbers of TNO test particles. It is useful to categorize the models that study the implantation and/or perturbation of the extant TNOs:

- Neptune smoothly migrates out via a numerically imposed extra force that expands its orbit. Most previous work is of this type, choosing the initial and final locations of Neptune and the timescale for outward motion. Steady circularization and *i*-damping forces can be included to permit Neptune's initial orbit to have large *e* and/or *i*.
- The presence of the other large bodies scattering off of Neptune is incorporated in a heuristic way by jumping Neptune (in and out, although out must dominate), with the jumps motivated by other limited simulations to produce a grainy semimajor axis evolution.
- Direct incorporation of a limited number of massive objects in a partial *N*-body sense (e.g., the massive objects gravitationally interact with the planets in the simulation to drive migration but do not interact with each other).
- Direct incorporation of an even more limited number of massive objects in a full *N*-body sense (i.e., all massive objects fully interact).

The last two categories are the least well-explored owing to computational expense, but such models are often used to help guide parameter choices for the other two categories. We summarize the main conclusions from these types of migration models below.

In situ excitation of the hot TNO populations during migration does not sufficiently explain their large inclinations and eccentricities (Section 4.2), nor the main belt's bimodal *i* distribution. Levison & Morbidelli (2003) suggested that all (hot and cold) TNO populations were implanted from initial locations inside ∼35 AU, which is a process that should increase dynamical excitation. This idea was followed by the proposal that planet–planet instabilities played an important role in giant planet migration (in addition to planetesimal-driven migration; Tsiganis et al. 2005). In these models, the giant planets' orbits were excited because of mutual interactions before being damped back down via dynamical friction to their final nearly circular and coplanar state. In early versions of these models (which were done in a partial *N*-body sense), the planetesimal disk was truncated at 30 AU as a mechanism to stop Neptune's migration at that distance (Gomes et al. 2004); otherwise the fuel of planetesimal scattering is just resupplied as more primordial cold planetesimals are accessed. TNO transplants originating from such a truncated disk via large-scale giant planet instabilities (e.g., Tsiganis et al. 2005) can produce some hot population features (e.g., Levison et al. 2008) though they fail to produce the hot–cold inclination dichotomy and enough low-*e* orbits (Petit et al. 2011). It is now clear that the hot populations also differ physically from the cold population, and the general consensus is that the cold population formed in situ (Sections 4.1 and 4.2). Recent work suggests an in situ cold population could be retained, maintaining their low *e* and *i* during instability-driven migration (though it places limits on these scenarios, particularly on the maximum excitation of Neptune's orbit; Batygin & Brown 2010, Dawson & Murray-Clay 2012, Gomes et al. 2018; see also the review by Nesvorný 2018). However, a currently unexplained dramatic drop-off in the planetesimal disk surface density is still required for Neptune's migration to stop at 30 AU (c.f., Section 6.2.3).

The idea that the main-belt hot TNOs were implanted from elsewhere in the disk remains one of the most secure ideas to have arisen from the above models (e.g., Levison et al. 2008). We note that this scenario was motivated by but does not rely on the so-called Nice model picture, which involved a giant planet instability and the much-discussed (and justly criticized) idea of an ≈0.5-Gyr metastable delay before dispersing the outer Solar System's massive primordial planetesimal disk; this delay has since been disavowed even by its original proponents as being inconsistent with constraints on the Solar System's dynamical history (e.g., Nesvorný et al. 2018). Newer planetary migration models posit very early instabilities (of varying severity) involving the most massive bodies followed by a long (possibly grainy) migration of Neptune (reviewed by Nesvorný 2018).

A recent trend favors a long-duration tail for Neptune's migration, as it may better produce the detached population (see Section 5.2) and high-inclination TNOs. Nesvorný (2015b) suggested that a long period of smooth migration is needed to allow enough time for scattering plus secular i excitation to produce the observed hot population inclinations, but not all migration simulations show such a timescale dependence (Volk & Malhotra 2019). Migration timescales also affect the detailed distributions of captured resonant objects (e.g., Murray-Clay & Chiang 2005) as well as the number and distribution of detached TNOs (e.g., Kaib & Sheppard 2016, Pike & Lawler 2017). The amount of graininess in late migration affects how well Neptune's resonances capture TNOs; too much graininess could underpopulate the resonances, but adding some graininess can help reduce the final resonant populations to be in line with what we observe (e.g., Murray-Clay & Chiang 2006, Nesvorný & Vokrouhlický 2016). Graininess also enhances the resonance drop-out mechanism for producing some of the detached TNOs (e.g., Lawler et al. 2019). Larger-scale jumps in Neptune's orbit (due to planet–planet scattering) have been invoked (Nesvorný 2015a) to explain the cold-belt kernel, though this picture has some less desirable consequences (see the sidebar title The Main-Belt Kernel in Section 4.1).

No single proposed migration scenario explains the entire TNO population, and many observed features can be produced by more than one mechanism, the differences between which are not yet observationally distinguishable. Generally, some sort of direct planet–planet interactions or dynamical upheaval (often invoked to produce the giant planets' orbits and possibly involving an additional ice giant; see Nesvorný 2018 review) allows enough excitation of planetary orbits to dynamically heat and scatter out a sufficiently high-i hot population; a subsequent period of slow migration (perhaps grainy) then implants some of those TNOs onto metastable orbits by raising their q or capturing them into resonances. Although models can easily populate the close-in resonances, they have thus far struggled to produce the large observed populations in Neptune's more distant resonances (Section 4.2.2). Section 5.2 discusses certain aspects of the TNO population that may result from rogue planets that were eventually scattered out of the Solar System.

A major challenge in finding an outer Solar System dynamical history that recreates the observed TNO populations, especially the hot populations, is that the implanted orbits are rare end states. Because simulations show efficiencies for hot TNO implantation are low (0.01–1% of the $a < 30$ AU reservoir), it is challenging to simulate enough particles to build a statistically meaningful predicted population, even for a single migration model. When added to the large number of free parameters in these models (initial planet number and positions, migration timescale, level of graininess, etc.), it becomes difficult to make robust model predictions. There is also an observational aspect to this challenge. For example, determining whether the detached population is predominantly implanted via resonant interactions during Neptune's migration or via interactions with rogue planets (see Section 5.2) requires improved constraints on the intrinsic orbital distribution of today's difficult-to-detect detached population.

Overall, significant observational and modeling work remains to determine what the surviving TNO populations tell us about planetary migration. There are promising leads on what physical

processes may have been important, but much of the work is fragmented; often, only one aspect of the belt is explained by each mechanism, and the ramifications for the rest of the belt are not sufficiently examined. As an example, Neptune jumps that strand formerly resonant particles should leave stranded clumps all across the belt, an as yet unexamined observational consequence. Similarly, models developed to populate one particular resonance must be generalized so predictions about the relative populations of resonances at all semimajor axes can be made. Future, characterized, large-area surveys (Section 6.3) that detect factors of several more TNOs will provide more model-testing opportunities. Given the biases in the known TNO sample (Section 2), it is critical that these are accounted for when testing models against observations; Lawler et al. (2018a) describe how models can be robustly compared with TNOs detected by characterized surveys.

Characterized surveys: surveys that record their pointing history, detection efficiency versus magnitude and rate, and tracking efficiency, allowing robust bias corrections

5.2. Detached Population Is Key for Understanding Presence of Other Large Objects

The production mechanism for the detached population is at the forefront of TNO research as a key constraint for understanding the Solar System's dynamical history because the required perihelion lifting must involve significant additional massive perturbers in the early Solar System. Whether this additional mass is concentrated in a few bodies or a more extended mass distribution is unclear. A number of scenarios have been proposed, including the following: the combination of resonant and secular dynamics during the end stages of Neptune's migration, gravitational influence of nearby stars in the Sun's birth cluster, interactions with rogue planets in the early Solar System, interactions and/or self-gravity with a massive distant disk of objects, and an extant undetected large planet in the outer Solar System (discussed in Section 5.3).

Some detached TNOs have semimajor axes near resonance locations (**Figure 4**), leading to suggestions that resonant dynamics played a role in their emplacement, much like the example shown in **Figure 7**. Within Neptune's mean-motion resonances, TNOs can experience a secondary effect; the so-called Kozai secular resonance (see Section 3) produces correlated changes in a TNO's q and i. During the late, presumably slow, stages of Neptune's outward migration, some scattering TNOs would stick to Neptune's distant resonances (Section 4.2.2) and experience Kozai evolution; in some instances, a TNO's q would increase and the TNO would drop out of resonance. As the planet continued outward migration, so did the resonance in question, thus permanently stranding the TNO at a higher q now-detached state (Gomes 2000, 2003). These resonant interactions plausibly explain a subset of the detached TNOs (Gomes et al. 2008), especially the observed high-q detached objects like Buffy (2004 XR190) located just interior to some of Neptune's strong resonances. Some modeling has studied the distribution of these detached objects near a variety of resonances under different migration scenarios, though the observational constraints are insufficient to usefully test these models (e.g., Kaib & Sheppard 2016, Pike & Lawler 2017).

Some simulations indicate that the resonant production of detached TNOs becomes less effective at $a \gtrsim 250$ AU and does not produce perihelia above $q \sim 70$ AU (Brasser & Schwamb 2015). Additionally, it cannot explain low-i detached TNOs at very large q because the increase in q because of Kozai interactions must be coupled to an increase in i (Gomes et al. 2005, 2008). The detached population's inclination distribution will thus constrain how dominant this mechanism is; detached objects produced this way should have inclinations skewed higher than that of the current scattering population. The current observational data are insufficient here (**Figure 4**), but the inclination distribution of the most distant TNOs is poorly known owing to relatively small numbers of characterized detections and the limited latitude range of characterized surveys (see, e.g., Petit et al. 2017).

The discovery of Sedna with $q = 76$ AU and $a \simeq 500$ AU, well inside today's $a \approx 2{,}000$ AU inner Oort cloud boundary but with q well beyond Neptune's influence, renewed interest into whether such detached TNOs could be members of a fossilized inner Oort cloud (e.g., Morbidelli & Levison 2004, Brasser et al. 2006). The Sun likely formed in a stellar cluster, and this denser early stellar environment may have led to more frequent close stellar flybys that perturb TNOs with $a \simeq 1{,}000$ AU orbits at that time (see Ida et al. 2000 and citations to it). If giant planet formation and the ensuing creation of a huge scattering and/or Oort cloud population occurred before the Sun's birth cluster dispersed, then the early inner Oort cloud (in this case, defined by stellar passage effects) could have extended further inward than today. This early inner Oort cloud would become fossilized after cluster dispersal because the objects in it would have perihelia well beyond Neptune's gravitational influence but aphelia well inside the subsequent influence of galactic tides and passing stars. Models of such scenarios (Brasser et al. 2012) found Sedna-like orbits being produced in the very densest birth cluster scenarios and even raise the possibility of such objects being captured from the intruding star's Oort cloud (Morbidelli & Levison 2004). These scenarios require tight timing constraints: The stellar flyby must occur late enough that the planets have formed and created an Oort cloud population to be lifted but before the cluster disperses. The stellar cluster must also be dense enough to make a deep encounter possible, but not so dense that likely subsequent birth-cluster encounters strip off the detached and Oort cloud objects (Pfalzner & Vincke 2020). With statistics of one Solar System, judging the probability of this scenario is challenging; this sequence of events will likely occur for a minority of G-type stars. Finally, there are high-q detached TNOs with smaller semimajor axes that are very hard to create via stellar encounters, thus requiring an additional mechanism.

Another early suggestion for detached population production was rogue planets in the early scattering population that have since been lost from the Solar System (Gladman & Chan 2006). It seems unlikely that nature created four giant planet cores but then nothing else larger than dwarf planets (Pluto and Eris scale) in the outer Solar System. Any Mars-Earth-scale objects that formed would have joined the initial scattering population; those with q values near the outermost ice giants can survive ~100 Myr. Their temporary presence in the scattering population, with a values of hundreds of astronomical units, causes secular oscillations that detach scattering TNOs (**Figure 8**). A high-e rogue that reaches $a \simeq 500$ AU could create all known detached objects, with the most-massive rogue having most of the q-lifting power. Should multiple rogues be present, their self-interactions could decouple one of them, with the most likely such case leaving a Mars-scale object stranded in the $a < 200$ AU detached population (Silsbee & Tremaine 2018).

Although most attention has been paid to the presence of other very large objects in the disk, it is also possible that a huge number of small objects have enough collective self-gravity to cause perihelion lifting. Madigan & McCourt (2016) describe an instability that can occur in a very massive disk of objects on eccentric orbits; during this instability, orbital e can decrease dramatically, offering another potential pathway for producing detached TNOs. It is unclear that the initial orbital distribution and required duration for the instability to function (Zderic et al. 2020) can be created by an outwardly scattered planetesimal disk.

5.3. Speculations About Other Still-Resident Planets

The search that produced Pluto's discovery was motivated by speculation that other large planets beyond Neptune existed; these were based on cosmogonic arguments about the extent of the protoplanetary disk or (incorrect) evidence for residuals in Neptune's motion due to a massive perturber. The idea that Mars-Earth-scale planets might have assembled beyond 50 AU (e.g., Stern 1996) continued into the early era of modern transneptunian studies. With the current belief that

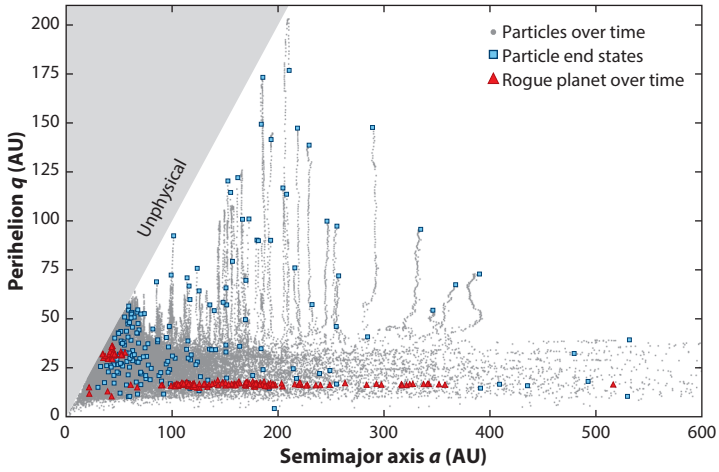

Figure 8

Example of a rogue planet raising TNO perihelia. This unpublished simulation, from the Gladman & Chan (2006) suite, shows the a, q evolution of a 2-$M_\oplus$ rogue planet (*red*), which is scattered out and spends 150 Myr in the $100 < a < 400$ AU scattering disk (with q near Uranus). Gray points track the evolution of an initial 20–50 AU disk of cold test particles to their final value (*blue squares*). When some of these particles are scattered to semimajor axes within a factor of two of the rogue's a, the secular (averaged) gravitational effect of the rogue causes their q to rise and decouple from Neptune; when the rogue is ejected, TNOs with $q >$ 40 AU freeze at their current orbits, leaving a detached population. Note the production of near-circular $a =$ 150–200 AU orbits. An animation of this same parameter space for another simulation, showing known large-a detached TNOs, is available in **Supplemental Video 3**. Abbreviation: TNO, transneptunian object.

Supplemental Material >

dwarf-planet-scale (and larger) TNOs all belong to the hot TNO population that formed closer to the Sun and were then moved outward and decoupled, in situ formation is no longer required.

More recent postulates about still-existing planets have been based on the detected TNO orbital distribution. Aspects of the dynamical structure and the 48-AU edge in the classical belt might be explained by the current presence of a planet beyond 50 AU (Gladman et al. 2001, Brunini & Melita 2002, Lykawka & Mukai 2008). The detached population could be due to a still-resident planet (Gladman et al. 2002) or a now-gone rogue (Section 5.2). The midplane of the TNO distribution beyond 50 AU has been suggested to be warped, with a resident Mars-scale object a possible explanation (Volk & Malhotra 2017), although a later independent data set favors no signal (Van Laerhoven et al. 2019).

More recently, two groups have posited that the orbital distribution of a selected subset of large-a TNOs point to the current presence of a super-Earth- to Neptune-mass planet in the 200–1,000-AU region (reviewed by Trujillo 2020 and Batygin et al. 2019). These arguments were based on apparently nonrandom distributions of the orbital elements of a small TNO subset, but there was no prior hypothesis driving the subset's definition. Although it can identify interesting trends worthy of investigation, this data dredging (or p-hacking) approach results in spurious estimates of the false positive probability (e.g., Young & Karr 2011) for the initially identified trend. Future work can, however, accept the now-claimed hypothesis and then (by not using the previous sample that contains the possibly spurious signal) test it. Analysis of independent surveys to assess the strength of the proposed nonrandomness derived from the previous large-a TNO subsample do not provide support for the claimed signal (Shankman et al. 2017a, Bernardinelli et al. 2020, Kavelaars et al. 2020); they also demonstrated how concentrations in orbital-element

space can be generated by the location of survey fields. We thus believe the evidence for another current large planet provided by the angular elements of known TNO orbits is weak.

There are additional uncomfortable consequences from an $\approx$10-$M_\oplus$ planet at several hundred astronomical units for the orbital evolution of other large-a TNOs. The planet's gravitational influence strongly alters the q and i distribution of these objects (Shankman et al. 2017b), and may thus alter the Solar System's orbital distribution in ways that do not agree with observational reality; because the biases against large a, q, and i are so strong, however, it is difficult to determine if the observed orbital distribution is discrepant (Lawler et al. 2017). As described earlier (Section 5.2), an unavoidable consequence of a massive planet at several hundred astronomical units is that it forces scattering objects to oscillate up and down in q. These oscillations would traverse the putative q gap (initially defined as $q = 50$–75 AU by Trujillo & Sheppard 2014 and contracted to $q = 50$–65 AU by Sheppard et al. 2019, as shown in **Figure 4**) that has also been attributed to the planet's presence; a gap in the q distribution of large-a TNOs would thus be evidence against an extant planet large enough to significantly perturb their orbits. We note that, for the orbits proposed for the planet, this q cycling would not extend inward enough to explain much of the $a < 100$ AU detached population, so other origins for those detached objects are still required (Section 5.2). Lastly, a large-mass planet would also subtly alter the movement of the known giant planets; high-precision tracking of spacecraft orbiting them rule out a 10-$M_\oplus$ planet within 650 AU (e.g., Fienga et al. 2020).

Independent of arguments based on large-a TNOs, there are observed exoplanets more than 50 AU from their host stars. To date all these have been detected via direct imaging and are thus Jupiter scale or larger (see, e.g., the review by Winn & Fabrycky 2015) and plausibly formed around stars that had massive disks that extended to these distances (see Section 6.2). In our Solar System, direct all-sky spacecraft surveys in the infrared largely rule out the existence of another Jupiter bound to the Solar System or another Saturn within 28,000 AU (Luhman 2014). Uranus- and/or Neptune-scale objects have more uncertainty in their infrared flux, but three-quarters of the sky has been searched to distances of 650–900 AU with no detections (Meisner et al. 2018). Although we find it impossible to imagine that astronomers have already discovered the largest objects in the transneptunian region (see Section 6.1), the evidence for any of the specific proposed additional planets is generally weak, and observational limits on the presence of the largest (Neptune scale) additional planets continue to shrink their potential hiding places.

5.4. Planetesimal and/or TNO Formation Mechanisms

A recent dramatic perception shift is the movement away from the idea that the main-belt TNOs represent the remnant of the accretion and then near-total destruction of an initially massive disk from 30–50 AU (reviewed by Kenyon et al. 2008). There is now substantial evidence (Section 4.1) that the cold TNOs have not been significantly collisionally or dynamically depleted since they formed, but instead had a primordial surface density not terribly different than today's.

Because traditional bottom-up accretion models require significantly more mass and density to build 100-km radius objects than is present today, this has led to new planetesimal formation models in low surface density environments. These more efficient formation mechanisms may also be needed to help form the numerous dwarf-planet-scale TNOs in the primordial hot population, which were substantially more abundant than in today's dynamically depleted hot population (Morbidelli & Nesvorný 2020). Lastly, the asteroid belt's size distribution has been used to argue for the idea that asteroids were born big (that is, mass was quickly assembled from submeter size to objects as large as 100 km; Morbidelli et al. 2009) and outer Solar System planetesimal formation would plausibly be similar.

The currently favored formation mechanism is that, once enough dust is condensed out of the cooling protosolar nebula, the streaming instability concentrates small (gas-coupled) solids together so that they coalesce into large objects (Youdin & Goodman 2005, Johansen et al. 2007) that are then able to avoid solar inspiral caused by gas drag. Estimates of the mass scale associated with this instability have varied, but these model dependencies might provide a way to access the nebula's physical conditions (e.g., solid-to-gas ratio, viscosity) based on the observed TNO size distribution. Li et al. (2019) provide recent streaming instability simulations showing that the mechanism can produce size distributions consistent with various functional forms used to describe the observed TNO H magnitude distributions. Morbidelli & Nesvorný (2020) provide a recent review of the development of this topic.

Another important constraint on TNO formation mechanisms is the prevalence of binaries in the cold TNO population (see Section 4.1). The observed cold classical binaries likely formed as multiples rather than becoming binaries through dynamical capture interactions after formation; the dynamical interactions that could generate binaries from populations of single objects are inefficient, making it difficult to explain the large binary fraction in the cold population (Section 4.1). Additionally, the known binary pairs have self-similar colors (indicating consistent compositions) and mutual orbits that are mostly prograde (i.e., the binary pair orbit each other in the same sense as the heliocentric orbit of their center of mass). Nesvorný et al. (2010) showed that gravitational collapse via the streaming instability can form a population with a very high fraction of prograde binary pairs; formation from the same localized area in the disk would naturally lead to the similar compositions seen in the known TNO binary population. Increasingly improved simulations of streaming instability continue to produce more detailed predictions (e.g., Li et al. 2019, Nesvorný et al. 2019) of the size distribution and binary orbital properties (such as binary separation and inclination) that can be tested against the improving observational constraints on binary TNOs. We refer the reader to recent reviews on TNO binaries by Noll et al. (2020) and Brunini (2020). Additionally, McKinnon et al. (2020) discusses how the recent *New Horizons* spacecraft's visit to the cold belt, contact-binary TNO Arrokoth fits into this framework.

6. PROSPECTS FOR THE NEXT DECADE

We finish by turning our attention to some exciting topics. These have been selected because of our optimism that the next decade will allow significant progress.

6.1. How Complete Is the Inventory of Large Objects in the Hot Population?

The inventory for large TNOs in the cold main belt is nearly complete (Section 4.1). The larger e and i values of the hot $a > 50$ AU hot population, however, mean that large hot objects remain to be found at greater distance. This is evidenced by continued discovery of large-diameter, high-a objects of moderate (e.g., Bannister et al. 2017, Holman et al. 2018, Sheppard et al. 2019) to high inclination, especially in surveys with large latitude coverage (Weryk et al. 2016, Petit et al. 2017, Becker et al. 2018). Even near-polar orbits (see the sidebar titled The Polar Corridor) with $D >$ 100 km are now steadily being discovered.

Although there is not strong evidence for super-Earth-mass planets currently in the transneptunian region (Section 5.3), we find it nearly inconceivable that there is not an object somewhat larger than Pluto–Eris scale still undetected in the scattering or detached populations. If the Solar System's retention efficiency is on the order of a percent, then the largest TNO likely to remain is the size where ~100 objects existed at the start of dispersal in the giant planet region. This is almost certainly larger than Pluto–Eris, for which ~1,000 are thought to have formed (see, e.g., Stern 1991, Morbidelli & Nesvorný 2020) but smaller than Earth scale. Our money is on a Mars-scale

object in the inner Oort cloud or large-a scattering and/or detached population, but direct constraints on objects of this scale beyond a distance of 300 AU are still weak (e.g., Ashton et al. 2020).

6.1.1. Is There an Entire Other Population? The historic (and intrinsic) observational bias against discovering TNOs with large i, a, and q values means that there could be TNO groups that have not yet been recognized. First, and with near certainty, large-area deep surveys will begin to find the largest objects in the inner Oort cloud, with $a > 2{,}000$ AU and $q \gg 40$ AU. In our view, the TNOs so-far claimed to be in this population do not quite satisfy the definition that the current galactic environment is capable of raising their inclinations and perihelia to their current values from a scattering state (which happens for $a \gtrsim 2{,}000$ AU). Instead, the known $a \simeq 500$–$1{,}500$ AU TNOs might be due to a fossilized inner Oort cloud or rogue planets or have been produced by other q-lifting mechanisms (Section 5.2). The two TNOs approaching the inner Oort cloud boundary are 2015 TG387 ($a \simeq 1{,}100$ AU and $q \simeq 65$ AU) and 2014 FE72 ($a \simeq 1{,}500$ AU and $q \simeq 36$ AU), which Sheppard et al. (2019) show have a variations of hundreds of astronomical units predominantly driven by planetary perturbations. Although it is possible these objects are related to an earlier, more compact inner Oort cloud, we do not believe nomenclature and the resulting interpretation should be driven by unproven dynamical histories of the Solar System, and thus applying the Oort cloud terminology clouds the issues (see the sidebar titled Nomenclature Matters). The most promising probe of the inner Oort cloud might not be its largest objects, however, but rather the hordes of its smaller (~10 km) members that might eventually be detected via serendipitous stellar occultations (Section 6.3).

There are several other possible new populations. There could be a reservoir population of very high-inclination (including retrograde) TNOs. The discovery of a population of $i \simeq 90$ deg TNOs is recent (see the sidebar titled The Polar Corridor) and at least partly due to ephemeris bias (Section 2). TNO 471325 (also known as 2011 KT19 and nicknamed Niku) was only recognized as having a very high i upon rediscovery by Chen et al. (2016); the provisional short-arc 2011 orbit was satisfactorily fit with a prograde assumption. Future discovery of more polar corridor objects will help characterize this population and indicate whether the observed objects represent the most observable members of a hidden reservoir uncoupled to Neptune.

There could also be as yet undiscovered groups in the detached TNOs. If there currently is a Neptune-scale planet at a few hundred astronomical units, its ability to detach scattering objects

THE POLAR CORRIDOR

The scattering population has a wide i distribution, with a long tail above 30 deg, including $i \sim 90$ deg. We adopt the term polar corridor (Namouni & Morais 2018), refining the definition to $i = 90 \pm 30$ deg and $q > 5$ AU. Because of very-high relative planetary encounter speeds, polar corridor objects with $a > 30$ AU only evolve on gigayear timescales (Gladman et al. 2009), migrating in q and a but maintaining large i. The three polar corridor TNOs (with $a > 30$ AU and $q > 19.2$ AU) in **Figure 4** are, in increasing i order: 127546 (also known as 2002 XU93; Elliot et al. 2005), the first retrograde TNO 528219 (also known as 2008 KV42; Gladman et al. 2009), and retrograde 471325 (i.e., 2011 KT19; Chen et al. 2016). Five more polar corridor TNOs have q between Saturn and Uranus and another seven have $5 < q < 10$ AU; no detached examples yet exist. Given the biases against identifying them, they represent a substantial population not created in previous hot population production models. Ideas discussed include a primordially emplaced high-i TNO reservoir (Gladman et al. 2009), feeding high-i objects using a distant planet (Batygin & Brown 2016), captured interstellar objects (Namouni & Morais 2018), or an eroding remnant that Jupiter emplaced long ago (Greenstreet et al. 2020).

from Neptune's grasp would result in an enhancement in the detached a distribution near the planet's location (Lawler et al. 2017). This could also be true for a now-gone rogue planet if it spent significant time at a localized a range. The large perihelia of the detached objects make it difficult to detect such an enhancement; very deep ($m_r > 25$) surveys will be required.

Pebble accretion: envisions gas-drag assisted accretion of mm- to cm-scale bodies to accelerate the final growth of planets

A more distant cold TNO population may also remain completely undetected. This would not be unprecedented in the context of observed debris disks around other stars (Section 6.2), and we discuss prospects for detecting it in Section 6.2.3.

6.2. Context with Star Formation

Last year's Annual Reviews article on protoplanetary disk structure (Andrews 2020) and a review of extrasolar Kuiper Belts (Wyatt 2020) provide up-to-date background and plentiful references for the amazing advances in our understanding over the past decade regarding the presence and conditions within both protoplanetary disks and evolved debris disks around other stars. Here, we concentrate mostly on implications for the radial structure and the extent of our Solar System's initial disk, as well as the TNO formation process.

6.2.1. Protoplanetary disks. As a brief primer: Protoplanetary disks surround forming stars, showing roughly power-law surface density (mass per unit area projected down onto the midplane) profiles as one moves away from the star. During the accretion disk phase, the local temperature at the (usually optically thick) midplane also drops with power-law behavior (set heavily by local dissipation and potentially some irradiation of the flared surface of the disk once the protostar's luminosity ramps up). Perhaps surprisingly, studies from our Solar System's meteorites, interplanetary dust particles, and especially in situ comet tail samples (reviewed by Brownlee 2014) show that presolar solids are incredibly rare; nearly all presently solid materials in the Solar System were subject to high-temperature events (of unknown nature, especially for silicate materials that appear to have condensed in the outer Solar System), then reequilibrated and recondensed. Solids are observed in emission and absorption in protoplanetary disks and their chemical composition and the physical state of the material will depend on both the local general physical conditions (especially temperature) and unclear processes of short duration (example: the chondrule heating events). A rough rule of thumb is that (except for ancient stars with very low-metallicity disks) the condensates will eventually be on the order of 1% of the gas mass, with about twice as much available in regions below ice condensation temperatures.

Gas disks are usually optically thick, but the condensed dust (which produces more emission than larger solid bodies of lower surface area to mass ratio) can be optically thin at some wavelengths. This allows modern studies at millimeter and radio wavelengths to now directly study the resolved radial profiles of the solid material out of which planetesimals will be built. This fundamental increase in observational capability has permitted detection of azimuthal and radial structures (gaps) related to disk processes (e.g., van der Marel et al. 2019). Smooth disk models are now insufficient representations of what is going on in early planetary systems; rapid mass migration of small solids in disks with pressure bumps and condensation fronts are actively discussed (Andrews 2020). Additionally, these new observations and recent modeling work indicate that the timescales for the final solidification of the condensable disk mass, for the formation of planetesimals, and for final growth to large planets (possibly due to pebble accretion, reviewed by Johansen & Lambrechts 2017) may be much faster than previously believed (see Tychoniec et al. 2020, ALMA Partnership et al. 2015, and citations therein). There are new puzzles arising from our improved observational constraints. One particularly relevant to TNO studies is whether circumstellar bands with less emission (often called gaps, although the flux within is not zero) are places in which condensation into dust is ineffective, or transport away is rapid, or instead are

places in which growth is so effective that most mass is already sequestered into larger objects invisible at these wavelengths. Could edges (or significant surface-density drops), such as the one at the 30-AU drop postulated for our Solar System (Section 6.2.3), be caused by pressure bumps in the disk that generate disk enhancements by inward migrating of small solids? The answer to these questions will have profound consequences for the importance of the streaming instability (Section 5.4) and pebble accretion. Our understanding of planetesimal formation should improve dramatically over the next decade as continued observational investigation of young disks advances to higher resolution and greater sensitivity.

From the data at hand, it is already clear that there is much variation in the scale and mass of protoplanetary disks; more massive stars tend to have larger gas disks (extending out to beyond 100 AU; Matrà et al. 2018), which is unsurprising given that Jupiter-sized planets are known to exist at these distances (Winn & Fabrycky 2015 review planetary system architectures). The streaming instability mechanism may allow planet formation to occur out to great distances in lower-density environments than previously thought. However, even disks with large amounts of material can have sharp edges due to the influence of external factors; for example, the famous truncated proplyd disks in Orion (O'dell & Wen 1994, and references therein) are due to photoionization in the birth cluster from nearby O/B stars that strips off the disk's gas and any solids still small enough to be dynamically coupled to it. We discuss edges in the Solar System context in Section 6.2.3.

6.2.2. Debris disks. After the protoplanetary gas is dispersed, most of the mass remaining around other stars becomes nearly invisible to us. The most easily detectable material is dust, which either scatters the central star's light (like the image of Fomalhaut in **Figure 9*d***) or reradiates absorbed energy at infrared and longer wavelengths. Unlike the planet-forming dust detected in protoplanetary disks, in older stars the dust is viewed as short-lived debris generated by collisions of larger objects. Our Solar System's transneptunian material, if it was detected from outside, would be of this type; the fraction of our material in dust is tiny, however, so detection is unlikely (Poppe et al. 2019; see also the review by Hughes et al. 2018).

The most famous debris disk is Beta Pictoris, with both a detected planet (Lagrange et al. 2010) and an abundant debris dust; the inferred orbital inclination distribution of the dust-generating planetesimals is bimodal (Matrà et al. 2019), implying that a vestigial cold population and overlain hot population may be a generic feature of extrasolar transneptunian-belt analogs (Wyatt 2020). Another frequent phenomenon is debris rings (like in Fomalhaut), where, even if not resolved, the spectral energy distribution implies that the dust is concentrated in certain stellocentric distance ranges where a planetesimal belt is undergoing collisional activity.

Our Solar System serves as a decent distribution template for at least some of these systems, although the total quantity of dust present in our system corresponds to an optical depth several orders of magnitude lower than any detected debris disk (Matthews & Kavelaars 2016, Hughes et al. 2018). **Figure 9** shows a debiased model of our TNO populations viewed from various perspectives, including a viewing angle similar to that of the Fomalhaut system. (The comparison is imperfect because dust presence is not perfectly correlated with where the planetesimals are and light scattering effects are not included.) Several features in this figure are worthy of note. First, the face-on view shows very little azimuthal structure despite the fact that the resonant population is large (Section 4.2.2). Resonant populations are often used to create models for observed azimuthal asymmetries in debris disks (Wyatt 2020); in our Solar System, however, obvious asymmetry is erased by the large range of libration amplitudes and the great variety of populated resonances (only small-amplitude libration in n:2 resonances shows the much-discussed pericenter confinement near ± 90 deg away from the planet). In any planetary system in which

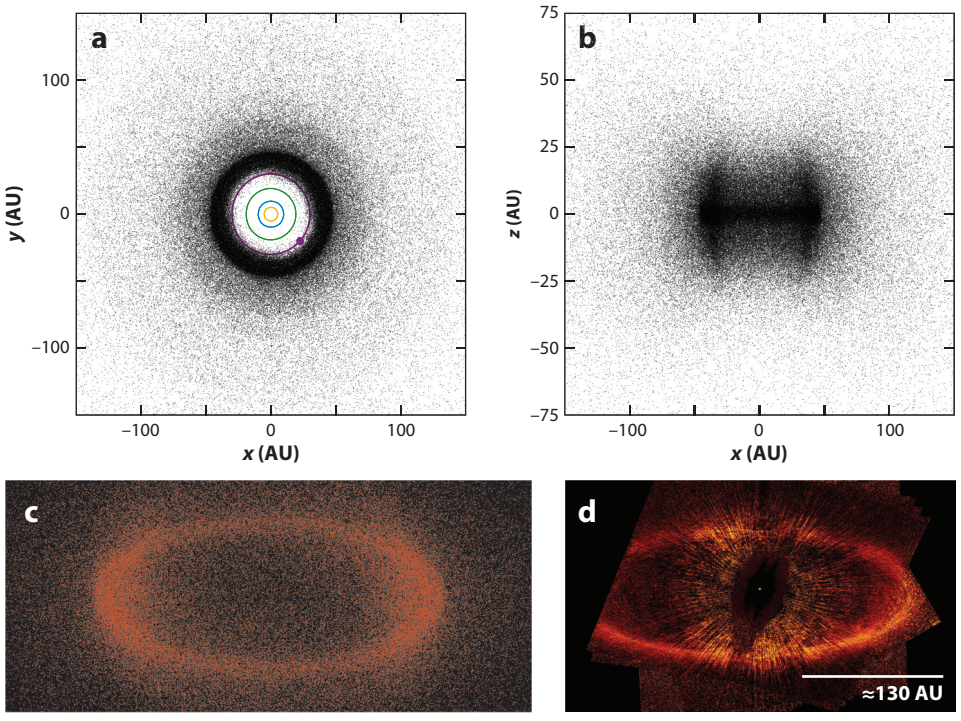

Figure 9

A comparison of an observationally calibrated TNO population model (Petit et al. 2011, Gladman et al. 2012) with the resolved Fomalhaut debris disk. The three model panels show projections of instantaneous TNO positions whose orbital distributions agree with the calibrated Canada-France Ecliptic Plane Survey, containing the inner, main, and outer classical belts along with resonant, detached, and scattering populations. (*a*) Face on projection, showing the four giant planet orbits for reference and a dot at Neptune's position. (*b*) Edge-on projection with a factor of two vertical exaggeration. The horns are real and due to the (limb brightened) contributions of the azimuthal cylinder of hot TNOs with perihelia and/or aphelia concentrated in the 35–47-AU range. (*c*) The same model scaled up in linear dimension by a factor of 2.2 and viewed obliquely at 25 deg to roughly match the scale and viewing angle of the Fomalhaut debris disk in panel *d*. (*d*) This final image shows light scattered by dust generated in a collisional ring as a classical Kuiper Belt analog; coronographic artifacts and scattering efficiency as a function of phase are present. Fomalhaut image in panel *d* reproduced from NASA, ESA, P. Kalas and J. Graham (University of California, Berkeley) and M. Clampin (NASA/GSFC). Abbreviation: TNO, transneptunian object.

large-scale planetesimal scattering has occurred, one expects many resonances to be populated, again subduing these azimuthal variations unless the dust can be preferentially migrated into just the *n*:2 resonances. Second, our TNO population shows a projected concentration when viewed edge on owing to the abundance of large-*i* perihelia in the main-belt distance range; we are unaware of any such feature yet seen in (rare) edge-on debris disks. This vertical structure is not evident in oblique view, which shows a version of our Kuiper belt scaled up only in heliocentric distance to match the size of the collisionally active ring in the Fomalhaut system (though we note that our main belt is not collisionally active enough to produce such a spectacularly observable dust ring).

6.2.3. An outer edge to the cold population. We consider now the important issue of an edge to our Solar System's protoplanetary disk, given what we know about our transneptunian

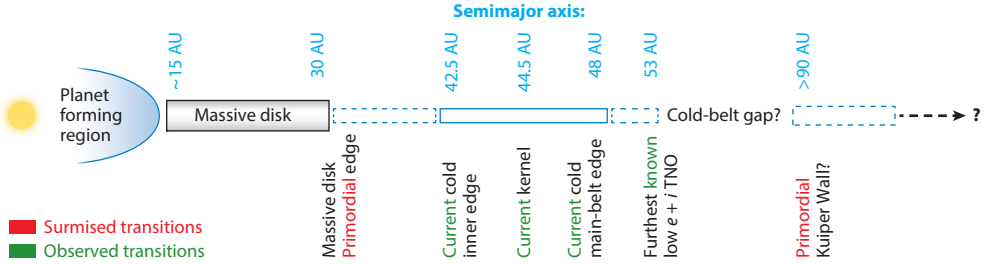

Figure 10

A schematic of surmised (*red*) and observed (*green*) transitions in the planetesimal population that formed from the Solar System's protoplanetary disk. The massive disk needs to abruptly ramp down at 30 AU to prevent Neptune's continued migration past its current location. Presumably a low-mass, small-body disk continued from this point, though all its members out to the current cold belt's inner-boundary starting at $a \simeq 42.5$ AU have been eliminated. Although a few known cold TNOs exist beyond the $a \simeq 48$ AU position of the 2:1 resonance, another drop in the surface density at this outer edge of the main belt is required. We modeled the possible extent of a cold-belt gap and find, based on the observed cold main-belt number density, a primordial depletion out to roughly 90 AU is required by the observations (assuming an a^{-1} radial falloff in number density and no change in cold-belt size distribution; a steeper falloff or decrease in the maximum size allows a closer Kuiper Wall). Abbreviation: TNO, transneptunian object.

population. The distinction between a gas-disk edge and a boundary interior to where a significant mass of solids have already condensed or concentrated may be important. The observed radial structures in protoplanetary disks (Section 6.2.1) indicate that dust is present in varying degrees, but whether this is due to more effective condensation or concentration of solids at some distances or more efficient incorporation of this dust into larger bodies is unclear. In either case, planetesimal building rates vary demonstrably at different distances.

We thus return to the issue of edges (or at least abrupt changes) in the radial distribution of the Solar System's primordial planetesimals. The evidence for a surface density discontinuity in the early Solar System at 30 AU relies entirely on the strong theoretical argument that if the surface density of small bodies had continued smoothly past this, Neptune should have just kept migrating outward (Section 5.1). A relatively sharp transition is required to halt Neptune's migration, although the transition's form is poorly constrained (Nesvorný et al. 2020 demonstrates a novel simulation weighting technique to investigate the edge's functional form). Observations of protoplanetary disks make it clear that the surface density profile is not monotonic, with both dips and bumps occurring, so perhaps the 30-AU feature is the end of such a bump; in scenarios such as this it becomes questionable whether reconstruction of the early surface density profile will ever be possible, especially when the complications of large-scale giant planet migration are included (along with the possibility of other now-lost giants!).

Proceeding outward (**Figure 10**), the fate of the early disk between 30 AU and the current inner cold-belt edge at 42.5 AU is unclear. It is thought that the current main-belt cold population is roughly representative of its original size distribution and surface density (Section 4.1). Thus, if there were initially ≈1,000 cold $D > 100$ km TNOs per astronomical unit (the current number per astronomical unit of the cold belt; Petit et al. 2011) covering the present position of the inner classical belt (36–39 AU), the 3:2 resonance would have efficiently captured them during Neptune's migration (e.g., Hahn & Malhotra 2005). This is inconsistent with the fact that the observed 3:2 members lack a cold component and have physical properties consistent with the hot TNO populations (Section 4.2.2). A plausible fix (Gladman et al. 2012) for this apparent inconsistency is that all cold-disk TNOs with $a < 42$ AU were pumped to Neptune-crossing orbits during the early planet-migration epoch when the ν_8 secular resonance migrated to its current position near 42 AU; this process could empty the inner classical belt's cold population before 3:2 resonance

capture can occur. In that scenario these removed cold primordial TNOs become a tiny fraction of the hot scattering population, and some should be reimplanted at other locations in the belt (e.g., Sheppard et al. 2012).

Starting at 42.5 AU, we enter the region (**Figure 10**) where extant cold TNOs constrain arguments. Under an assumption that the surface density of main-belt cold objects is primordial, the next drop in the surface density of the cold belt is at 44.5 AU (Morbidelli et al. 2008, Kavelaars et al. 2009), where the kernel (see the sidebar titled The Main-Belt Kernel) ends. This lower density then continues out to the current 48-AU location of Neptune's 2:1 resonance, where the observed cold population either has another (even larger) number-density drop or ends entirely. The coincidence of this transition with the resonance location is a long-standing puzzle. If one posits 48 AU as a primordial edge of where planetesimals condensed, there is no natural reason for this to coincide with the 2:1's final location after migration (which is set by the 30-AU boundary). The elegant solution (Levison & Morbidelli 2003) would be if the 2:1 dominates TNO emplacement for hot and cold populations alike, thus explaining the 2:1 as a marked outer edge. However, such a model does not implant a sufficiently low-e and low-i cold population and is inconsistent with current evidence that the cold population formed in situ (Section 5.1). It is possible that the true primordial edge was at 44.5 AU and that the cold population's extension out to 47 AU is due to the migrating 2:1 capturing $a < 44.5$ AU TNOs and then populating the 45–48-AU range in the final stages of migration via resonance drop-off (Gladman et al. 2012); this still would not eliminate the need for two edges in the primordial disk (one at 30 AU and one at 44.5 AU), but it would eliminate the uncomfortable coincidence of the outer edge with the 2:1. Perhaps the 44.5-AU drop is again a feature created by a pressure bump in the gas disk.

This leads us to the region beyond the 2:1. Early in TNO science, the entire 30–50-AU region (where the first TNOs were detected) was thought to be collisionally or dynamically depleted by 2–3 orders of magnitude relative to a hypothetical extension of the then estimated mass distribution needed to build the cores of the giant planets; Stern (1996) suggested that this depletion could be viewed as a Kuiper gap that was created at that distance range, and somewhere beyond 50 AU perhaps the depletion ceased and the planetesimals that accreted there would exist at the primordial solid surface mass-density in the protoplanetary disk. However, even with a moderate radial falloff of the disk's surface density, if the surface density were to come back up to a massive primordial level by $\simeq 75$ AU, detections beyond this "Kuiper Wall" (Chiang & Brown 1999, pg. 1421) should have already occurred in early surveys (c.f., Trujillo & Brown 2001). An ad hoc suggestion was that, though objects that formed closer than 50 AU had a size distribution extending up to dwarf-planet sizes, if for some reason only very small ($D < 30$ km) objects formed beyond the wall, the primordial mass density could still be present in a more distant disk (Chiang & Brown 1999).

This discussion changes qualitatively given our modern understanding of the TNO region. The past arguments above regarding expected numbers of objects assumed that any distant objects had the same size distribution as those then known (dominated by the hot population). Instead, the evidence now suggests that the cold population exhibits a maximum size cutoff and that it formed at its current low surface density. Therefore, most TNO detections beyond 50 AU are in fact of large hot objects that have invaded the region and thus have nothing to do with the primordial in situ density. Adopting instead the intrinsic number density of the cold main-belt TNOs, with a steep size distribution of only $H_r > 4.5$ TNOs, we used the OSSOS survey simulator (Lawler et al. 2018a) to show that, assuming an r^{-1} surface density profile, the cold Kuiper Wall could still be as close as 90 AU and remain undetected even with the vaster sky coverage of today's deep surveys; if the $a > 45$ AU solid surface density was steeper than r^{-1} or the brightest TNO locally built (e.g., by the streaming instability) becomes fainter than $H_r = 4.5$, the Kuiper Wall could move even closer and have escaped detection.

Regardless of whether a cold distant population remains undetected, the cold population's surface density must have dropped significantly somewhere in the range of 44.5–48 AU. If this is a primordial edge, a plausible idea would be that the young Solar System began to form planetesimals but then photoevaporation from a nearby O/B star boiled off the gas beyond 30 AU (see, e.g., the review by Adams 2010), removing the gas and the majority of the heavy elements (still in gas phase or in tiny gas-coupled solids) before they could be incorporated into planetesimals. Inside 30 AU, accretion could continue and build dwarf-planet and planetary objects, but in the primordial cold belt the density was (and is) so low that no additional accretion or collisional evolution happens over the Solar System's age. Although this scenario would explain the main-belt region and stop Neptune's migration, it provides no reason why there would be an additional depletion beyond the 2:1 resonance. It has also been suggested that a close stellar flyby could have stripped the disk, leaving an edge at 48 AU (reviewed by Morbidelli et al. 2008), but this process would have difficulty leaving the cold population intact.

If a distant outer classical cold belt is found at $\simeq$ 90 AU, a mechanism will be required to remove the cold population in the 50–90-AU region. The formation or the temporary implantation of a moderately sized planet near 80 AU has been suggested for the depletion beyond the 48-AU edge (Brunini & Melita 2002, Lykawka & Mukai 2008); if it was able to excite the cold population in that range, it might help solve the puzzle of overpopulated resonances from 55–100 AU (Section 4.2.2). No published model yet satisfactorily demonstrates this process. Alternatively, a gap here could potentially be created by inefficient dust condensation and/or planetesimal formation, like those suggested to be causing dust deficits in protoplanetary disks (Section 6.2.1), or extremely efficient inward dust migration that emptied this region with inefficient resupply from further out. The resolution of these puzzles will provide strong interplay between Solar System modeling and protoplanetary disk studies over the next decade.

6.3. Other Future Studies

The coming decade is likely to bring significant advancements to our understanding of the transneptunian region due to additional observational platforms. The *Vera C. Rubin Observatory*'s Legacy Survey of Space and Time (LSST) is expected to observe $\approx$30,000 TNOs (compared to the current inventory of $\approx$2,000), with the brighter TNOs being detected several hundred times over the course of its ten-year primary survey (Schwamb et al. 2018, Ivezić et al. 2019). Although the LSST survey strategy is still being defined (so the limiting magnitude and visit distribution are not yet known), it should dramatically increase the number of known TNOs with well-determined orbits. The wide sky coverage will also yield important constraints on the existence of new dynamical populations (Section 6.1.1). LSST also provides an opportunity to detect bright objects that have thus far evaded detection due to them previously being positioned against the $\approx$30-deg wide dense stellar background of the galactic plane. Cold classicals move $\approx$1 deg year^{-1}, so previously hidden objects within $\approx$50 AU will emerge to be discovered in the next decade or so.

Although LSST will provide some characterization of its TNO detections, much room remains for smaller follow-up programs. LSST will yield broadband color estimates for some of the brighter TNOs, though it will not be simultaneous and thus could be strongly affected by light curve effects; dedicated photometric studies (e.g., Schwamb et al. 2019) will still be required to generate high-precision observations needed to explore TNO surface properties. Similarly, LSST will provide only sparsely sampled light curves, making it difficult to identify binaries and contact binaries (see, e.g., Noll et al. 2020) or provide high-quality constraints on physical properties such as shapes and rotation periods. This is another area in which dedicated follow-up studies (e.g., Alexandersen et al. 2019) are needed. The LSST data will provide useful initial estimates for light curve and color properties that can be used to refine target lists for these dedicated programs.

Going beyond photometric colors to characterize TNOs requires spectroscopy. Currently high signal-to-noise ratio spectra can be obtained only for the brightest TNOs, but next-generation instruments such as the Near Infrared Spectrograph on the *James Webb Space Telescope* (JWST) can dramatically improve our understanding of TNO surface compositions. Many of the ices expected to be present on TNOs (based on predicted ice-line locations in the protoplanetary disk) have absorption bands in the near-IR, some of which will be observationally accessible for larger TNOs with JWST; see Métayer et al. (2019) for details and simulated example spectra. For TNOs that are not bright enough for spectra, JWST can provide photometry in a range of filters useful for diagnosing surface compositions (Parker et al. 2016). These studies will provide useful context for larger, ground-based photometric studies (e.g., **Figure 6**). We also note that ground-based spectroscopy capabilities (the current state of which for TNOs is reviewed by Barucci & Merlin 2020) will receive a huge boost from the increased apertures of the next generation of extremely large telescopes.

One of the more promising ways to refine the sizes and shapes of a sample of TNOs is through occultation studies, which can be done with small ground-based telescopes for occultations of reasonably bright stars. Accurate shapes and sizes of known $D \simeq$ 10–1,000-km TNOs have been obtained by measuring occultations along multiple chords. This was successful, for example, in gathering shape and size information on Arrokoth before the *New Horizons* spacecraft flyby (Buie et al. 2020). Occultations have also successfully detected debris rings around outer Solar System small bodies (reviewed by Sicardy et al. 2020). Predicting occultations of specific stars by specific TNOs requires both highly precise orbits and highly precise stellar catalogs. The precision of the latter has been dramatically improved by *Gaia* (Gaia Collab. et al. 2018), and the orbits of known TNOs will continue to improve in the LSST era. This should yield more accurate occultation predictions and thus more frequent occultation measurements. Building networks of small telescopes capable of recording these occultations (such as the network described by Buie & Keller 2016) will be important in increasing the number of TNOs with known sizes. Ortiz et al. (2020) provides a recent review of this topic.

Occultations can also detect rather than characterize TNOs. This will be an important tool for measuring the number density of very small TNOs (in the kilometer-scale size range), which are far too faint to be directly detected in reflected light. In these surveys, fields of stars are monitored for serendipitous occultation events (e.g., Roques et al. 1987, Zhang et al. 2013) to measure the on-sky density of very small TNOs. The Trans-Neptunian Automated Occultation Survey (TAOS II) uses multiple telescopes to reduce false positives and should operate over several years in the coming decade (Lehner et al. 2016). For such detections, tracking the objects to determine orbits is unlikely, but the occultation will yield an estimate of geocentric distance (which can be used to determine if an object is at transneptunian distances) and the latitude of the occulted star provides a limit on the detected object's inclination, allowing one to distinguish (to some extent) the hot and cold TNO populations.

Finally, there will still be significant room for important science from surveys that are deeper than LSST. Ground-based imaging facilities such as Subaru and the Canada-France-Hawaii Telescope can achieve significantly fainter limiting magnitudes (e.g., $m_r = 25.5$ in Sheppard et al. 2019 and Bannister et al. 2018) than typical LSST depths. Such surveys are critical for probing the faint-end H distribution of the main-belt and detached TNOs; unlike the scattering population, the larger q values in these populations mean that, even at perihelion, smaller TNOs ($H >$ 8–9, $D <$ 100 km) in these populations will be fainter than LSST's limiting magnitude. Results from smaller but deeper optical surveys will be important for constraining the size distributions, which are a critical test for TNO formation models.

SUMMARY POINTS

1. There are many detailed substructures in the transneptunian region that point to a rich and complex process of TNO creation and transplantation.
2. The low-e, low-i cold main-belt population formed in situ and has not experienced significant collisional evolution or dynamical depletion; this has led to new models of planetesimal formation in low surface density environments.
3. The $a > 50$ AU resonances are observed to be significantly more populated than predicted by any published model of a planetesimal population scattered outward and then sculpted by Neptune's migration; this remains an outstanding challenge.
4. The detached TNO population is very important for constraining arguments for past or present additional large planets; increasing the number of detections from characterized deep surveys will help tremendously to distinguish between models.
5. The $i > 40$ deg TNOs in the hot classical and detached populations, along with the polar corridor, imply that there are incompletely understood dynamical processes in the outer Solar System.
6. The reasons for various drops in the cold-disk's primordial surface density are unclear, but future advances in our understanding of planetesimal accretion with constraints from protostellar studies are promising.
7. LSST will make a major contribution to TNO science, but its moderate depth will not probe all TNO populations; other targeted science programs will be needed to resolve other questions (for example, pushing limits on the Kuiper Wall's possible inner edge and detecting the faint detached population).

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Jiaqing Bi, Peter Brown, Don Brownlee, Yukun Huang, Theodore Kareta, Nienke van der Marel, Scott Tremaine, Ewine van Dishoeck, and Mariana Yadkoo for input or providing comments on a draft version of this review. B.G. acknowledges support from NSERC of Canada, and K.V. acknowledges support from NASA and NSF (grants 80NSSC19K0785 and AST-1824869).

LITERATURE CITED

Adams FC. 2010. *Annu. Rev. Astron. Astrophys.* 48:47–85
Alexandersen M, Benecchi SD, Chen YT, et al. 2019. *Ap. J. Suppl.* 244:19
Alexandersen M, Gladman B, Kavelaars JJ, et al. 2016. *Astron. J.* 152:111
Allen RL, Bernstein GM, Malhotra R. 2001. *Ap. J. Lett.* 549:L241–44
Allen RL, Gladman B, Kavelaars JJ, et al. 2006. *Ap. J. Lett.* 640:L83–86
ALMA Partnership, Brogan CL, Pérez LM, et al. 2015. *Ap. J. Lett.* 808:L3
Andrews SM. 2020. *Annu. Rev. Astron. Astrophys.* 58:483–528
Ashton E, Gladman B, Kavelaars J, et al. 2020. *Icarus* 356:113793

Bannister MT. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 439–53. Amsterdam: Elsevier
Bannister MT, Gladman BJ, Kavelaars JJ, et al. 2018. *Ap. J. Suppl.* 236:18
Bannister MT, Kavelaars JJ, Petit J-M, et al. 2016. *Astron. J.* 152:70
Bannister MT, Shankman C, Volk K, et al. 2017. *Astron. J.* 153:262
Barucci MA, Merlin F. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 109–26. Amsterdam: Elsevier
Batygin K, Adams FC, Brown ME, Becker JC. 2019. *Phys. Rep.* 805:1–53
Batygin K, Brown ME. 2010. *Ap. J.* 716:1323–31
Batygin K, Brown ME. 2016. *Ap. J. Lett.* 833:L3
Becker JC, Khain T, Hamilton SJ, et al. 2018. *Astron. J.* 156:81
Bernardinelli PH, Bernstein GM, Sako M, et al. 2020. *Planet. Sci. J.* 1:28
Bernstein GM, Trilling DE, Allen RL, et al. 2004. *Astron. J.* 128:1364–90
Brasser R, Duncan MJ, Levison HF. 2006. *Icarus* 184:59–82
Brasser R, Duncan MJ, Levison HF, Schwamb ME, Brown ME. 2012. *Icarus* 217:1–19
Brasser R, Schwamb ME. 2015. *MNRAS* 446:3788–96
Brown ME. 2001. *Astron. J.* 121:2804–14
Brown ME. 2008. In *The Solar System Beyond Neptune*, ed. MA Barucci, H Boehnhardt, DP Cruikshank, A Morbidelli, R Dotson, pp. 335–44. Tucson, AZ: Univ. Ariz. Press
Brown ME. 2012. *Annu. Rev. Earth Planet. Sci.* 40:467–94
Brown ME, Bannister MT, Schmidt BP, et al. 2015. *Astron. J.* 149:69
Brown ME, Schaller EL, Fraser WC. 2011. *Ap. J. Lett.* 739:L60
Brown ME, Trujillo C, Rabinowitz D. 2004. *Ap. J.* 617:645–49
Brownlee D. 2014. *Annu. Rev. Earth Planet. Sci.* 42:179–205
Brunini A. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 225–47. Amsterdam: Elsevier
Brunini A, Melita MD. 2002. *Icarus* 160:32–43
Buie MW, Keller JM. 2016. *Astron. J.* 151:73
Buie MW, Porter SB, Tamblyn P, et al. 2020. *Astron. J.* 159:130
Chen YT, Gladman B, Volk K, et al. 2019. *Astron. J.* 158:214
Chen YT, Lin HW, Holman MJ, et al. 2016. *Ap. J. Lett.* 827:L24
Chiang E, Choi H. 2008. *Astron. J.* 136:350–57
Chiang EI, Brown ME. 1999. *Astron. J.* 118:1411–22
Chiang EI, Jordan AB. 2002. *Astron. J.* 124:3430–44
Chiang EI, Lovering JR, Millis RL, et al. 2003. *Earth Moon Planets* 92:49–62
Cohen CJ, Hubbard EC. 1965. *Astron. J.* 70:10–13
Davies JK, McFarland J, Bailey ME, Marsden BG, Ip W-H. 2008. In *The Solar System Beyond Neptune*, ed. MA Barucci, H Boehnhardt, DP Cruikshank, A Morbidelli, R Dotson, pp. 11–23. Tucson, AZ: Univ. Ariz. Press
Dawson RI, Murray-Clay R. 2012. *Ap. J.* 750:43
Di Ruscio A, Fienga A, Durante D, et al. 2020. *Astron. Astrophys.* 640:A7
Dones L, Brasser R, Kaib N, Rickman H. 2015. *Space Sci. Rev.* 197:191–269
Dones L, Weissman PR, Levison HF, Duncan MJ. 2004. In *Comets II*, ed. MC Festou, HU Keller, HA Weaver, pp. 153–74. Tucson, AZ: Univ. Ariz. Press
Doressoundiram A, Peixinho N, de Bergh C, et al. 2002. *Astron. J.* 124:2279–96
Duncan M, Quinn T, Tremaine S. 1987. *Astron. J.* 94:1330–38
Duncan MJ, Levison HF. 1997. *Science* 276:1670–72
Duncan MJ, Levison HF, Budd SM. 1995. *Astron. J.* 110:3073–81
Elliot JL, Kern SD, Clancy KB, et al. 2005. *Astron. J.* 129:1117–62
Fernández J. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 1–22. Amsterdam: Elsevier
Fernández JA, Ip W-H. 1984. *Icarus* 58:109–20

Fienga A, Di Ruscio A, Bernus L, et al. 2020. *Astron. Astrophys.* 640:A6
Fraser WC, Brown ME, Morbidelli AR, Parker A, Batygin K. 2014. *Ap. J.* 782:100
Gaia Collab., Brown AGA, Vallenari A, et al. 2018. *Astron. Astrophys.* 616:A1
Gallardo T. 2006. *Icarus* 181:205–17
Gladman B. 2005. *Science* 307:71–75
Gladman B, Chan C. 2006. *Ap. J. Lett.* 643:L135–38
Gladman B, Holman M, Grav T, et al. 2002. *Icarus* 157:269–79
Gladman B, Kavelaars J, Petit J-M, et al. 2009. *Ap. J. Lett.* 697:L91–94
Gladman B, Kavelaars JJ, Petit J-M, et al. 2001. *Astron. J.* 122:1051–66
Gladman B, Lawler SM, Petit J-M, et al. 2012. *Astron. J.* 144:23
Gladman B, Marsden BG, Vanlaerhoven C. 2008. In *The Solar System Beyond Neptune*, ed. MA Barucci, H Boehnhardt, DP Cruikshank, A Morbidelli, R Dotson, pp. 43–57. Tucson, AZ: Univ. Ariz. Press
Gomes R, Nesvorný D, Morbidelli A, Deienno R, Nogueira E. 2018. *Icarus* 306:319–27
Gomes RS. 2000. *Astron. J.* 120:2695–707
Gomes RS. 2003. *Icarus* 161:404–18
Gomes RS, Fernández JA, Gallardo T, Brunini A. 2008. In *The Solar System Beyond Neptune*, ed. MA Barucci, H Boehnhardt, DP Cruikshank, A Morbidelli, R Dotson, pp. 259–73. Tucson, AZ: Univ. Ariz. Press
Gomes RS, Gallardo T, Fernández JA, Brunini A. 2005. *Celest. Mech. Dyn. Astron.* 91:109–29
Gomes RS, Morbidelli A, Levison HF. 2004. *Icarus* 170:492–507
Greenstreet S, Gladman B, McKinnon WB. 2015. *Icarus* 258:267–88
Greenstreet S, Gladman B, McKinnon WB, Kavelaars JJ, Singer KN. 2019. *Ap. J. Lett.* 872:L5
Greenstreet S, Gladman B, Ngo H. 2020. *Astron. J.* 160:144
Hahn JM, Malhotra R. 2005. *Astron. J.* 130:2392–414
Hainaut OR, Boehnhardt H, Protopapa S. 2012. *Astron. Astrophys.* 546:A115
Holman MJ, Payne MJ, Fraser W, et al. 2018. *Ap. J. Lett.* 855:L6
Hughes AM, Duchene G, Matthews BC. 2018. *Annu. Rev. Astron. Astrophys.* 56:541–91
Ida S, Larwood J, Burkert A. 2000. *Ap. J.* 528:351–56
Ivezić Ž, Kahn SM, Tyson JA, et al. 2019. *Ap. J.* 873:111
Jewitt DC, Luu JX. 1995. *Astron. J.* 109:1867–76
Johansen A, Lambrechts M. 2017. *Annu. Rev. Earth Planet Sci.* 45:359–87
Johansen A, Oishi JS, Mac Low MM, et al. 2007. *Nature* 448:1022–25
Jones RL, Parker JW, Bieryla A, et al. 2010. *Astron. J.* 139:2249–57
Kaib NA, Sheppard SS. 2016. *Astron. J.* 152:133
Kavelaars JJ, Jones L, Gladman B, Parker JW, Petit J-M. 2008. In *The Solar System Beyond Neptune*, ed. MA Barucci, H Boehnhardt, DP Cruikshank, A Morbidelli, R Dotson, pp. 59–69. Tucson, AZ: Univ. Ariz. Press
Kavelaars JJ, Jones RL, Gladman BJ, et al. 2009. *Astron. J.* 137:4917–35
Kavelaars JJ, Lawler SM, Bannister MT, Shankman C. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 61–77. Amsterdam: Elsevier
Kenyon SJ, Bromley BC. 2020. *Planet. Sci. J.* 1:40
Kenyon SJ, Bromley BC, O'Brien DP, Davis DR. 2008. In *The Solar System Beyond Neptune*, ed. MA Barucci, H Boehnhardt, DP Cruikshank, A Morbidelli, R Dotson, pp. 293–313. Tucson, AZ: Univ. Ariz. Press
Khain T, Becker JC, Lin HW, et al. 2020. *Astron. J.* 159:133
Lagrange AM, Bonnefoy M, Chauvin G, et al. 2010. *Science* 329:57–59
Lan L, Malhotra R. 2019. *Celest. Mech. Dyn. Astron.* 131:39
Lawler SM, Gladman B. 2013. *Astron. J.* 146:6
Lawler SM, Kavelaars JJ, Alexandersen M, et al. 2018a. *Front. Astron. Space Sci.* 5:14
Lawler SM, Pike RE, Kaib N, et al. 2019. *Astron. J.* 157:253
Lawler SM, Shankman C, Kaib N, et al. 2017. *Astron. J.* 153:33
Lawler SM, Shankman C, Kavelaars JJ, et al. 2018b. *Astron. J.* 155:197
Lehner MJ, Wang SY, Reyes-Ruiz M, et al. 2016. *Proc. SPIE Conf. Ser.* 9906:99065M
Levison HF, Duncan MJ, Dones L, Gladman BJ. 2006. *Icarus* 184:619–33

Levison HF, Morbidelli A. 2003. *Nature* 426:419–21
Levison HF, Morbidelli A, Van Laerhoven C, Gomes R, Tsiganis K. 2008. *Icarus* 196:258–73
Levison HF, Stern SA. 2001. *Astron. J.* 121:1730–35
Li R, Youdin AN, Simon JB. 2019. *Ap. J.* 885:69
Luhman KL. 2014. *Ap. J.* 781:4
Luu J, Marsden BG, Jewitt D, et al. 1997. *Nature* 387:573–75
Luu JX, Jewitt DC. 2002. *Annu. Rev. Astron. Astrophys.* 40:63–101
Lykawka PS, Mukai T. 2005. *Earth Moon Planets* 97:107–26
Lykawka PS, Mukai T. 2007a. *Icarus* 186:331–41
Lykawka PS, Mukai T. 2007b. *Icarus* 192:238–47
Lykawka PS, Mukai T. 2008. *Astron. J.* 135:1161–200
Madigan AM, McCourt M. 2016. *MNRAS* 457:L89–93
Malhotra R. 1993. *Nature* 365:819–21
Malhotra R. 1995. *Astron. J.* 110:420–29
Malhotra R. 2019. *Geosci. Lett.* 6:12
Malhotra R, Volk K, Wang X. 2016. *Ap. J. Lett.* 824:L22
Matrà L, Marino S, Kennedy GM, et al. 2018. *Ap. J.* 859:72
Matrà L, Wyatt MC, Wilner DJ, et al. 2019. *Astron. J.* 157:135
Matthews BC, Kavelaars J. 2016. *Space Sci. Rev.* 205:213–30
McKinnon WB, Richardson DC, Marohnic JC, et al. 2020. *Science* 367:eaay6620
Meisner AM, Bromley BC, Kenyon SJ, Anderson TE. 2018. *Astron. J.* 155:166
Metayer R, Guilbert-Lepoutre A, Ferruit P, et al. 2019. *Front. Astron. Space Sci.* 6:8
Morbidelli A. 1997. *Icarus* 127:1–12
Morbidelli A, Bottke WF, Nesvorný D, Levison HF. 2009. *Icarus* 204:558–73
Morbidelli A, Levison HF. 2004. *Astron. J.* 128:2564–76
Morbidelli A, Levison HF, Gomes R. 2008. In *The Solar System Beyond Neptune*, ed. MA Barucci, H Boehnhardt, DP Cruikshank, A. Morbidelli, R Dotson, pp. 275–92. Tucson, AZ: Univ. Ariz. Press
Morbidelli A, Nesvorný D. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 25–59. Amsterdam: Elsevier
Morbidelli A, Thomas F, Moons M. 1995. *Icarus* 118:322–40
Müller T, Lellouch E, Fornasier S. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 153–81. Amsterdam: Elsevier
Murray CD, Dermott SF. 1999. *Solar System Dynamics*. Cambridge, UK: Cambridge Univ. Press
Murray-Clay RA, Chiang EI. 2005. *Ap. J.* 619:623–38
Murray-Clay RA, Chiang EI. 2006. *Ap. J.* 651:1194–208
Namouni F, Morais MHM. 2018. *MNRAS* 477:L117–21
Nesvorný D. 2011. *Ap. J. Lett.* 742:L22
Nesvorný D. 2015a. *Astron. J.* 150:68
Nesvorný D. 2015b. *Astron. J.* 150:73
Nesvorný D. 2018. *Annu. Rev. Astron. Astrophys.* 56:137–74
Nesvorný D, Brož M, Carruba V. 2015. In *Asteroids IV*, ed. P Michel, FE Demeo, WF Bottke, pp. 297–321. Tucson, AZ: Univ. Ariz. Press
Nesvorný D, Li R, Youdin AN, Simon JB, Grundy WM. 2019. *Nat. Astron.* 3:808–12
Nesvorný D, Vokrouhlický D. 2016. *Ap. J.* 825:94
Nesvorný D, Vokrouhlický D, Alexandersen M, et al. 2020. *Astron. J.* 160:46
Nesvorný D, Vokrouhlický D, Bottke WF, Levison HF. 2018. *Nat. Astron.* 2:878–82
Nesvorný D, Youdin AN, Richardson DC. 2010. *Astron. J.* 140:785–93
Noll K, Grundy WM, Nesvorný D, Thirouin A. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 201–24. Amsterdam: Elsevier
Noll KS, Grundy WM, Stephens DC, Levison HF, Kern SD. 2008. *Icarus* 194:758–68
O'dell CR, Wen Z. 1994. *Ap. J.* 436:194–202
Ortiz JL, Sicardy B, Camargo JIB, Santos-Sanz P, Braga-Ribas F. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 413–37. Amsterdam: Elsevier

Parker A, Pinilla-Alonso N, Santos-Sanz P, et al. 2016. *Publ. Astron. Soc. Pac.* 128:018010
Parker AH. 2015. *Icarus* 247:112–25
Parker AH, Kavelaars JJ. 2010. *Ap. J. Lett.* 722:L204–8
Parker AH, Kavelaars JJ. 2012. *Ap. J.* 744:139
Peixinho N, Thirouin A, Tegler SC, et al. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 307–29. Amsterdam: Elsevier
Petit J-M, Kavelaars JJ, Gladman BJ, et al. 2011. *Astron. J.* 142:131
Petit J-M, Kavelaars JJ, Gladman BJ, et al. 2017. *Astron. J.* 153:236
Pfalzner S, Vincke K. 2020. *Ap. J.* 897:60
Pike RE, Fraser WC, Schwamb ME, et al. 2017. *Astron. J.* 154:101
Pike RE, Kavelaars JJ, Petit J-M, et al. 2015. *Astron. J.* 149:202
Pike RE, Lawler SM. 2017. *Astron. J.* 154:171
Pinilla-Alonso N, Stansberry JA, Holler BJ. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 395–412. Amsterdam: Elsevier
Poppe AR, Lisse CM, Piquette M, et al. 2019. *Ap. J. Lett.* 881:L12
Roques F, Moncuquet M, Sicardy B. 1987. *Astron. J.* 93:1549–58
Schwamb ME, Brown ME, Fraser WC. 2014. *Astron. J.* 147:2
Schwamb ME, Fraser WC, Bannister MT, et al. 2019. *Ap. J. Suppl.* 243:12
Schwamb ME, Jones RL, Chesley SR, et al. 2018. arXiv:1802.01783
Shankman C, Gladman BJ, Kaib N, Kavelaars JJ, Petit J-M. 2013. *Ap. J. Lett.* 764:L2
Shankman C, Kavelaars JJ, Bannister MT, et al. 2017a. *Astron. J.* 154:50
Shankman C, Kavelaars JJ, Lawler SM, Gladman BJ, Bannister MT. 2017b. *Astron. J.* 153:63
Sheppard SS, Ragozzine D, Trujillo C. 2012. *Astron. J.* 143:58
Sheppard SS, Trujillo C. 2016. *Astron. J.* 152:221
Sheppard SS, Trujillo CA, Tholen DJ, Kaib N. 2019. *Astron. J.* 157:139
Sheppard SS, Udalski A, Trujillo C, et al. 2011. *Astron. J.* 142:98
Sicardy B, Renner S, Leiva R, et al. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 249–69. Amsterdam: Elsevier
Silsbee K, Tremaine S. 2018. *Astron. J.* 155:75
Singer KN, McKinnon WB, Gladman B, et al. 2019. *Science* 363:955–59
Spencer JR, Stern SA, Moore J-M, et al. 2020. *Science* 367:eaay3999
Stansberry J, Grundy W, Brown M, et al. 2008. In *The Solar System Beyond Neptune*, ed. MA Barucci, H Boehnhardt, DP Cruikshank, A Morbidelli, R Dotson, pp. 161–79. Tucson, AZ: Univ. Ariz. Press
Stern SA. 1991. *Icarus* 90:271–81
Stern SA. 1996. *Astron. J.* 112:1203–10
Tegler SC, Romanishin W. 2000. *Nature* 407:979–81
Tegler SC, Romanishin W, Consolmagno GJ. 2003. *Ap. J. Lett.* 599:L49–52
Thommes EW, Duncan MJ, Levison HF. 1999. *Nature* 402:635–38
Tiscareno MS, Malhotra R. 2009. *Astron. J.* 138:827–37
Torbett MV. 1989. *Astron. J.* 98:1477–81
Trujillo CA. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 79–105. Amsterdam: Elsevier
Trujillo CA, Brown ME. 2001. *Ap. J. Lett.* 554:L95–98
Trujillo CA, Sheppard SS. 2014. *Nature* 507:471–74
Tsiganis K, Gomes R, Morbidelli A, Levison HF. 2005. *Nature* 435:459–61
Tychoniec Ł, Manara CF, Rosotti GP, et al. 2020. *Astron. Astrophys.* 640:A19
van der Marel N, Dong R, di Francesco J, Williams JP, Tobin J. 2019. *Ap. J.* 872:112
Van Laerhoven C, Gladman B, Volk K, et al. 2019. *Astron. J.* 158:49
Volk K, Malhotra R. 2017. *Astron. J.* 154:62
Volk K, Malhotra R. 2019. *Astron. J.* 158:64
Volk K, Murray-Clay RA, Gladman BJ, et al. 2016. *Astron. J.* 152:23
Volk K, Murray-Clay RA, Gladman BJ, et al. 2018. *Astron. J.* 155:260

Weryk RJ, Lilly E, Chastel S, et al. 2016. arXiv:1607.04895
Winn JN, Fabrycky DC. 2015. *Annu. Rev. Astron. Astrophys.* 53:409–47
Wyatt M. 2020. In *The Trans-Neptunian Solar System*, ed. D Prialnik, MA Barucci, L Young, pp. 351–76. Amsterdam: Elsevier
Youdin AN, Goodman J. 2005. *Ap. J.* 620:459–69
Young SS, Karr A. 2011. *Significance* 8:116–20
Yu TYM, Murray-Clay RA, Volk K. 2018. *Astron. J.* 156:33
Zderic A, Collier A, Tiongco M, Madigan AM. 2020. *Ap. J. Lett.* 895:L27
Zhang ZW, Lehner MJ, Wang JH, et al. 2013. *Astron. J.* 146:14