




































# The Zwicky Transient Facility Bright Transient Survey. II. A Public Statistical Sample for Exploring Supernova Demographics\*

Daniel A. Perley<sup>1</sup> , Christoffer Fremming<sup>2</sup> , Jesper Sollerman<sup>3</sup> , Adam A. Miller<sup>4,5</sup> , Aishwarya S. Dahiwalé<sup>2</sup>,  
Yashvi Sharma<sup>2</sup> , Eric C. Bellm<sup>6</sup> , Rahul Biswas<sup>7</sup> , Thomas G. Brink<sup>8</sup> , Rachel J. Bruch<sup>9</sup> , Kishalay De<sup>2</sup> ,  
Richard Dekany<sup>10</sup> , Andrew J. Drake<sup>2</sup>, Dmitry A. Duev<sup>2</sup> , Alexei V. Filippenko<sup>8,11</sup> , Avishay Gal-Yam<sup>12</sup> , Ariel Goobar<sup>7</sup> ,  
Matthew J. Graham<sup>2</sup>, Melissa L. Graham<sup>6</sup> , Anna Y. Q. Ho<sup>2,8,13</sup> , Ido Irani<sup>12</sup> , Mansi M. Kasliwal<sup>2</sup> , Young-Lo Kim<sup>14</sup> ,  
S. R. Kulkarni<sup>2</sup> , Ashish Mahabal<sup>2,15</sup> , Frank J. Masci<sup>16</sup> , Shaunak Modak<sup>8</sup> , James D. Neill<sup>2</sup> , Jakob Nordin<sup>17</sup> ,  
Reed L. Riddle<sup>10</sup> , Maayane T. Soumagnac<sup>9,18</sup> , Nora L. Strotjohann<sup>12</sup> , Steve Schulze<sup>12</sup> , Kirsty Taggart<sup>1</sup> ,  
Anastasios Tzanidakis<sup>2</sup> , Richard S. Walters<sup>2</sup>, and Lin Yan<sup>10</sup> 

<sup>1</sup> Astrophysics Research Institute, Liverpool John Moores University, Liverpool Science Park, 146 Brownlow Hill, Liverpool L3 5RF, UK; [d.a.perley@ljmu.ac.uk](mailto:d.a.perley@ljmu.ac.uk)

<sup>2</sup> Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

<sup>3</sup> Department of Astronomy, The Oskar Klein Center, Stockholm University, AlbaNova, SE-10691 Stockholm, Sweden

<sup>4</sup> Center for Interdisciplinary Exploration and Research in Astrophysics and Department of Physics and Astronomy, Northwestern University, 1800 Sherman Avenue, Evanston, IL 60201, USA

<sup>5</sup> The Adler Planetarium, Chicago, IL 60605, USA

<sup>6</sup> DIRAC Institute, Department of Astronomy, University of Washington, 3910 15th Avenue NE, Seattle, WA 98195, USA

<sup>7</sup> Department of Physics, The Oskar Klein Center, Stockholm University, AlbaNova, SE-10691 Stockholm, Sweden

<sup>8</sup> Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA

<sup>9</sup> Department of Particle Physics and Astrophysics, Weizmann Institute of Science, 234 Herzl Street., Rehovot 76100, Israel

<sup>10</sup> Caltech Optical Observatories, California Institute of Technology, Pasadena, CA 91125, USA

<sup>11</sup> Miller Senior Fellow, Miller Institute for Basic Research in Science, University of California, Berkeley, CA 94720, USA

<sup>12</sup> Benoziyo Center for Astrophysics, The Weizmann Institute of Science, Rehovot 76100, Israel

<sup>13</sup> Miller Institute for Basic Research in Science, University of California, Berkeley, CA 94720, USA

<sup>14</sup> Université de Lyon, Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, F-69622, Villeurbanne, France

<sup>15</sup> Center for Data Driven Discovery, California Institute of Technology, Pasadena, CA 91125, USA

<sup>16</sup> IPAC, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA

<sup>17</sup> Institut für Physik, Humboldt-Universität zu Berlin, Newtonstr. 15, D-12489 Berlin, Germany

<sup>18</sup> Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

Received 2020 September 2; revised 2020 September 28; accepted 2020 September 29; published 2020 November 19

## Abstract

We present a public catalog of transients from the Zwicky Transient Facility (ZTF) Bright Transient Survey, a magnitude-limited ( $m < 19$  mag in either the  $g$  or  $r$  filter) survey for extragalactic transients in the ZTF public stream. We introduce cuts on survey coverage, sky visibility around peak light, and other properties unconnected to the nature of the transient, and show that the resulting statistical sample is spectroscopically 97% complete at  $< 18$  mag, 93% complete at  $< 18.5$  mag, and 75% complete at  $< 19$  mag. We summarize the fundamental properties of this population, identifying distinct duration–luminosity correlations in a variety of supernova (SN) classes and associating the majority of fast optical transients with well-established spectroscopic SN types (primarily SN Ibn and II/Iib). We measure the Type Ia SN and core-collapse (CC) SN rates and luminosity functions, which show good consistency with recent work. About 7% of CC SNe explode in very low-luminosity galaxies ( $M_i > -16$  mag), 10% in red-sequence galaxies, and 1% in massive ellipticals. We find no significant difference in the luminosity or color distributions between the host galaxies of SNe Type II and SNe Type Ib/c, suggesting that line-driven wind stripping does not play a major role in the loss of the hydrogen envelope from their progenitors. Future large-scale classification efforts with ZTF and other wide-area surveys will provide high-quality measurements of the rates, properties, and environments of all known types of optical transients and limits on the existence of theoretically predicted but as yet unobserved explosions.

*Unified Astronomy Thesaurus concepts:* [Supernovae \(1668\)](#); [Catalogs \(205\)](#); [Surveys \(1671\)](#); [Transient sources \(1851\)](#); [Time domain astronomy \(2109\)](#)

## 1. Introduction

Recent years have brought an unprecedented expansion in our ability to survey for transient astronomical phenomena. The optical sky is now being scanned on an almost nightly basis by several different telescope networks around the world, including the All-Sky Automated Survey for Supernovae (ASAS-SN; Shappee et al. 2014), the Asteroid Terrestrial Last-Alert System (ATLAS; Tonry et al. 2018), and the Zwicky Transient Facility

(ZTF; Bellm et al. 2019a, 2019b; Graham et al. 2019; Dekany et al. 2020). New projects, such as BlackGEM (Groot 2019) and the Gravitational-wave Optical Transient Observer (GOTO; Dyer et al. 2018), are beginning operations, and earlier surveys such as the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS; Kaiser et al. 2002), continue. In 2019 almost 20,000 new and unique optical transients were reported via official channels (Kulkarni 2020), an increase of two orders of magnitude from a decade prior (Gal-Yam et al. 2013). The Legacy Survey of Space and Time (LSST; Ivezić et al. 2019) at the Vera C. Rubin Observatory is

\* An interactive catalog with all data used in this paper is available at <https://sites.astro.caltech.edu/ztf/bts> and is updated in real time.

expected to increase these numbers by another order of magnitude within a few years.

Large numbers of transients are of limited scientific value without secure classifications and redshifts (Kulkarni 2020). Despite recent advances in photometric classification (e.g., Muthukrishna et al. 2019; Villar et al. 2019, 2020; Dauphin et al. 2020; Hosseinzadeh et al. 2020), the only ground truth for this remains spectroscopy, an observationally expensive endeavor. Deciding which transients to spectroscopically classify and which to ignore typically involves extensive human decision-making, potentially introducing complex biases and diminishing the value of large statistical samples for studies of (for example) volumetric rates, luminosity functions, or ensemble host-galaxy properties.

In Fremling et al. (2020) we introduced the ZTF Bright Transient Survey (BTS), which aims to provide a large and purely magnitude-limited ( $m < 19$  mag for discovery and  $m < 18.5$  mag for classification<sup>19</sup>) sample of extragalactic transients in the northern sky, suitable for detailed statistical and demographic analysis. In that work we described some of the aims of the project and presented early results on the fraction of supernovae (SNe) at this magnitude level hosted by galaxies with known, cataloged redshift (44%), along with a catalog of the first 761 SNe found by the project.

The BTS is an ongoing effort that will continue in its current form through the end of the public ZTF Northern Sky Survey (Bellm et al. 2019b) in October 2020. Work is ongoing to provide final photometric and spectroscopic data releases (and associated scientific papers) spanning this entire period. Preliminary photometric and spectroscopic data are also released in real time on a nightly basis via the ZTF brokers<sup>20</sup> and the Transient Name Server (TNS<sup>21</sup>). In this paper, which is accompanied by an online web portal, we supplement these basic, continuous data releases with a live catalog of higher-level properties of our sample measured from the real-time public data—in particular, peak luminosities, rise and decay times, and host-galaxy associations—and demonstrate the use of the sample for a variety of scientific aims.

The paper is structured as follows. In Section 2 we describe additional improvements to the filtering and screening process implemented since 2018, and detail a series of post-facto selection cuts that we employ to remove poorly observed transients and variables without imposing selection biases on the remaining sample. Section 3 provides spectroscopic completeness statistics for the resulting subset, demonstrating that it is  $\sim 93\%$  complete for transients peaking above  $m < 18.5$  mag. In Section 4 we highlight how in only two years ZTF has mapped out a vast swathe of the observational transient parameter space, providing the largest and most reliable look at the diversity of luminous transient phenomena in the universe. We also provide preliminary characterizations of the SN luminosity function, the core-collapse (CC) SN rate, and a color–magnitude analysis of host-galaxy properties of the major SN classes. We summarize our work in Section 5 and

provide additional documentation of the BTS Sample Explorer, a public webpage that serves our real-time transient catalog.

## 2. BTS Sample Selection and Characterization

### 2.1. Alert Filter

The fundamental criteria for inclusion of an event in the BTS are that it is a genuine transient (a well-defined *event* with a beginning and end, as distinct from a variable star or active galactic nucleus (AGN) for which changes in flux are always occurring), that it is extragalactic, and that it is brighter than  $m < 19$  mag in either the *g* or *r* filter at some point in its observed ZTF light curve. On any given night, genuine transients at this magnitude level within the ZTF difference-image alert stream (Masci et al. 2019; Patterson et al. 2019) are vastly outnumbered by variables (stars and AGNs), artifacts, and moving objects. These must be efficiently filtered out without losing any of the genuine transients the survey seeks to catalog.

All ZTF Avro<sup>22</sup> packets contain two machine-learning scores to aid this process: a real–bogus score (`rb_score`; Mahabal et al. 2019) to separate PSF-like sources from artifacts in ZTF subtractions, and a star–galaxy separation score (`sg_score`; Tachibana & Miller 2018), based on a cross-match with the Pan-STARRS (Chambers et al. 2016) catalog, to aid in the rejection of stars. Neither metric is perfect, and producing a transient catalog free of variables and artifacts requires substantial additional effort using both software filters and human attention.

Our first-year in-stream software filter, which reduced the  $\sim 10^6$  Avro alert packets produced each night to  $\sim 500$  viable transient candidates, was first described by Fremling et al. (2020). The cuts employed by this filter were relatively basic: two detections at least 0.02 days ( $\sim 30$  minutes) apart, a high `rb_score`, no underlying counterpart ( $< 2''$ ) with high `sg_score`, no bright star in the vicinity ( $< 20''$ ), and a difference magnitude brighter than  $m < 19$ . While this filter had the benefit of being straightforward to implement and understand, the false-positive rate was significant (several hundred stars/AGNs would pass the filter nightly) and on rare occasions it would miss SNe near bright galactic nuclei mistakenly flagged as stars by `sg_score`. Thus, beginning in June 2019 we have made several additional adjustments.

Events with a long history of previous detections coincident with a bright PS1 or Gaia source are now rejected, since these tend to be AGNs flagged as galaxies. The exclusion radius around bright stars has been reduced (to an extent depending on the star’s brightness and color) to reduce the risk of rejecting objects around galaxies mistakenly flagged as stars. We also remove slow-moving asteroids using a catalog cross-match and employ the new deep-learning real–bogus `drb_score` (Duev et al. 2019) to better remove artifacts. A full list of changes is given in Appendix A. These adjustments reduced the typical number of false positives from several hundred per night to  $\lesssim 50$ , reducing (but not entirely eliminating) the need for human vetting.

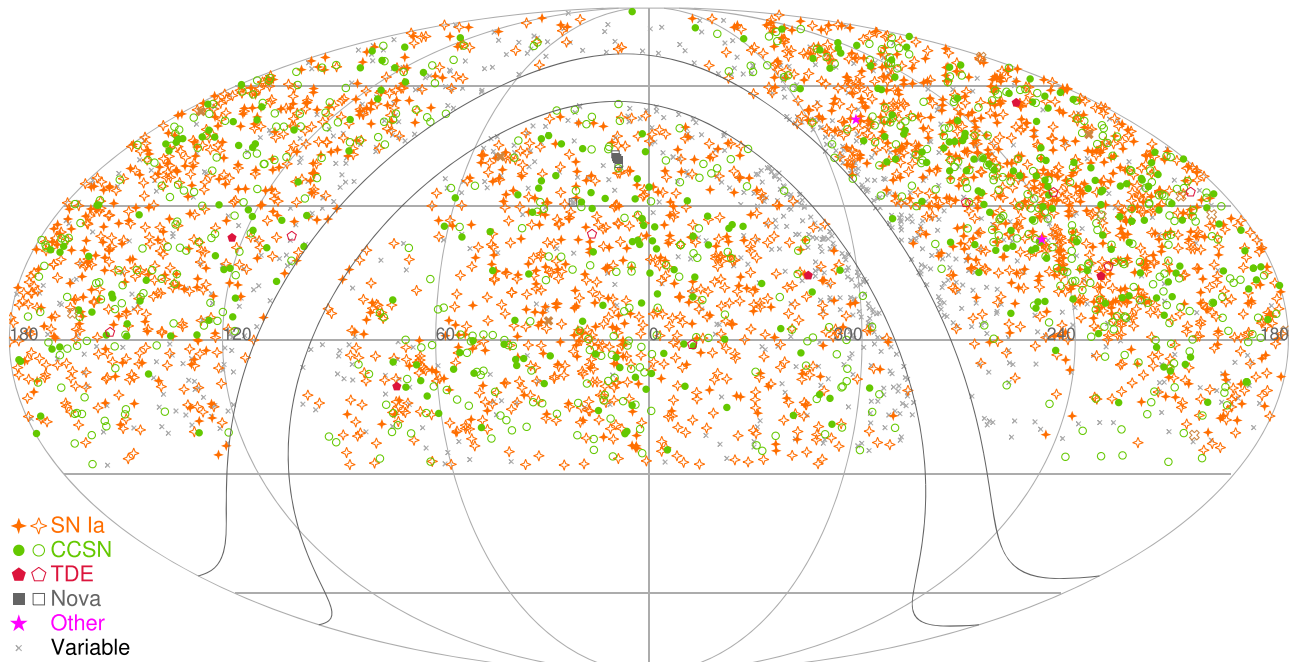
The BTS filter runs in parallel on the GROWTH Marshal (Kasliwal et al. 2019) and on AMPEL (Nordin et al. 2019). Candidates passing the filter (and not already saved to the BTS program) are reviewed by human scanners nightly using the

<sup>19</sup> A parallel volume-limited survey, the ZTF Census of the Local Universe (CLU) experiment, extends the classification threshold to  $m < 20$  mag for transients occurring in known galaxies within  $D < 200$  Mpc (De et al. 2019, 2020).

<sup>20</sup> Currently operating, fully featured public brokers include ANTARES (<https://antares.noao.edu/>), LASAIR (<https://lasair.roe.ac.uk/>), ALERCE (<http://alerce.science/>), and MARS (<https://mars.lco.global/>).

<sup>21</sup> <https://wis-tns.weizmann.ac.il>

<sup>22</sup> <https://avro.apache.org>



**Figure 1.** Skymap of classified transients and variables within BTS. Transients at  $m_{\text{peak}} < 18.5$  mag satisfying our sample cuts (described in Sections 2.3 and 2.4) are shown as solid symbols; other transients are shown as open symbols. BTS covers the entire northern sky outside the Galactic plane, although there is some variation in transient sky-areal density as a result of seasonal variations in survey coverage.

GROWTH Marshal scanning tool, as described by Fremling et al. (2020). Candidates assessed to definitely be AGNs, variable stars, artifacts, or other false positives are ignored; the remaining candidates are registered to the BTS program within the GROWTH Marshal database (“saved”). Typically 5–10 candidates are saved on an average clear night. A skymap of saved events with classifications is shown in Figure 1.

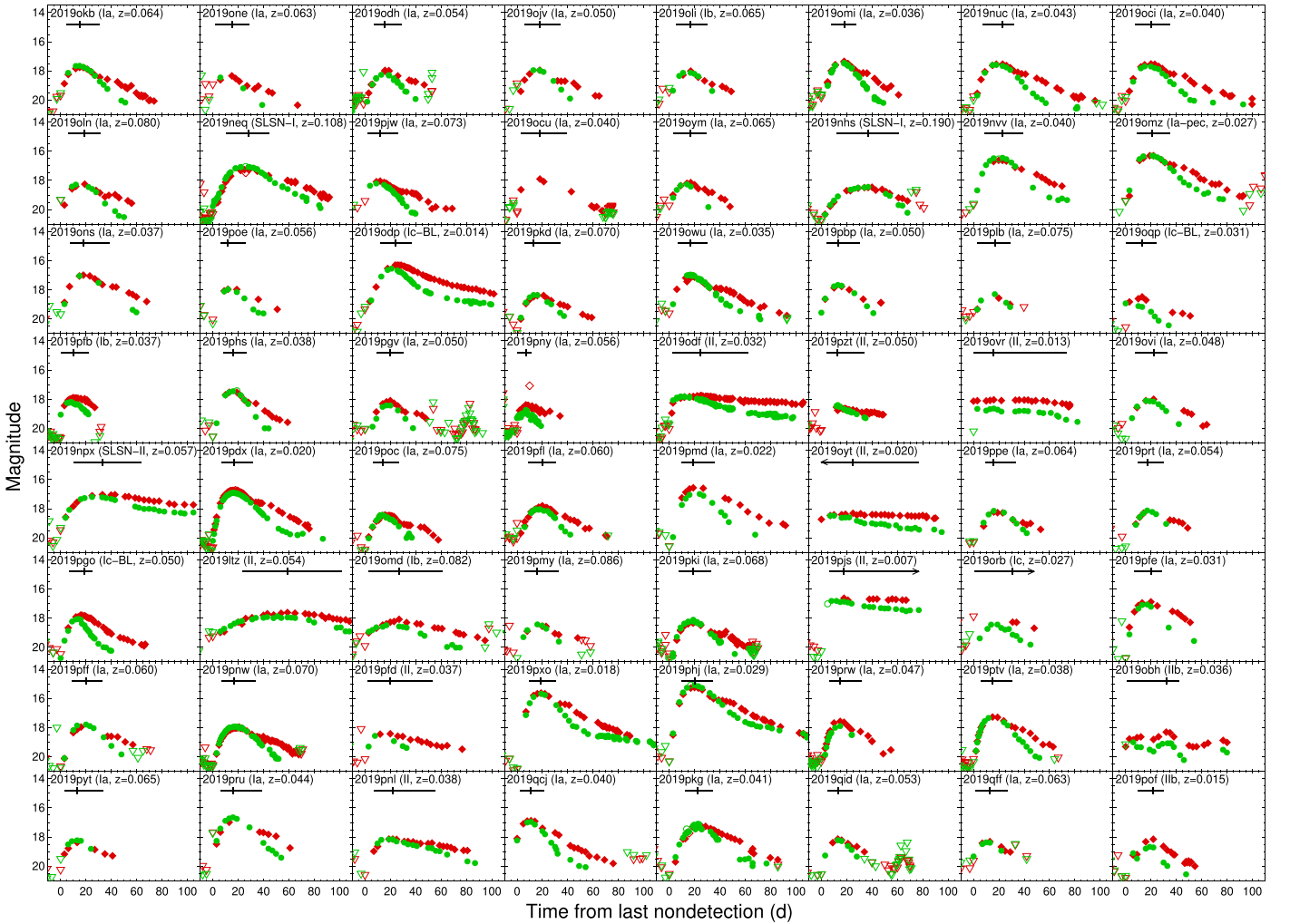
## 2.2. Candidate Characterization

Once saved, candidates are subjected to additional scrutiny: visual inspection of the full alert-based light curve and cross-matches to various catalogs and imaging data. This is sometimes enough to classify false positives (for example, if the transient is coincident with a known AGN or a Wide-field Infrared Survey Explorer source with AGN-like colors and the light curve shows normal AGN variability, or if a faint star-like counterpart is visible in the image and previous flares from this location are seen in the ZTF light curve or reported on TNS). Otherwise, the candidate is reported to TNS and spectroscopic follow-up observations with the Spectral Energy Distribution Machine (SEDM; Blagorodnova et al. 2018; Rigault et al. 2019) on the Palomar 60 inch telescope or the Spectrograph for the Rapid Acquisition of Transients (SPRAT; Piascik et al. 2014) on the Liverpool Telescope (Steele et al. 2004) are requested according to our priority-ranking system (Fremling et al. 2020) and reviewed on an approximately weekly basis. Targets that cannot be classified with these facilities are scheduled for observations during scheduled classical observing runs at larger telescopes—in particular with the Double-Beam Spectrograph (DBSP; Oke & Gunn 1982) at the Palomar 5 m Hale telescope, with the Dual Imaging Spectrograph (DIS) at Apache Point Observatory, the Kast spectrograph on the Shane 3 m telescope at Lick Observatory (Miller & Stone 1993), and occasionally with the Low-Resolution Imaging Spectrometer (LRIS; Oke et al. 1995) at the 10 m

Keck I telescope on Maunakea. Spectra are reduced and resulting classifications are registered with TNS, generally within 24 hr of observations. If a classification is first reported to TNS by another group or acquired by another ZTF program unrelated to our project, we avoid duplicating effort unless there is an indication that the external classification is unreliable. (See Fremling et al. 2020 for additional details and statistics on classifications from the first year of the project.)

All saved candidates are also automatically analyzed by an independent script that continuously downloads the public Avro packet data for every transient in our program. Light curves are built from the packet data using the `jd` (time), `fid` (filter), `magpsf` (magnitude), `sigmapsf` (uncertainty), and `diffmaglim` (field limiting magnitude) values; examples are shown in Figure 2. Measurements in poor observing conditions (indicated by limiting magnitudes shallower than 19) are ignored. For both the *g*-band and *r*-band light curves, we measure the time (JD) and magnitude of observed maximum light as the brightest measurement in the relevant light curve. The “rise time” and “fade time” are calculated as the time elapsed in days between this peak and the point where the light curve drops to 0.75 mag below peak (equivalent to half the peak flux); this is calculated using simple linear interpolation between data points. Upper limits from the alert-packet history are used only if they occur before the first detection, in which case the interpolation is performed between the limit and the first detection in the light curve.<sup>23</sup> This calculation is run separately in each filter, but because filter coverage is often irregular (e.g., the rise may be sampled in the

<sup>23</sup> Limits in packet data refer to the limiting magnitude for empty regions of the whole ZTF field, and only designate that an alert was not generated. This usually means a true nondetection, but can also indicate a data-quality issue, leading to apparent nondetections occasionally being interspersed throughout a light curve.



**Figure 2.** Example two-filter ZTF light curves for all events that passed our quality cuts and that peaked at  $m < 18.5$  mag during a 22-day period in summer 2019 (JD 2458727.8–2458749.8). Green circles are  $g$  band and red diamonds are  $r$  band. Open triangles signify upper limits and unfilled circles/diamonds are low-quality measurements not used in the light-curve measurements. The cross symbol at upper left of each panel shows the rise, peak, and fade times (the left end, crossbar, and right end of the cross, respectively) measured using the technique as defined in Section 2.2.

$g$  band but the decline only in the  $r$  band) we also calculate a hybrid timescale measurement by shifting the “fainter” band (as defined at peak) to match the “brighter” one at its peak, and using whichever band has the fastest evolution in each direction. (This hybrid measurement is the value used in the cuts and subsequent analysis discussed later.)

We also obtain the Galactic extinction along the line of sight using the NASA Extragalactic Database (NED) extinction tool<sup>24</sup> (based on the dust map of Schlafly & Finkbeiner 2011). If the redshift of the event has been measured, we calculate the peak (observed) absolute magnitude by applying the extinction correction, the distance modulus calculated from the redshift assuming a cosmological model with  $\Omega_M = 0.3$ ,  $\Omega_\Lambda = 0.7$ , and  $h = 0.7$  (peculiar-motion corrections are ignored, although for transients in M31 and M33 we use standard distance moduli of 24.4 and 24.55 mag, respectively), and apply a uniform  $K$ -correction of  $2.5 \log_{10}(1+z)$ . This estimate will be close to the real absolute magnitude for transients passing our quality cuts and with characteristic timescales longer than several days, but is a lower limit for transients for which the peak was not sampled or for very fast events. Note also that redshifts

measured from SN features have significant uncertainty ( $\Delta z \approx 0.005$ ; Fremling et al. 2020) and the impact of peculiar motions (up to  $\Delta v \approx 600 \text{ km s}^{-1}$ , or  $\Delta z \approx 0.002$ ) can also be significant at low redshift, so there is additional scatter in these measurements for nearby transients (up to 0.5 mag at  $z \approx 0.02$  if the redshift is not precisely known, or 0.2 mag if a host-galaxy redshift is available). About 6% of our classified sample is at  $z < 0.02$ ; while in the majority of these cases ( $\sim 75\%$ ; see also Fremling et al. 2020) precise host-galaxy redshifts are available, this uncertainty should be kept in mind when dealing with low-luminosity populations or in interpreting luminosity outliers.

We additionally cross-match the location of the transient to a variety of catalogs. We use the nearest three PS1 cross-matches from the Avro packets (which contain `sgscore` values). We also download the list of potential cross-matches registered on the event summary page on Lasair (Smith et al. 2019). Lasair also provides star/galaxy classifications, although we employ this information only for objects brighter than  $g < 21$  or  $r < 21$  mag. If a cataloged star exists within  $1''$  of the candidate we associate it with the (probable) star. Otherwise, we calculate the most likely galaxy counterpart using a chance-probability calculation: we exclude stars and choose the cross-match for

<sup>24</sup> [https://ned.ipac.caltech.edu/extinction\\_calculator](https://ned.ipac.caltech.edu/extinction_calculator)

which the probability of finding a galaxy as bright (or brighter) as close (or closer) to the transient position by random chance is lowest (Appendix C). This is done independently for both cross-match lists and the results are used for our purity cut (Section 2.4).

To provide additional counterpart photometry for purposes of host-galaxy analysis, we also download the complete Sloan Digital Sky Survey (SDSS) and Pan-STARRS (PS1) photometric catalogs for sources near the transient location and calculate the most likely cross-match in this region. Stars and galaxies are distinguished using the SDSS star-galaxy code, or the PS1 PSF-Kron magnitude method.<sup>25</sup> The method for identifying the probable counterpart association is similar to the Avro/Lasair cross-match method above, although we also use the transient classification to restrict the search to exclude matches of inappropriate type (e.g., potential stellar associations are excluded for classified SNe). The cross-match is based only on the SDSS/PS1 photometry; spectroscopic information or other cross-matched catalogs are not yet considered. The search is currently limited to within  $<90''$  and  $<30$  kpc (projected) of the transient location to minimize false positives, although these restrictions will be loosened in the future as more catalogs are added. If there is no source meeting these criteria the event is designated as hostless and an upper limit is calculated using the relevant limiting magnitudes of the survey.

### 2.3. Quality Filter: Removal of Poorly Observed Candidates

Our general aim is to acquire spectra of every event that *could* be an SN and that is brighter than  $m \leq 18.5$  mag at any point in its evolution when observed by the survey, regardless of any other properties of the object. However, owing to poor weather, seasonal gaps, or a suboptimal sky location at the time of discovery, some events may be especially difficult to characterize even if they are bright.

These limitations are unavoidable for an all-sky, ground-based survey, but their impact on the completeness and quality of our survey can be minimized by introducing additional, post-facto cuts to remove events discovered at times and sky locations that were heavily affected by coverage gaps or poor observability. Care must be taken that these cuts are unbiased with respect to the properties of the transient to the maximum extent possible. In particular, they should be minimally dependent on duration: for example, requiring a large number of detections would introduce a strong bias against short-timescale transients.

Our current set of quality cuts is as follows.

1. The transient must have an observation significantly prior to peak light. Specifically, P48 must have observed the field at least once between 7.5 and 16.5 days prior to the time when the brightest detection in the light curve was recorded. The transient need not be detected in this observation, but the observation must be deep enough to be constraining ( $m_{\text{lim}} > 19$  mag).
2. The transient must have two observations around the time of peak light. Specifically, in addition to the observation at (apparent) maximum, the transient must have a second observation either 2.5–7.5 days before *or* 2.5–7.5 days after this measurement. These measurements must also have  $m_{\text{lim}} > 19$  mag.
3. The transient must have an observation after peak light. Specifically, it must have an observation between 7.5 and 16.5 days after the observed time of maximum, or alternatively an observation 2.5–7.5 days after maximum *and* an observation between 16.5 and 28.5 days after maximum. These measurements must also have  $m_{\text{lim}} > 19$  mag.
4. The location of the transient in the sky must be conducive to follow-up spectroscopy. Specifically, it must remain above  $30^\circ$  elevation for at least 2 hr at  $>12^\circ$  twilight during the night occurring 30 days after the observed time of peak light.
5. The transient must not be present in its reference image. Specifically, there must be at least a 30 day span between the last exposure in its reference image and the first registered detection in its light curve.
6. Galactic extinction toward the transient should be low ( $A_V < 1.0$  mag).
7. At least one packet in the alert history must pass the most recent version of the alert-stream filter (2.1), even if the candidate was saved under an earlier version of the filter.

Cuts 1–3 limit the sample to events for which the peak time and luminosity are well constrained, and approximately weekly cadence is maintained between two weeks prior to peak and two weeks after peak. Cuts 4–6 remove events that present other types of difficulties (e.g., hard to obtain spectra, photometry is contaminated, large/uncertain extinction correction). The final cut is not strictly a quality cut but is applied for consistency. For the purposes of this paper, we additionally restrict the sample to transients with a time of peak between 2018 June 1 (the public start of the BTS survey) and 2020 July 15, inclusive.

We emphasize that these criteria make no reference to the timescale or behavior of the transient itself. The only assumption is that a well-defined single peak does exist and can be recognized based on three or more observations spanning two weeks around this peak. Transients with durations shorter than the permissible coverage gaps (about 10 days) will be mildly<sup>26</sup> selected against and transients with durations shorter than the survey cadence itself (3 days) will be heavily selected against. Very slow transients with durations comparable to the survey (2 yr to date) will also be preferentially missed. However, events with durations between  $\sim 10$  and 100 days should be selected with equal efficiency as long as they have only a single peak.

Multi-peaked events or events with an extended, flat plateau are somewhat *more* likely to be selected than single-peaked events, since they have multiple opportunities for the peak to fall in a window that passes our criteria. These events are relatively rare, so this effect is not large, although a detailed measurement of duration-dependent rates would require additional corrections.

Together these cuts remove about half (48%) of the sample. About 25% of the down-selection can be attributed to weather and instrument-downtime gaps, 12% to other coverage gaps

<sup>25</sup> <https://outerspace.stsci.edu/display/PANSTARRS/How+to+separate+stars+and+galaxies>

<sup>26</sup> We use the term “mildly” selected against to refer to types of events that may be undercounted by a factor of up to two; events that are “heavily” selected against may be undercounted by a factor of more than two. Exact selection losses will depend sensitively on light-curve shape and peak magnitude as well as on duration. A detailed quantification of selection losses in these regimes is beyond the scope of this paper but will be addressed in future work using simulations.

(lunar, seasonal, or scheduling), 5% to visibility, 5% to reference imaging, and 1% to Galactic extinction. Significant losses of this type are unavoidable, given our goal to establish a universal set of criteria for demographic studies of transients spanning a wide variety of potential behaviors. Studies focused on a narrow range of events would be able to achieve a higher yield, at the expense of this generality, by employing different criteria tuned to the anticipated properties of the sample of interest.

#### 2.4. Purity Filter: Removal of False Positives

It is inevitable that many candidates saved by the human scanners eventually turn out to be variables, rather than genuine transients: typically, AGNs or cataclysmic variable stars (CVs). While these are not part of our project, they do pass the filter routinely and cannot always be distinguished from a transient, given the information available at the time of scanning. Some are observed and classified spectroscopically by our observing programs, but whenever possible photometric and contextual information is used to preserve scarce spectroscopic resources for genuine extragalactic transients.

AGNs can often be recognized immediately after being saved via cross-matches to pre-existing spectroscopic surveys or to multiwavelength (radio, X-ray, mid-IR) catalogs. Alternatively, it is possible to eliminate them photometrically after further monitoring: continuous slow variability (both upward and downward) lasting more than a year—with no evidence of a return to the pre-detection flux seen in the reference image—is also considered sufficient for a photometric-only classification.

CV eruptions (dwarf novae) usually exhibit a distinctive fast rise to peak followed by a somewhat slower decay during which a constant, blue  $g - r$  color is maintained, often followed by a sudden drop; no known class of SN shows this behavior. Dwarf novae are also expected to have a star-like counterpart, or no detectable counterpart, in PS1 reference imaging. Candidates matching all of these criteria are considered to be securely classified as dwarf novae and are removed from the sample, even in the absence of spectroscopy. Repeated flares (including flares in other surveys) separated by long stretches of inactivity are an even more definitive indicator of a CV origin.

However, it is not always possible to obtain a definitive assessment of a false positive from the light curve alone: the short durations of dwarf novae in particular make it easy to miss one or more of the key phases above during brief gaps in coverage, introducing a potential to be confused on an individual basis with fast-evolving SNe. Given our goal to produce a purely magnitude-limited, unbiased, and highly complete transient sample, it is still desirable to separate these from genuine transients in a systematic way.

To remove highly probable false positives, we apply an additional criterion (which we will refer to as a purity cut): *either* the potential transient must be coincident with a cross-matched galaxy (but not with its nucleus) *or* its light curve must have a vaguely SN-like timescale. The cross-match criterion is documented in Appendix B. To satisfy the light-curve criterion, the rise time (see Section 2.2) must be  $<120$  days and the fade time must be  $<200$  days but more than  $\sim 11$  days<sup>27</sup> (the region within the dashed line in Figure 3). This selection in principle imposes a small bias against very

fast transients in undetected galaxies or very compact galaxies (which mimic CVs), or against extremely slow transients in galactic nuclei (which mimic AGNs). However, it is essentially unbiased to the parameter space occupied by the vast majority of real extragalactic transients.

The impact of our purity cut is visualized in Figure 3. Nearly all spectroscopic SNe have SN-like light-curve properties by our definition (although there are a few exceptions). Nearly all unclassified events with SN-like properties are associated with galaxies. Since we only require that a candidate meets one of these criteria, this means that it will be very rare for a genuine transient to fail the purity cut, presuming its light curve is well sampled and free of bad measurements.

In practice, among events with classifications (and which pass the quality cuts), the purity cut removed only two classified transients ( $<0.2\%$  of the sample): one SN for which the timescale measurements were compromised by sparse data, and one (probable) nova in the far outskirts of M31 with an erratically flaring light curve. Notably, none of the known fast transients failed the purity cut because all are associated with galaxies. We also visually inspected all unclassified  $m < 18.5$  mag events that passed the quality cuts but failed the purity cut, and confirmed that none of them is likely to be an SN or some other extragalactic transient.

### 3. Sample Completeness

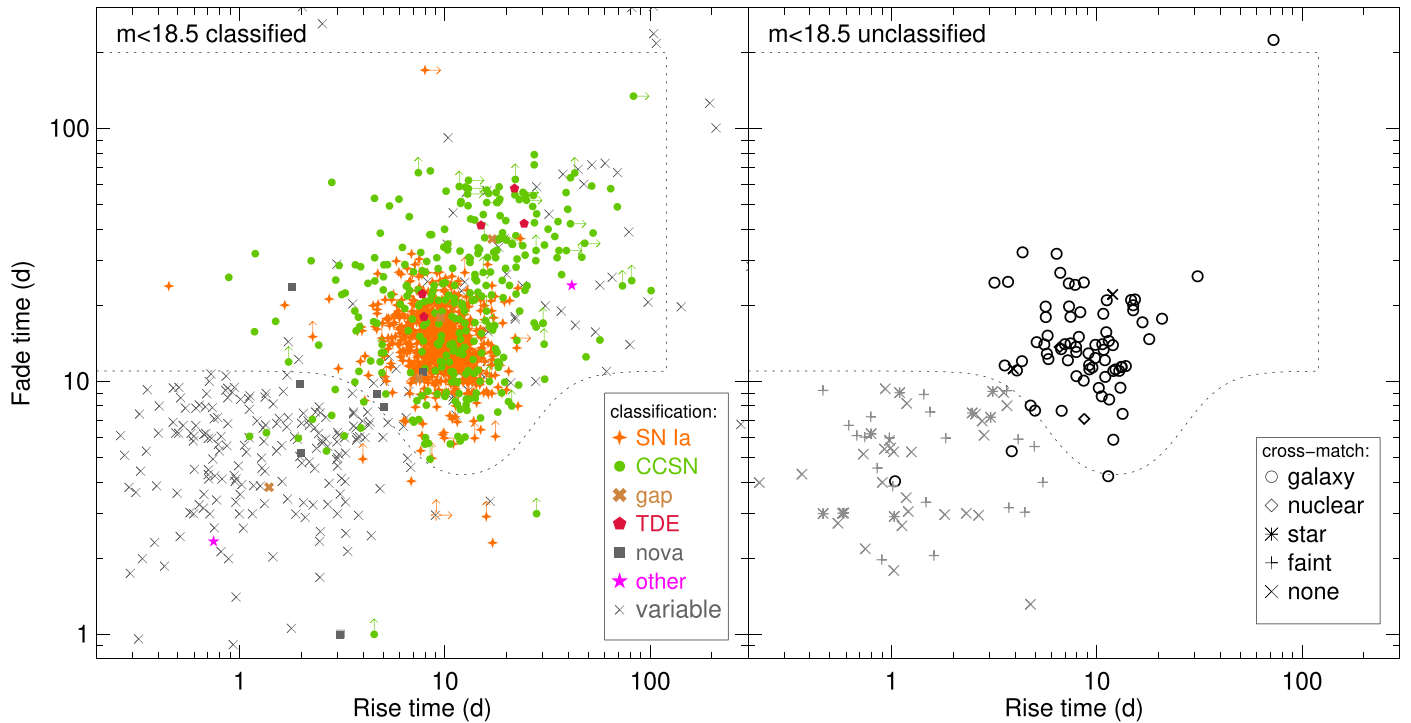
The BTS project registered 3147 spectroscopically classified transients during the 25.5 month period of this study, 1865 of which satisfy our quality and purity cuts (a partial breakdown of these by type is given in Table 1). Given the stated goal of our project to provide a spectroscopically complete magnitude-limited sample of extragalactic transients, we also need to know how many genuine transients met our selection criteria but could not be classified.

To check this we performed several tests. First, we took every event saved to the program that satisfies the quality and purity cuts and is not classified as a star/AGN (or other false positive); we refer to this population as the “statistical sample.” A histogram of these objects by peak magnitude is shown in Figure 4. The overall distribution of these events (solid black line) is consistent with the  $\Delta N \propto f^{3/2}$  power law predicted for a flux-limited survey in a homogeneous, Euclidean universe, except for the faintest bin ( $m > 18.9$  mag) where we anticipate being incomplete. This indicates that we are saving the expected numbers of events independent of their magnitude.

To quantify our success at spectroscopically classifying this population, a histogram of classified events is also shown in Figure 4. Expressed in cumulative terms (dotted line in the upper panel of Figure 4), 93% of events passing our selection cuts at  $m < 18.5$  mag have successful, public classifications. Completeness improves to 97% at  $m < 18$  mag and 100% at  $m < 17$  mag. Since we do not systematically target events fainter than  $m > 18.5$  mag, classification completeness drops sharply beyond this point: to 85% at  $m < 18.75$  mag and 75% at  $m < 19$  mag.

Inspection of the light curves and locations of unclassified events (at  $m < 18.5$  mag) confirms that most are likely to be ordinary SNe with properties that generally reflect the demographics of the rest of the sample (i.e., most appear to be SNe Ia) and were missed solely because no spectrum could be obtained, or one was obtained and had low quality, usually in association with periods of bad weather (Appendix D). A

<sup>27</sup> The limit on fade time is slightly less stringent if the rise time is in the range of normal SNe. The exact equation for the lower limit on fade time is  $t_{\text{fade}} > 4 + 7/(1 + e^{(\log_{10}(5) + \log_{10}(t_{\text{rise}})/0.1)}) + 7/(1 + e^{\log_{10}(30) - \log_{10}(t_{\text{rise}})/0.1})$ .



**Figure 3.** Rise time (half-peak to peak) vs. fade time (peak to half-peak) for  $m_{\text{peak,obs}} < 18.5$  mag events in BTS that pass the quality filter. Events with public classifications (which are always spectroscopic for SNe/transients) are shown in the left panel. Transients and variables cluster in different regions of the diagram, with some overlap. The right panel shows events we were unsuccessful at classifying, which exhibit a similar bimodal distribution. The symbol indicates the nature of the cross-matched catalog object. Events that *either* pass the light-curve cuts indicated by the dashed line *or* have a credible cross-matched galaxy and are not coincident with its nucleus pass our purity cut. At  $m < 18.5$  mag, 93% of such events are classified and only 7% are unclassified. A small number of outlier SNe Ia with apparent very fast rise/fade times are present. This is usually due to erroneous upper limits associated with the failure of the transient to generate an alert, but can also be due to additional photometric scatter from subtraction residuals in bright, point-like galactic nuclei.

few unclassified events have peculiar light curves that do not resemble any known SN type, including a handful with only single-night detections that may be particularly fast transients (but could also be the result of data-quality issues affecting the rest of the light curve). A discussion of these will be deferred until forced photometry is available at the end of the survey.

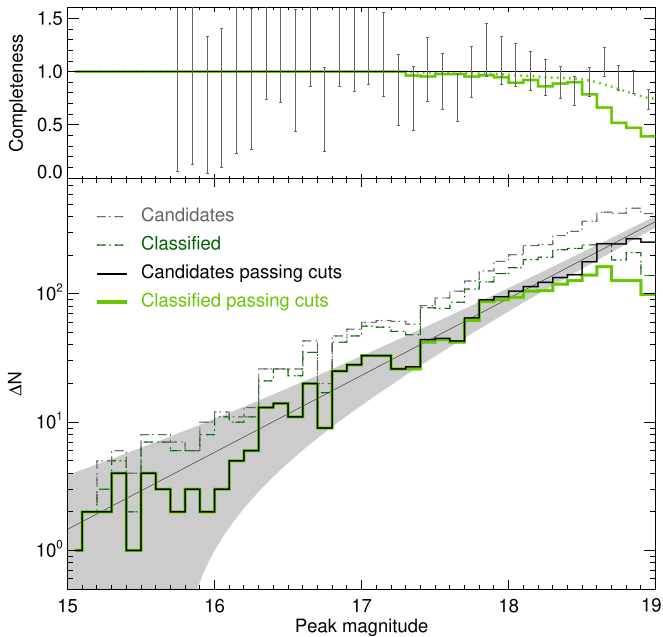
To check for events we may have failed to save to our program in the first place, we downloaded the entire catalog of classified TNS transients reported between 2018 January 1 and 2020 August 12 with decl.  $\delta > -30^\circ$  (4350 in total). Of these, 1015 were not in the BTS catalog. We cross-matched the coordinates against the complete database of public-program ZTF Avro alerts using the GROWTH Marshal, and for all matches we checked how many had exceeded 19 mag in ZTF public data in more than two observations. After removing matches that were not actually SNe we found 75 missing transients. Most of these either (a) passed the filter on only a single night owing to extremely sparse coverage and/or a light-curve peak just above  $m = 19$  mag, (b) occurred during the spring 2018 science validation period before the formal public survey began, (c) were contaminated by SN light in the reference image or by a coincident foreground star, or (d) were close to bright foreground stars. These would, by design, not have passed our alert filter and/or our selection cuts. We did identify 13 SNe that nominally satisfy our selection goals (eight at  $m < 18.5$  mag and five at  $18.5 < m < 19$  mag). Five passed the alert filter but were not saved by scanners. The remaining eight did not pass the filter because the nuclei of their host galaxies were incorrectly treated as stars owing to a high or ambiguous  $sgscore$ . These represent possible examples of incompleteness (with those in the latter category also

**Table 1**  
Classification Totals

Class	All	Passing cuts	Passing cuts and $m < 18.5$
Transients	3147	1865	1206
SN Ia	2232	1352	875
SN CC	878	490	313
H-rich	671	357	226
II	516	273	171
IIb	45	28	15
IIc	89	45	32
SLSN-II	21	11	8
H-poor	207	133	87
Ib/Ic	141	86	51
Ic-BL	27	21	17
Ibn	11	9	8
SLSN-I	28	17	11
TDE	13	8	5
Gap	11	5	4
Novae	11	8	7
Other	2	2	2
Unclassified	1596	627	82

**Note.** Totals include only public classifications available on TNS or other open sources and should be considered preliminary, pending reanalysis. “Gap” transients include Ca-rich events, luminous red novae, intermediate-luminosity red transients, and LBV eruptions. “Other” transients are AT2018cow and AT2019cmw, which have extensive spectroscopy but resemble no well-established transient type and likely belong to new categories of object.

potentially imposing some bias, since they are close to bright galactic nuclei). However, they represent a very small fraction of the overall sample total (0.6% of  $m_{\text{peak}} < 18.5$  mag



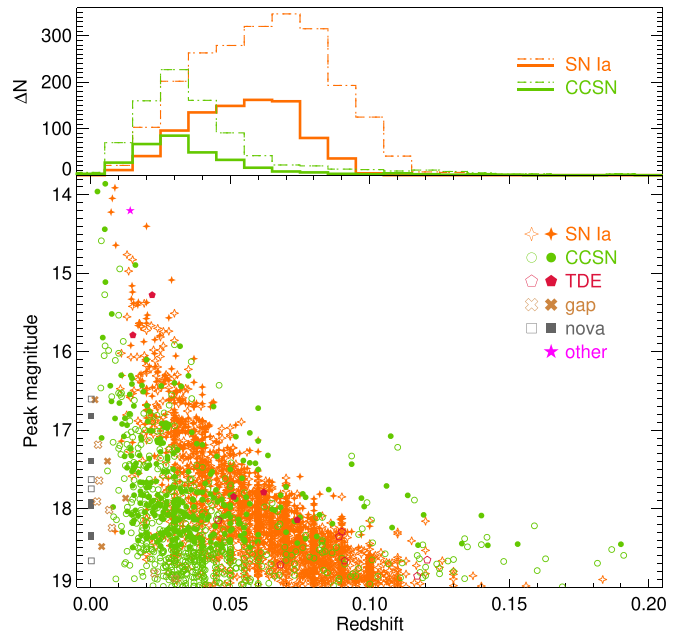
**Figure 4.** Histograms of transient candidate and confirmed transient counts by magnitude. The dashed histogram lines show all events saved to the survey and not known to be variables. The solid histogram lines show only events passing our sample cuts. A fit assuming a simple  $\frac{dN}{d \log f} \propto f^{3/2}$  power law, with  $2\sigma$  prediction intervals for each magnitude bin, is also illustrated. Completeness fractions are shown in the upper panel. Error bars indicate recovery completeness for saved sample transients with respect to the  $f^{3/2}$  prediction (95% Poisson confidence interval); the green line indicates spectroscopic completeness for saved transients passing sample cuts (the solid line shows completeness per bin, the dotted line shows cumulative completeness down to a particular limiting magnitude).

statistical-sample transients). Even accounting for the likelihood that some additional bright SNe may have been missed by both us *and* the broader community, we expect any impact on our scientific goals to be minimal.

The above checks suggest that our program has largely succeeded in reaching its goal of producing a magnitude-limited sample with no significant biases relating to the duration, behavior, or host-galaxy environment of the transient.

The statistics above are appropriate for “ordinary” SNe of the type that dominate the overall transient rate. We do expect to fare less well in other circumstances. Fast events that mimic CVs (and are either not in galaxies or are in galactic nuclei), SNe coincident ( $\lesssim 1''$ ) with AGNs, and long-timescale transients from the central regions of galaxies that resemble AGNs are all unlikely to be classified by other observers for the same reasons that they are much more likely to be missed by our selection process or excluded by our sample cuts. Certain rare transient categories are particularly likely to be heavily impacted: fast and luminous transients at high redshift such as on-axis afterglows of gamma-ray bursts (GRBs), or transients specific to galactic nuclei such as tidal disruption events (TDEs). Finding these events effectively and studying their demographics in an unbiased way requires different selection methods. Parallel efforts within ZTF focusing on these populations are ongoing; these are described in other works (van Velzen et al. 2019, 2020; Ho et al. 2020a).

Additionally, as previously noted, the limited cadence of the ZTF public survey itself results in a milder bias against the shortest-duration transients ( $< 10$  days, and especially  $< 3$  days)



**Figure 5.** Apparent magnitude and redshift distribution of classified transients in the BTS sample. Symbol conventions are the same as in Figure 1. A histogram by redshift for SNe Ia and CC SNe is shown in the top panel; thick lines show  $m < 18.5$  mag transients satisfying all sample cuts and dashed lines indicate other classified transients. Our survey probes CC SNe out to approximately  $z < 0.05$  and SNe Ia out to  $z < 0.1$ , and superluminous events beyond these limits.

and longest-duration transients ( $> 200$  days). Although we are still sensitive to transients with these properties, additional corrections would be necessary to accurately calculate their rates or study their demographics in a complete sense.

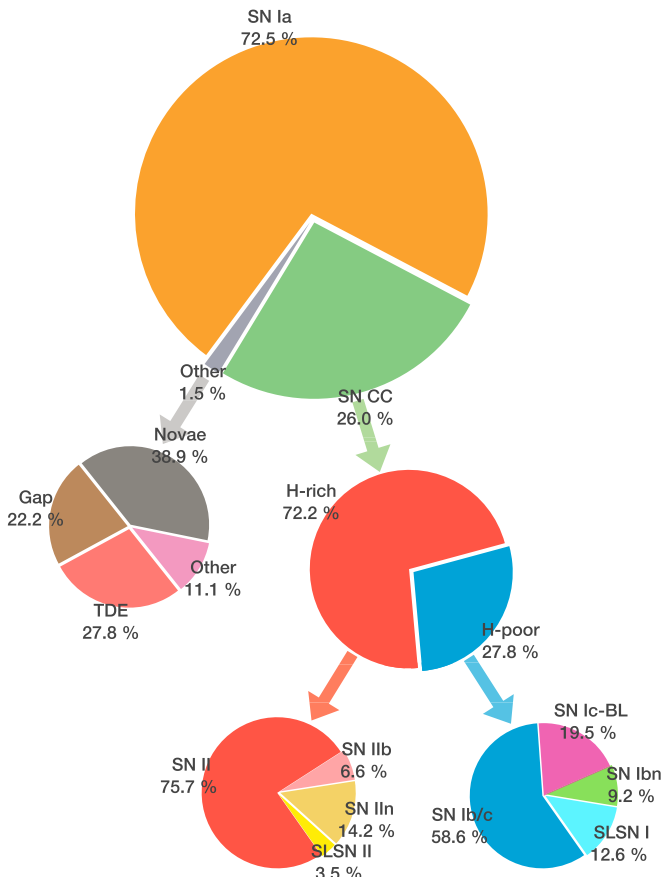
The redshift and magnitude distribution of the sample, before and after applying the selection cuts, is presented in Figure 5. A breakdown of transients by classification category is presented in Table 1 and Figure 6.

## 4. Results

The BTS catalog is the first large, highly complete, untargeted sample of transients for which spectroscopy and high-quality light curves are simultaneously available. This provides opportunities to examine the complete observational parameter space occupied by these explosions, and to examine correlations between key parameters of interest, without being limited by selection bias.

A complete investigation of all potential scientific uses of this sample is beyond the immediate scope of this paper, and will be reserved for a variety of follow-up works once the first phase of ZTF is finished and complete forced-photometry light curves are available, along with final spectroscopic (re)classifications. A complete and impartial spectroscopic analysis will be essential for uncommon or easily confused subclasses in particular (e.g., Ia-CSM versus II<sub>n</sub>, II versus II<sub>b</sub>, or Ib versus Ic). For the vast majority of transients, however, neither the light curves nor the classifications are expected to change significantly. In this section we will make use of the existing data products to provide a preliminary exploration of a variety of topics to demonstrate the scientific capabilities of the sample.





**Figure 6.** Pie charts showing fractional number counts in the  $m < 18.5$  mag BTS statistical sample (1206 events in total) within various categories and subcategories.

#### 4.1. The Landscape of Stellar Death

Kasliwal (2011) summarized the state of knowledge of transient “parameter space” (in luminosity and characteristic timescale), highlighting advances provided by wide-field surveys in discovering events with properties different from those of typical SNe: very luminous events (SLSNe), events intermediate in luminosity between novae and typical SNe (“gap” transients), and very fast events. More recently, Villar et al. (2017) approached the issue theoretically, providing a physical explanation for the luminosity and timescale distributions of many types of known and predicted transients.

We are now in a position to provide an unbiased look at this topic using a complete sample of real transients. This is provided in Figure 7, calculated using the BTS sample. Filled points show events that pass the quality and purity cuts and peak at  $m \leq 18.5$  mag. We also show other events as open circles as long as they either pass the quality/purity cuts *or* have a useful measurement of both their rise and fade times even in the presence of poorer sampling or a fainter peak, though this supplementary sample is not unbiased and the associated measurements typically have larger uncertainties. The timescale (rest-frame time above half-maximum light, calculated by adding the rise and fade times and dividing by the time dilation factor of  $1 + z$ ) and peak luminosity are calculated using the basic interpolation method described in Section 2.2. Events that only have lower limits on their timescales have been omitted from the diagram if the limit is not “constraining” in comparison to the bulk of the SN

population ( $>16$  days), *unless* the rise time has been measured and is  $<8$  days (i.e., unless there is reason to think it is an actual fast transient whose decay was not well captured, rather than a transient with a sparse light curve). A small number of SNe with data-quality issues identified by manual inspection were also removed.

The region of the diagram with characteristics of typical SNe (timescales of about a month and absolute magnitude close to  $-18$ ) is extremely well populated, as expected. However, smaller numbers of events do populate the diagram in all directions except the longest durations ( $\gg 100$  days), to which we are not yet sensitive because of the limited duration of the survey.

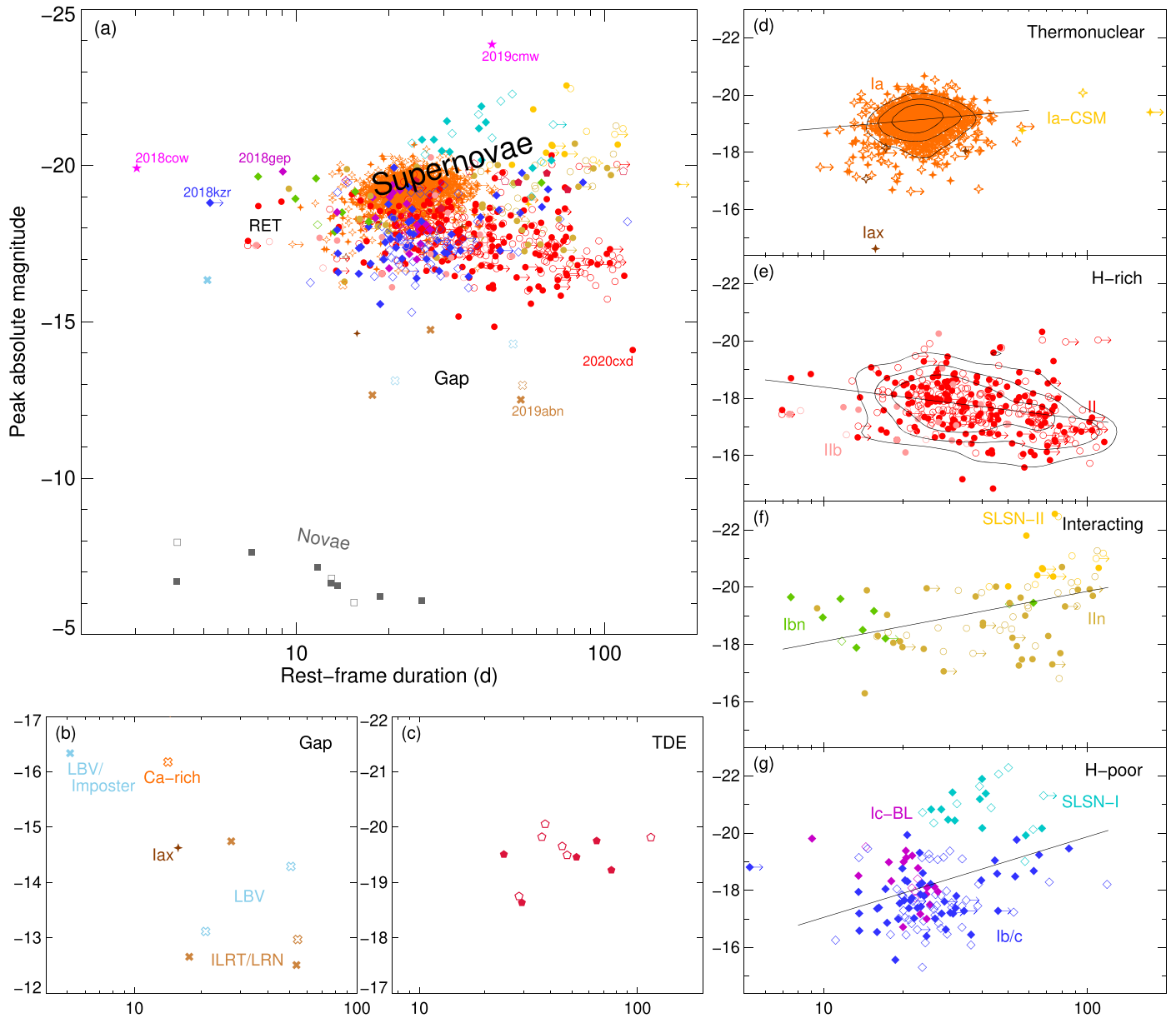
##### 4.1.1. Rapidly Evolving Transients

Transients in the leftmost part of the diagram ( $t < 10$  days) are expected to be somewhat undersampled owing to the limited cadence (Section 3). Even so, it is clear that very fast events (sometimes called rapidly evolving transients or RETs; Drout et al. 2014; Arcavi et al. 2016; Pursiainen et al. 2018; Wiseman et al. 2020<sup>28</sup>) are quite rare: there are only 14 with total durations of  $<10$  days and peak absolute magnitudes  $M < -16$  in the (statistical) sample. The most striking such object is AT2018cow (Prentice et al. 2018; Ho et al. 2019a; Margutti et al. 2019; Perley et al. 2019), which independently passed our sample cuts and appears as an outlier in parameter space, even in comparison to the other RETs. The physical nature of this event is still unknown. The next-fastest luminous event in the diagram, SN 2018kzr, faded too rapidly for its timescale to be measured precisely, but its rapid nature is confirmed by its fast rise. Nominally spectroscopically classified as an SN Ic, this event has been suggested to be a white dwarf accretion-induced collapse or a merger of a neutron star with a white dwarf (McBrien et al. 2019; Gillanders et al. 2020).

The remaining 12 events (Figure 8) are much less extreme (all have  $>5$  days duration) and show a variety of, generally ordinary, spectroscopic classifications: Type II and IIb SNe are most common (four and two examples, respectively), but there are also two Type Ibn, one fast Type IIc, one peculiar Type Ic-BL (SN 2018gep; Ho et al. 2019b), and one luminous blue variable (LBV) SN impostor. Most of these events are quite luminous ( $M < -18$  mag), making them incompatible with radioactive heating as the primary energy source (Arnett et al. 1989) for the main peak, although a few show subsequent second peaks or plateaus on a classical SN timescale. Shock-breakout into an extended envelope or circumstellar medium (CSM) has been suggested as the likely explanation for previous fast-peaking, luminous events (both for AT2018cow and for less-extreme RETs: Ofek et al. 2010; Drout et al. 2014; Pursiainen et al. 2018; Ho et al. 2019b; Margutti et al. 2019; Perley et al. 2019), and it is reasonable to hypothesize that this process is also responsible for most or all of the transients in this portion of the diagram in our sample.

As most of these evolved into relatively ordinary SN classes later in their evolution, it is not obvious that any additional special conditions are required beyond dense, extended CSM in most cases to produce a rapidly evolving transient, and we infer that this is also the case for most short-timescale events in the

<sup>28</sup> Other acronyms used include fast-evolving luminous transients (FELTs; Rest et al. 2018) or fast blue optical transients (FBOTs; Margutti et al. 2019).



**Figure 7.** (a) Duration–luminosity plot for 1197 classified BTS transients at  $m < 18.5$  mag satisfying our quality cut (filled points) and 966 additional transients with usable timescale measurements (unfilled points). Durations are time above half-peak and absolute magnitude is at the observed peak. The surrounding panels, (b)–(g), break the population into general spectral types, with non-SN populations on the bottom and SN populations at right. Panels (e)–(g) refer to CC SNe, with strongly CSM-interacting members of H-rich and H-poor populations shown in panel (f). Most SN types and subtypes occupy distinct (if overlapping) regions within duration–luminosity parameter space. Contours show 50%, 75%, and 90% containment of the kernel density estimate for SNe Ia and H-rich SNe. Correlations between duration and luminosity are observed for most SN populations.

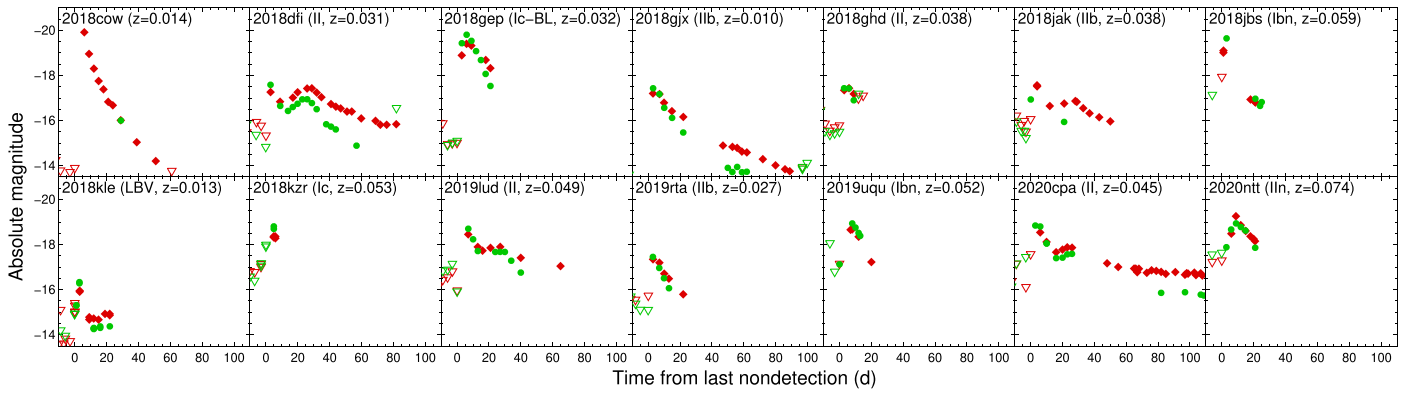
Dark Energy Survey, Pan-STARRS, and other earlier surveys. AT2018cow-like events, with their luminous multiwavelength emission (Ho et al. 2019a, 2020b; Margutti et al. 2019; Coppejans et al. 2020), represent a dramatic but rare exception, and certain SNe may also require an engine-driven jet to power the fast, energetic shock inferred from radio and X-ray observations (Ho et al. 2020c).

#### 4.1.2. Low-luminosity (“Gap”) Transients

The low-luminosity region of the diagram in the “gap” between novae and SNe ( $-9 > M > -15$  mag) is sparsely sampled—as expected for a magnitude-limited survey, since the volume to which events with these properties can be

detected is very limited. We do in fact detect similar numbers of low-luminosity transients and classical SNe at very low redshifts (two of each at  $z < 0.004$  or  $d < 17$  Mpc, where we are complete to  $M_{\text{lim}} \approx -12.7$  mag), suggesting that the volumetric rates of dim transients are at least comparable to those of classical SNe (see also Frohmaier et al. 2018).

Some low-luminosity events are clearly SNe themselves. SN 2020cxd at  $M_{\text{peak}} = -14.1$  mag is the most notable such example; with  $H\alpha$  velocity widths of  $5000 \text{ km s}^{-1}$  in its spectrum and an SN IIP-like light curve, it is almost certainly the explosion of a massive star, despite being an outlier relative to other SNe II in the sample. A handful of similar events are known in the literature (e.g., Pastorello et al. 2004, 2009; Gal-Yam et al. 2011).



**Figure 8.** ZTF light curves for rapidly evolving (time above half-peak  $t < 10$  days) transients in the sample. Symbol conventions are as in Figure 2. Most fast transients belong to known SN classes and many show clear second peaks or late-time phases of slow decline. These late-time features may not be obvious in surveys operating closer to the detection limit.

The remaining objects in the “gap” luminosity band do not match well-established SN templates. The classification system for these types of events is still evolving, and the progenitor interpretation of these classes remains an active area of research. Two are LBVs in a very luminous, high-activity state.<sup>29</sup> One is classified as an SN 2002cx-like SN Ia (“SN Iax”; Li et al. 2003; Foley et al. 2009, 2013), possibly an incomplete SN Ia that does not fully destroy the white dwarf. One is classified as a luminous red nova (LRN), a class of event generally interpreted as stellar mergers (Kulkarni et al. 2007; Pejcha et al. 2016; Pastorello et al. 2019; Blagorodnova et al. 2020). Three are classified as intermediate-luminosity red transients (ILRTs), a broadly defined observational class sometimes attributed to electron-capture SNe (Botticella et al. 2009; Thompson et al. 2009; Moriya et al. 2014), although this remains controversial; the distinction between these events and LRNs is not always obvious (Cai et al. 2019), and a variety of other models exist (Pastorello et al. 2007; Berger et al. 2009; Bond et al. 2009; Tsuna et al. 2020). The lowest-luminosity non-nova transient in the BTS sample, ILRT AT2019abn, is discussed in detail by Jencson et al. (2019).

A few additional events in our sample belong spectroscopically to low-luminosity classes that are outside the traditional SN scheme, but which individually have significantly higher luminosities and overlap with the SN distribution ( $M < -15$  mag). Among Ca-rich events (Filippenko et al. 2003; Perets et al. 2010; Kasliwal et al. 2012) in BTS, only SN 2019hty (De et al. 2020), with  $M = -16.1$  mag, passed our selection cuts. There is also one particularly luminous LBV eruption (SN impostor; Maund et al. 2006; Smith et al. 2011; Pastorello & Fraser 2019), and several luminous SNe Iax. A compilation of all events that are not standard SN types and have luminosities in the range  $-12 > M > -17$  mag is shown in Figure 7(b), with SNe Iax shown in comparison to the general SN Ia distribution in Figure 7(d).

At this stage the sample of low-luminosity events remains too small for a detailed examination of their population properties. The ZTF volume-limited survey (CLU) provides a much larger sample of transients in this regime, and a significantly expanded discussion of this population will be provided in forthcoming work in association with that effort.

<sup>29</sup> This category stretches our definition of “transient,” and our catalog is unlikely to be complete to such events.

An analysis of the hydrogen-poor subset of low-luminosity transients from the first two years of ZTF, with an emphasis on Ca-rich events, can be found in De et al. (2020). A discussion of low-luminosity, hydrogen-rich SNe will be provided by A. Tzanidakis et al. (2020, in preparation).

#### 4.1.3. Superluminous Transients

At the high-luminosity end, the superluminous supernova (SLSN; for reviews see Gal-Yam 2012, 2019) population is clearly visible as a group extending to the top and right of the general SN population, although with no indication of a gap between the SLSN and general SN populations (in agreement with De Cia et al. 2018). Hydrogen-poor and hydrogen-rich SLSNe form distinct regions in duration space: SLSNe-II are universally longer in duration. The SLSN population has been the focus of intensive efforts within ZTF to provide high completeness to significantly deeper limits than  $m > 18.5$  mag (e.g., Lunnan et al. 2020), and further discussion of this population will be reserved for a series of upcoming papers by Yan et al., Perley et al., and Chen et al.

No securely classified transient in the BTS sample is more luminous than  $M < -23$  mag, although we have found one featureless, slow transient with  $M = -23.6$  mag (AT2019cmw) inferred from a redshift measurement via intergalactic-medium absorption lines. It is not yet clear whether this represents an extreme SN, a TDE, or a particularly extreme AGN accretion phenomenon; it will be addressed by a subsequent study.

#### 4.1.4. Tidal Disruption Events

The number of TDEs within the BTS sample is relatively small (Figure 7(c)) and as yet insufficient for a detailed statistical investigation. While all are quite luminous ( $M_{\text{peak}} < -18$  mag, and all but two are at  $M_{\text{peak}} < -19$  mag), the absence of lower-luminosity examples does not yet rule out the possibility that fainter events (e.g., Blagorodnova et al. 2017) comprise the bulk of the population. Timescales range between approximately 20 and 120 days, although we are unlikely to be sensitive to any longer-duration events because they would be indistinguishable from AGNs to our filter (Section 2.1). A more complete overview of TDEs within ZTF can be found in the sample study of van Velzen et al. (2020).

#### 4.2. SNe and Luminosity–Duration Correlations

Events traditionally defined as SNe broadly occupy a common region of the diagram (of typically 10–100 days duration and absolute magnitudes between  $-16$  and  $-21$ ) and there is overlap between all SN types in this region. Even so, there are clear distinctions between different classes in this parameter space, with different trends emerging among different groups. In the right four panels ((d)–(g)) of Figure 7 we have separated SNe into four general categories: thermonuclear (Ia), H-rich (“ordinary” II and Ib), interacting (II, SLSN-II, and Ibn), and H-poor (Ib/c and SLSN-I); see, e.g., Filippenko (1997) for a review of SN spectral classification.

SNe Ia, as expected, have relatively standard properties and cluster in a small locus (although additional scatter is introduced due to host-galaxy extinction, distance uncertainties, sampling gaps, and the use of both  $g$  and  $r$  data to determine timescales). While our general duration parameterization differs from the  $\Delta m_{15}$  decline parameter often used in SN cosmology, the well-known correlation of Phillips (1993) between timescale and luminosity is nevertheless qualitatively replicated at high statistical significance: a simple linear regression to the SN Ia population shown in Figure 7 gives a slope ( $b \equiv \Delta m / \Delta \log t$ ) of  $-0.80 \pm 0.15 \text{ mag dex}^{-1}$  ( $1\sigma$  uncertainties).

Interestingly, other transient classes also display similar correlations. In particular, SNe II show a reasonably tight (Pearson coefficient  $r = 0.32$ ) correlation between magnitude and luminosity: longer events are dimmer ( $b = 1.14 \pm 0.21$ ). This is in agreement with other studies based on much smaller samples (Anderson et al. 2014; Sanders et al. 2015; Galbany et al. 2016a; Rubin & Gal-Yam 2016; Valenti et al. 2016; de Jaeger et al. 2019).

SNe II<sub>n</sub> (which for this purpose we take to include all events classified as SLSN-II on TNS, most but not all of which show SN II<sub>n</sub>-like narrow features) obey the opposite correlation: more slowly declining events are on average *more* luminous, in agreement with Ofek et al. (2014) and Nyholm et al. (2020). This relation may turn over toward the short-timescale end of the diagram: Type Ibn SNe form a small cluster of luminous CSM-interacting hydrogen-poor transients, of which the shortest events are generally *more* luminous than the longer ones, but the sample is small and the trend is not significant. A linear fit to the entire interacting SN population as shown in Figure 7 gives  $b = -1.75 \pm 0.42$  and  $r = -0.50$ .

Hydrogen-poor SNe, like interacting SNe, show a positive correlation between duration and luminosity ( $b = -2.82 \pm 0.64$ ,  $r = -0.34$ ). There is a hint that this population may cluster into separate subpopulations: a cluster of ordinary SNe Ib/c but also a population of much slower and brighter SNe Ic including SLSN-I. A  $k$ -means clustering analysis did not confirm that these clusters are statistically significant, so larger samples will be necessary to confirm this hypothesis.

These correlations may have useful cosmological applications, and the distinguishability of different subpopulations (even without color or shape information) may be encouraging for the use of photometric techniques to classify transients in future surveys. They may also provide further insight into the physics and progenitor populations of explosions of different types.

#### 4.3. Rate Measurements

Another key advantage of a magnitude-limited, untargeted, spectroscopically complete survey is that the measurement of volumetric rates is relatively straightforward.

For an ideal survey that is able to scan the entire sky to a given magnitude limit  $m_{\text{lim}}$  without interruption and is unaffected by Galactic dust, collecting a sample of  $N$  events over a total survey time  $T$  much longer than the duration of any individual transient, the volumetric rate can be estimated from

$$R = \frac{1}{T} \sum_{i=1}^N \frac{1}{\left(\frac{4\pi}{3} D_{\text{max},i}^3\right)}, \quad (1)$$

where  $D_{\text{max},i}$  is the distance out to which the  $i$ th transient can be detected above  $m_{\text{lim}}$  at peak light in the absence of extinction, given its peak absolute magnitude  $M_i$ .

Any real SN survey does not cover the whole sky, operates over a finite time window with a complex cadence structure, must contend with Galactic extinction, and does not recover all of the transients it “detects.” This requires additional correction terms to compensate for the effective loss of survey volume, and for the gain of additional transients that “occurred” (peaked) outside the survey time window but were detected on the rise or the decline. These corrections can potentially be difficult to apply in practice since the loss/gain factors may vary by transient type, sky location, and other factors. In the case of BTS, we have strictly chosen a sample such that the peak is well determined using unbiased sample cuts and guaranteed to occur within the survey window, making this task much simpler. A revised equation is

$$R = \frac{1}{T} \sum_{i=1}^N \frac{1}{\left(\frac{4\pi}{3} D_{\text{max},i}^3\right) f_{\text{sky}} f_{\text{ext}} f_{\text{rec}} f_{\text{cl},i}}. \quad (2)$$

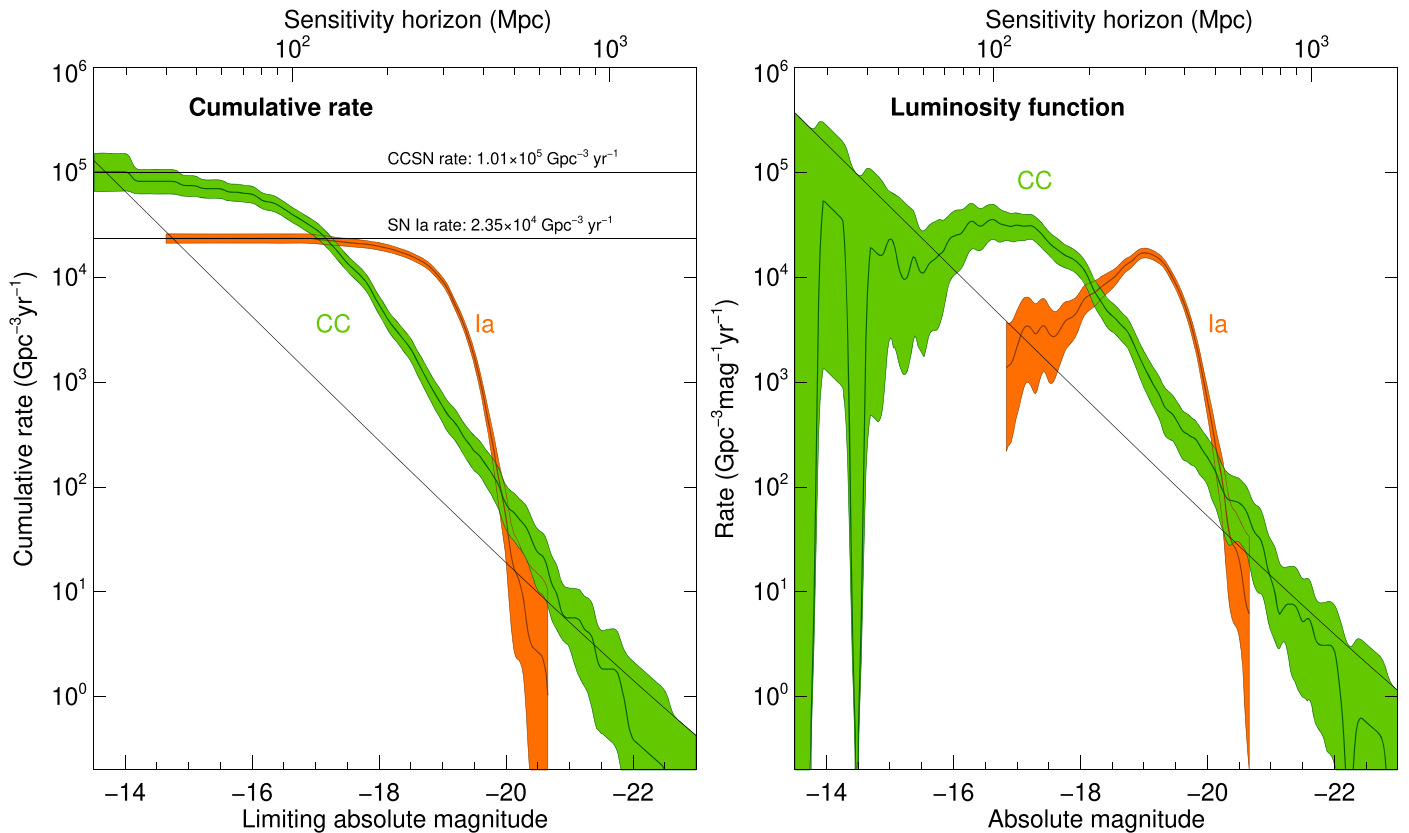
The loss factors are as follows.

1.  $f_{\text{sky}}$  is the average active survey footprint expressed as a fraction of the full sky.
2.  $f_{\text{ext}}$  is the average reduction in effective survey volume owing to Galactic extinction.
3.  $f_{\text{rec}}$  is the average recovery efficiency for a detectable transient within the survey footprint: the probability that it is found and included in the sample.
4.  $f_{\text{cl},i}$  is the classification efficiency. (This may depend on apparent magnitude, so the subscript is retained.)

Using the exposure history from the public survey, we estimate the average active area across the three-night cadence cycle over the period considered here of  $14,400 \text{ deg}^2$  ( $f_{\text{sky}} = 0.35$ ).

The Galactic extinction correction  $f_{\text{ext}}$  can be calculated by averaging the reduction in volume associated with the extinction toward each separate ZTF field (excluding fields with  $A_V < 1.0 \text{ mag}$  that are omitted from the sample). We infer  $f_{\text{ext}} = 0.82$ .

The recovery fraction is the most uncertain parameter. We previously estimated (Section 2.3) that 52% of our candidate transients passed our quality cuts, but this is not an ideal estimate because some candidates were not transients, or may have occurred (peaked) outside the active sky region and been classified much later. To provide a better estimate of this parameter, we took all events classified as SNe Ia with peak



**Figure 9.** Rate measurements for Type Ia and CC SNe. The plot on the left shows the volumetric rate of SNe brighter than a particular absolute magnitude (in either the  $g$  or  $r$  band, without correcting for host-galaxy extinction); the plot on the right shows the luminosity function calculated from SNe peaking within  $\pm 0.25$  mag of a particular magnitude. Colored bands correspond to 95% statistical confidence intervals. Horizontal lines show the total SN Ia rate (to  $-16.5$  mag) and CC SN rate (to  $-14$  mag) estimated from this work. The black diagonal line indicates statistical upper limits (95% confidence) for the case of zero detected events at a given magnitude.

absolute magnitudes of  $< -18.5$  and peak apparent magnitudes of  $< 18$ , a set of conditions that effectively ensures that the transient peaked within the active area and that it would have been very easy to classify even in suboptimal observing circumstances. Of these, 412/690 pass the quality cut, so we estimate  $f_{\text{rec}} = 0.60$ .

Our classification completeness was addressed in Section 3. It is close to 100% for bright transients but declines to about 90% at  $m = 18.5$  mag and drops quickly afterward. We assume  $f_{\text{cl}} = 1.0$  if  $m < 17.2$  mag and  $f_{\text{cl}} = 0.9$  at  $m = 18.5$  mag, with a linear decline in between.

The sample as presented in this paper spans  $t = 2.12$  yr of ZTF. We assume a uniform  $K$ -correction ( $K = 2.5 \times \log_{10}(1+z)$ ) and ignore cosmological effects.<sup>30</sup>

Based on these assumptions (and  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ), we infer an SN Ia rate of  $(2.35 \pm 0.24) \times 10^4 \text{ Gpc}^{-3} \text{ yr}^{-1}$  (95% confidence interval, statistical errors only). Caution should be taken in interpreting this as a truly independent measurement of the SN Ia rate: the survey parameters above were not chosen entirely blindly of the result and the true uncertainty will be dominated by systematics, which are not easy to quantify. Even so, it is encouraging that this value is very close to the value from several large-scale studies over the

past ten years (Dilday et al. 2010; Graur et al. 2011; Li et al. 2011; Frohmaier et al. 2019). We are not able to confirm the claim by Smith et al. (2019) of a much higher rate.

This method can be generalized to any population and any limiting magnitude or magnitude range. In Figure 9 we show both cumulative rates and luminosity functions for the SN Ia and CC SN populations. We replace the lower limit with a limit calculated from the  $m < 19.0$  mag sample (a limiting value of  $f_{\text{cl}} = 0.9$  is assumed beyond  $m > 18.5$  mag) if this is more constraining than that from the  $m < 18.5$  mag sample. Note that luminosities are as observed: host-galaxy extinction is *not* removed, although Galactic extinction is corrected for.

Calculating the total CC SN rate is more challenging than calculating the SN Ia rate, both because the number counts are less and because the luminosity function is broader, with a significant fraction of the population coming from very dim events that are not detectable except in small volumes within the nearby universe (see also Taylor et al. 2014 and forthcoming work by Tzanidakis et al.). Assuming a minimum luminosity of  $M < -14$  mag we infer a rate of  $(10.1^{+5.0}_{-3.5}) \times 10^4 \text{ Gpc}^{-3} \text{ yr}^{-1}$ .

This value is fully consistent with predictions based on the low-redshift star formation rate density (Madau & Dickinson 2014). This is in agreement with other works arguing that the “SN rate problem” (Horiuchi et al. 2011, originally motivated by the lower CC SN rate of Li et al. 2011) is resolved using galaxy-untargeted surveys and including the faint end of the luminosity function, without requiring a large population of completely optically obscured SNe (although

<sup>30</sup> The contraction of the control time window ( $\Delta t \propto (1+z)^{-1}$ ) is approximately compensated for by the increase in the star formation rate density ( $\text{SFRD} \propto (1+z)^{+1.2}$  in the low-redshift limit of Equation (9) of Madau & Dickinson 2014) and redshift-dependent SN rates (e.g., Dahlen et al. 2004, 2008; Barris & Tonry 2006; Dilday et al. 2010; Graur et al. 2011; Melinder et al. 2012; Strolger et al. 2015).

such events may exist; Mattila et al. 2012; Jencson et al. 2019) or direct collapses.

This same methodology could be applied to other, rarer transients in the sample. For example, inspection of the high-luminosity end of the CC SN rate curve implies an SLSN rate (above  $M < -21$  mag) of  $5.6^{+5.4}_{-2.8}$  Gpc $^{-3}$  yr $^{-1}$ , which is much lower than the commonly cited estimate by Quimby et al. (2013) of  $199^{+137}_{-86}$  Gpc $^{-3}$  yr $^{-1}$ .<sup>31</sup> We will present further calculations of the SLSN rate and luminosity function (using the full ZTF SLSN sample of  $>150$  events, extending to much fainter limiting magnitudes and with appropriate cosmological and  $K$ -corrections) in forthcoming work by Yan et al.

Rate calculations can, in principle, be extended even to classes of transients we do not detect at all. The diagonal lines in Figure 9 show 95% confidence upper limits on the intrinsic rate of any transient for which we have found no examples in BTS so far, assuming that the transient does not have properties that make it systematically selected against by the survey cadence or our selection cuts (e.g., very short duration, occurs near variable AGNs, etc.) and that it is not mistaken for another class of object. For example, we have not detected any event with properties consistent with an off-axis GRB afterglow or a kilonova in BTS to date. Assuming that these events are not selectively missed, this would imply that the rate and luminosity functions lie below these limits; e.g.,  $<5 \times 10^3$  Gpc $^{-3}$  yr $^{-1}$  for kilonovae more luminous than  $M < -16$  mag, or  $<7$  Gpc $^{-3}$  yr $^{-1}$  for off-axis afterglows peaking above  $M < -20.5$  mag. The kilonova estimate is consistent with the rate estimated by Andreoni et al. (2020) using all ZTF data although not as constraining, since Andreoni et al. (2020) do not require spectroscopic classifications and probe much deeper than  $m > 18.5$  mag. The GRB rate is consistent with expectations given the on-axis GRB rate of  $\sim 1$  Gpc $^{-3}$  yr $^{-1}$  (Wanderman & Piran 2010), assuming we would only be sensitive to events seen within a viewing angle of  $\theta < 3\theta_{\text{jet}}$  at this luminosity threshold.

#### 4.4. Host-galaxy Properties

The host-galaxy population of a specific transient subtype offers a valuable clue to the nature of its progenitor. The hosts of SNe Ia provide a means to estimate the distribution of the delay time between formation and explosion, and thereby the nature of the progenitor binary system. The association of CC SNe with generally young stellar populations is well established, as expected for massive stars, but measuring the (much shorter) delay time sufficiently precisely to constrain theory is much more challenging when only galaxy-integrated measurements are available. However, comparisons between different CC SN subtypes can still provide a powerful constraint on their respective origins. In particular, evolutionary models of single stars generally predict that mass loss will be more effective at high metallicity than at low metallicity (Maeder & Meynet 2000), increasing the fraction of stripped-envelope SNe in metal-rich galaxies. Very low metallicities have been suggested to be conducive to powering engine-powered and other rare transients that require a hydrogen-poor progenitor *without* strong winds (e.g., Yoon & Langer 2005).

<sup>31</sup> It is also lower than the estimate at  $z = 1.1$  by Prajs et al. (2017) of  $91^{+76}_{-36}$  Gpc $^{-3}$  yr $^{-1}$ , although if the factor of  $\sim 6$  increase in the star formation rate density with cosmic time is taken into account, our estimates are marginally consistent within the uncertainties.

The BTS volume extends well beyond the redshifts at which spectroscopic galaxy catalogs are complete (Fremming et al. 2020), and a thorough statistical investigation of the transient–host-galaxy connection will require significant spectroscopic follow-up observations and spectral energy distribution (SED) fitting to aperture-matched multiwavelength data. However, BTS redshifts are well within the range at which all-sky *photometric* galaxy catalogs are largely complete for all but the lowest-luminosity galaxies, making it possible to study the basic characteristics of the sample with these data alone.

In Figure 10 we plot the luminosities ( $M_i$ ) and colors ( $g - i$ ) of the host galaxies for 593 SNe Ia and 321 CC SNe at  $0.015 < z < 0.1$  within the SDSS footprint,<sup>32</sup> calculated using a probabilistic cross-match (and subsequent manual vetting). Both values are corrected for Galactic extinction. For comparison, we also plot the full set of SDSS (spectroscopic) galaxies at  $z < 0.03$  from the NASA-SDSS Atlas (Blanton et al. 2011). The completeness limit for the SDSS spectroscopic sample is approximately  $M_i < -17.9$  within this volume.

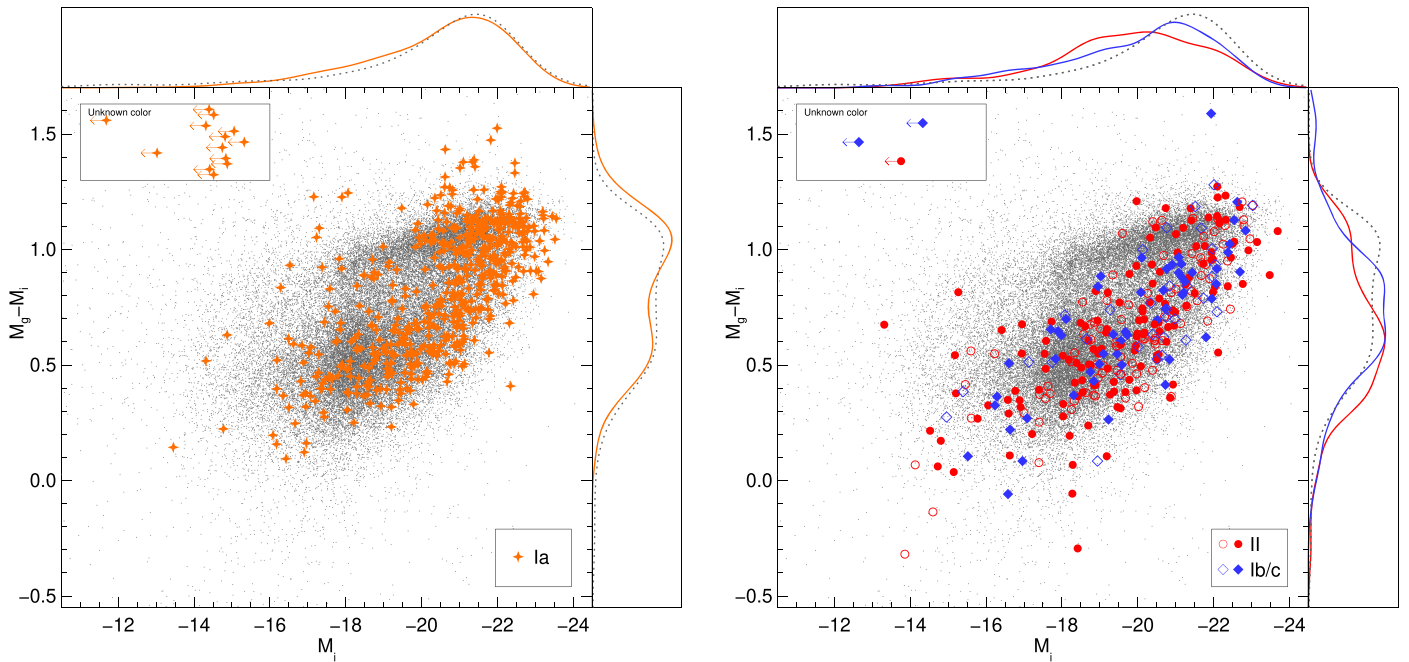
Low-redshift field-selected galaxies show a bimodal color–magnitude distribution: a “red sequence” dominated by passive, early-type galaxies and a “blue cloud” dominated by spirals and irregulars. (The region between these populations is sometimes termed the “green valley.”) This is evident in the SDSS field population in Figure 10. It is less obvious in the SN hosts, although the SN Ia color distribution (right subpanel of left panel) shows the associated bimodality clearly.

We subdivided the SN host population in color–magnitude space between red-sequence, green-valley, blue-cloud, and subdwarf galaxies as defined in Table 2. “Hostless” events are designated separately. For CC SNe these “hostless” events probably do have low-luminosity, coincident hosts fainter than the SDSS detection threshold (e.g., Zinn et al. 2012). For most “hostless” SNe Ia we anticipate that there is no coincident host and the progenitor has traveled a sufficient distance from the galaxy in which the system formed such that there is no probabilistically secure association, although it is possible that a few may have undetected low-luminosity, coincident hosts (Strolger et al. 2002).

Red-sequence galaxies contribute about a third (31%<sup>33</sup>) of the SN Ia population but also significantly contribute to the CC SN population (11%). We inspected the imaging of all SDSS-matched CC SN host galaxies with  $g - i > 0.9$  mag; most have an early-type morphology or contain H II regions or other signatures of ongoing star formation at or near the SN site, as would be expected given the short lifetimes of their progenitors. However, two CC SN hosts (SN 2019ape and SN 2020oce, representing 1% of the sample) are featureless ellipticals with no visible signs of star formation. This could indicate that these events are not actually CC SNe, that some CC SNe have long progenitor lifetimes, or that even classical ellipticals contain some residual star formation (Hakobyan et al. 2008, 2012; Graham et al. 2012; Irani et al. 2019). These

<sup>32</sup> PS1 catalog photometry is affected by aperture differences and galaxy shredding to a much greater extent than for SDSS and was found to be unreliable for this purpose, so we restrict our analysis to SDSS fields for now. We introduce a lower limit on the redshift to avoid background over-subtraction and host-galaxy mismatches, and an upper limit to ensure that nondetections are always constraining.

<sup>33</sup> Numbers presented in this section are based on the  $m < 18.5$  mag, quality-cut sample, to avoid the possibility of host-dependent spectroscopic confirmation bias. However, we obtain consistent results for the full classified sample.



**Figure 10.** Color–magnitude diagrams (rest-frame  $g - i$  color vs.  $r$ -band absolute magnitude) for the host galaxies of SNe Ia (left) and of CC SNe (right). We restrict the SN host population to a redshift range of  $0.015 < z < 0.1$  and show only events in SDSS fields passing the quality cuts. For SNe Ia we show only  $m < 18.5$  mag events but for CC SNe we also show host galaxies of fainter events (unfilled symbols). SDSS galaxies from the NASA-SDSS Atlas at  $z < 0.03$  are shown as gray points. The inset box at upper left shows SNe with no host association, which have a magnitude upper limit but no color constraint (their  $y$ -axis positions within the inset are arbitrary). The side panels show kernel-density curves for each SN host population and for the SDSS galaxies (weighted by  $u$ -band luminosity divided by  $V_{\max}$ ). CC SNe are grouped into only two general classes: SNe II (including all subtypes) in red and SNe Ib/c (including all subtypes) in blue. The luminosity and color differences between the Type II and Type Ib/c host-galaxy populations are not statistically significant.

**Table 2**

Contributions of Different Galaxy Populations to SNe Ia and CC SNe for SDSS Fields

Population	$N_{\text{Ia}}$	$f_{\text{Ia}}$	$N_{\text{CC}}$	$f_{\text{CC}}$
Red sequence <sup>a</sup>	181	$31\% \pm 2\%$	21	$11\% \pm 2\%$
Green valley <sup>b</sup>	112	$19\% \pm 2\%$	37	$19\% \pm 3\%$
Blue cloud <sup>c</sup>	281	$47\% \pm 2\%$	128	$64\% \pm 3\%$
Subdwarf <sup>d</sup>	6	$1.0\% \pm 0.4\%$	10	$5.0\% \pm 1.6\%$
Hostless <sup>e</sup>	13	$2.2\% \pm 0.6\%$	3	$1.5^{+1.1}_{-0.7}\%$

**Notes.** Contributions are given both as total counts and as a fraction. Uncertainties are approximate  $1\sigma$  binomial confidence intervals.

<sup>a</sup>  $M_i < -16$  and  $g - i > \max\{0.85 - 0.05(M_i + 18), 0.85\}$ .

<sup>b</sup>  $M_i < -16$  and  $0.85 < g - i < 0.85 - 0.05(M_i + 18)$ .

<sup>c</sup>  $M_i < -16$  and  $g - i < 0.85$ .

<sup>d</sup>  $M_i > -16$ .

<sup>e</sup> No host association found. May include SNe at large offset and intracluster SNe.

(and other candidate non-star-forming SN hosts found by ZTF) will be discussed in further detail by I. Irani et al. (2020 in preparation).

Very low-luminosity galaxies contribute negligibly to the SN Ia population (1.0%, not including “hostless” events) but represent a more significant fraction of the CC SN population (5.1%, plus 1.5% “hostless” events that are probably undetected dwarfs). While there are relatively few low-luminosity galaxies in the SDSS spectroscopic sample by raw numbers, this is largely a result of Malmquist biases inherent in redshift measurement. These galaxies are quite common by volume and are responsible for a significant fraction of cosmic star formation (although far from a majority; Brinchmann et al. 2004).

The SN Ia population is noticeably skewed toward being hosted within redder and more luminous galaxies than CC SNe, as expected for a population that largely traces stellar mass (although a component also traces young stars; e.g., Mannucci et al. 2006; Sullivan et al. 2006). Notably, however, there is no significant difference in the luminosities or colors of SN II and SN Ib/c hosts (a Kolmogorov–Smirnov two-sample test gives  $p_{\text{KS}} = 0.55$  for  $m < 18.5$  mag and  $p_{\text{KS}} = 0.08$  at  $m < 19$  mag), which may suggest that the importance of metallicity-sensitive channels (such as single-star wind stripping) is minor. While surprising given earlier work on this topic (Boissier & Prantzos 2009; Arcavi et al. 2010; Graur et al. 2017a, 2017b), this is in agreement with some other recent observational studies (Anderson et al. 2015; Galbany et al. 2016b; Kuncarayakti et al. 2018; Taggart & Perley 2019; Schulze et al. 2020).

As the SDSS-matched sample is relatively small, we cannot rule out smaller differences in luminosity (the difference in mean absolute magnitude is  $0.06 \pm 0.35$ ) or color. We have also not yet investigated explosion-site properties of the sample, which are likely to better reflect differences in delay times among young transients (Anderson & James 2008; Kelly & Kirshner 2012; Galbany et al. 2014; Maund 2018; Xiao et al. 2019).

## 5. Summary and Online Catalog

In this paper we have summarized the status of the ZTF BTS after two years of operation, and illustrated several basic cuts with which to establish a large, high-quality, unbiased, and nearly spectroscopically complete sample of extragalactic transients. Using this sample we have provided a preliminary exploration of transient parameter space on timescales of 6–200

days, a new estimate of the CC SN rate and luminosity function, and constraints on the fraction of star formation in rare environments such as low-mass and red-sequence galaxies.

Our results are based on ZTF alert-packet data and TNS reports, both of which are susceptible to occasional errors. Also, the analyses we have employed are based on simplified general techniques, chosen based on applicability to a wide range of SN properties and the need to avoid human intervention. These results will eventually be superseded by focused papers on all of these topics using additional classifications, improved redshift measurements, analysis of SN subtypes, superior reference template images and forced photometry, light-curve modeling, host SED fitting, and many other enhancements. However, we emphasize that the analysis presented here can be updated continuously in real time using public data as the sample continues to expand, and we invite the community to explore the properties of SN parameter space using our public data releases.

To this end, we have created a new web resource summarizing the key properties of our sample and provided an interactive interface to explore it in detail. Titled the BTS Sample Explorer, it provides a sortable web table containing the time and magnitude of peak light for each transient, classifications and redshifts, timescales measured from the light curve, extinction and luminosity measurements, and host-galaxy photometric properties. P48 stamp images, light-curve plots, and colorized Pan-STARRS images of the field and host galaxy are also provided. An alternative viewing mode allows instant collages of light curves, Pan-STARRS cutouts, or both, for public presentation or visual data exploration. The data table can be downloaded as a .csv file for offline exploration. All of these resources are updated nightly to add new transient discoveries, classifications, and measurements. This resource can be accessed at <https://sites.astro.caltech.edu/ztf/bts/explorer.php>.

Looking further to the future, while the SN catalog we present here is by far the largest of its type to date, it is clear that even larger-scale efforts will be required to fully address many key science areas: for example, unbiased population studies of uncommon transients (of which we have only a few examples so far) and at the faint end of the luminosity function for common transients. Several additional years of ZTF operations (with improved cadence and scheduling software, as planned for the second phase of ZTF, which began in October 2020) will be key to this: BTS has been operational for only two years so far, with much of the first year devoted to reference-building. Combining ZTF with other surveys will also be of significant benefit: close to half of our sample could not be fully utilized due (in part) to light-curve gaps, but data from telescopes at other sites may be able to fill these gaps.

Key to the success of BTS so far has been the availability of dedicated robotic spectrographs, especially the SED Machine at Palomar. The commissioning of more such facilities, alongside new high-throughput multichannel spectrographs on existing telescopes, would enable highly complete spectroscopic coverage of the transient sky to be extended to greater depths and larger sky areas. Several projects of this type are in development, including a proposed SED Machine clone at Kitt Peak, the New Robotic Telescope on La Palma (Gutiérrez et al. 2019), the Next Generation Palomar Spectrograph (Jiang et al. 2018), and the Son of X-shooter at La Silla (Schipani et al. 2018). A concerted, organized effort by these facilities

could easily increase the size of highly complete samples similar to BTS by an order of magnitude or more by the middle of the decade. This would allow for unambiguous rate and luminosity measurements and strong progenitor constraints for virtually all currently known classes of transients, discoveries of (or constraints on) even extremely rare and/or fast-evolving events, and would greatly advance our understanding of the explosive universe.

Even with these new generations of spectrographs, the larger but fainter populations of transients found by the Vera Rubin Observatory will pose particular challenges for obtaining similarly complete samples from the transients found by that facility, and heavy reliance on photometric classification will be unavoidable. Against this backdrop, bright-end surveys supported by large-scale spectroscopic classification efforts (such as ZTF/BTS) remain critical: the wealth of photometric training data and improved rate measurements will aid in developing reliable photometric classification tools and motivating follow-up strategies for the LSST era.

We would like to acknowledge the contributions of the worldwide community of observational astronomers in obtaining and reporting classification spectra of transients within this sample. We specifically acknowledge observing and classification contributions from Yuhan Yao, Viraj Karambelkhar, Igor Andreoni, and Alison Dugas (Palomar/Keck), WeiKang Zheng, Kishore Patra, and Andrew Hoffman (Lick), Brigitta Spiőcz, Zach Golkhou, Keaton Bell, and James Davenport (APO). We also acknowledge the contributions of BTS and other ZTF alert-stream scanners, including Raphael Baer-Way, Teagan Chapman, Matt Chu, Asia deGraw, Suhail Dhawan, Alison Dugas, Nachiket Girish, Samantha Goldwasser, Andrew Hoffman, Connor Jennings, Evelyn Liu, Emily Ma, Emma McGinness, Yukei Murakami, Derek Perera, Druv Punjabi, James Sunseri, Abel Yagubyan, and Erez Zimmerman. We thank Christopher Cannella and Ashot Bagdasaryan for developing and managing the GROWTH Marshal, and Peter Nugent and Eran Ofek for useful comments on this manuscript. We also thank Suvi Gezari and Sjoert van Velzen for discussions on the nature of AT2019cmw and AT2019fdr. We thank the anonymous referee for useful comments.

Based on observations obtained with the Samuel Oschin 48-inch Telescope and the 60-inch Telescope at the Palomar Observatory as part of the Zwicky Transient Facility project. ZTF is supported by the National Science Foundation (NSF) under grant AST-1440341 and a collaboration including Caltech, IPAC, the Weizmann Institute for Science, the Oskar Klein Center at Stockholm University, the University of Maryland, the University of Washington, Deutsches Elektronen-Synchrotron and Humboldt University, Los Alamos National Laboratories, the TANGO Consortium of Taiwan, the University of Wisconsin at Milwaukee, and the Lawrence Berkeley National Laboratory. Operations are conducted by COO, IPAC, and UW. The SED Machine is based upon work supported by the NSF under grant 1106171. The Liverpool Telescope is operated on the island of La Palma by Liverpool John Moores University in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias with financial support from the UK Science and Technology Facilities Council. Research at Lick Observatory is partially supported by a generous gift from Google.



Some of the work associated with this paper was carried out at the Aspen Center for Physics (ACP). The ACP is supported by NSF grant PHY-1607611. This work was partially supported by a grant from the Simons Foundation.

This work was supported by the GROWTH (Global Relay of Observatories Watching Transients Happen) project funded by the NSF under PIRE grant 1545949. GROWTH is a collaborative project among California Institute of Technology (USA), University of Maryland College Park (USA), University of Wisconsin Milwaukee (USA), Texas Tech University (USA), San Diego State University (USA), University of Washington (USA), Los Alamos National Laboratory (USA), Tokyo Institute of Technology (Japan), National Central University (Taiwan), Indian Institute of Astrophysics (India), Indian Institute of Technology Bombay (India), Weizmann Institute of Science (Israel), The Oskar Klein Centre at Stockholm University (Sweden), Humboldt University (Germany), Liverpool John Moores University (UK), and University of Sydney (Australia).

A.A.M. is funded by the LSST Corporation, the Brinson Foundation, and the Moore Foundation in support of the LSSTC Data Science Fellowship Program; he also receives support as a CIERA Fellow by the CIERA Postdoctoral Fellowship Program (Center for Interdisciplinary Exploration and Research in Astrophysics, Northwestern University). A.G.-Y.'s research is supported by the EU via ERC grant No. 725161, the ISF GW excellence center, an IMOS space infrastructure grant and BSF/Transformative and GIF grants, as well as The Benozio Endowment Fund for the Advancement of Science, the Deloro Institute for Advanced Research in Space and Optics, The Veronika A. Rabl Physics Discretionary Fund, Paul and Tina Gardner, Yeda-Sela and the WIS-CIT joint research grant; A.G.-Y. is the recipient of the Helen and Martin Kimmel Award for Innovative Investigation. Y.-L.K. has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 759194—USNAC). A.V.F.'s group is grateful for funding from the TABASGO Foundation, the Christopher J. Redlich Fund, and the Miller Institute for Basic Research in Science (U.C. Berkeley). M.L.G. acknowledges support from the DiRAC Institute in the Department of Astronomy at the University of Washington. The DiRAC Institute is supported through generous gifts from the Charles and Lisa Simonyi Fund for Arts and Sciences, and the Washington Research Foundation.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the NSF, the U.S. Department of Energy, the National Aeronautics and Space Administration (NASA), the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton

University, the United States Naval Observatory, and the University of Washington.

The Pan-STARRS1 Surveys (PS1) have been made possible through contributions of the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, NASA under grant NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the NSF under grant AST-1238877, the University of Maryland, and Eotvos Lorand University (ELTE).

This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA.

Some of the data that contributed to this paper were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and NASA. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community.

*Facilities:* PO:1.2 m, PO:1.5 m, Liverpool:2 m, ARC, Shane, Hale, Keck:I (LRIS).

## Appendix A Filter Details

The 2019 version of the BTS filter is designed to pass all genuine transients that would also be passed by the original filter while reducing the false-positive rate by an order of magnitude. In detail, requirements of the filter are as follows.

1. The alert must have  $\text{magpsf} < 19$ , or a detection in the history with  $\text{magpsf} < 19$  in the last 18 hr.
2. The subtraction must be positive ( $\text{isdifffpos} = t$  or  $\text{isdifffpos} = 1$ ).
3. The location must be outside the Galactic plane:  $|b| > 7^\circ$ .
4. The alert must have  $\text{rbpscore} > 0.2$ . If close to a bright catalog object this is increased to  $\text{rbpscore} > 0.3$  if an  $m < 17$  mag source is within  $1''$  and to  $\text{rbpscore} > 0.45$  if an  $m < 15.5$  mag source is within  $1''.5$ . (The source magnitude  $m$  can be in any filter and can be from PS1 or Gaia.)
5. The alert must have  $\text{drb} > 0.1$ .
6. The alert must not be within  $\text{distpsnr} < 2''$  of a high-probability PS1 star ( $\text{sgscore} > 0.76$ ). It must also not be within  $0''.5$  of a bright PS1 object with uncertain stellarity ( $\text{sgscore} = 0.5$  and  $m < 17$  mag in any PS1 filter), or within  $1''.0$  of a very red PS1 source ( $r - i > 3$  or  $r - z > 3$  mag, with  $\text{sgscore} > 0.2$ ).
7. The alert must not be close to a bright potential star among any of the three PS1 sources in the packet. The exclusion radius depends on the star's magnitude and  $\text{sgscore}$ . It is  $\text{distpsnr} < 20''$  for stars with  $r < 15$

or  $i < 14.5$  mag and  $sgscore > 0.8$ , and for stars with  $r < 12$  or  $i < 11.5$  mag and  $sgscore > 0.49$ . It is  $distpsnr < 10''$  for stars with  $z < 14.0$  mag and  $sgscore > 0.8$ , or  $z < 11.5$  mag and  $sgscore > 0.49$ . It is  $distpsnr < 5''$  for stars with  $r < 15$  or  $i < 14.5$  mag and  $sgscore > 0.49$ . It is  $distpsnr < 2.5''$  for stars with  $z < 14$  mag and  $sgscore > 0.49$ . It is  $distpsnr < 1.1''$  for stars with  $r < 16.5$  or  $i < 16.0$  mag and  $sgscore > 0.49$ . It is  $distpsnr < 0.9''$  for stars with  $z < 15.5$  mag. The most restrictive (largest exclusion radius) is always used. While an improvement over our 2018 filter, this criterion was found to occasionally reject real SNe and has been further revised in mid-2020.

8. To remove moving objects, there must be another alert at the same location in the history more than 0.02 days prior.
9. The alert must also not have a cross-match in the minor planet catalog within  $ssdistnr < 15''$ .
10. The object must not be a long-term variable, as determined by the presence of a coincident counterpart and a first detection well before the alert (generally  $>90$  days, based on  $dt = jd - jdstarthist$ ). The counterpart matching radius is  $distnr < 0.4''$  for  $magnr < 19.5$  mag,  $distnr < 0.8''$  for  $magnr < 17.5$  mag,  $distnr < 1.2''$  for  $magnr < 15.5$  mag, and  $distnr < 9.5''$  for  $magnr < 13$  mag; it is  $neargaia < 0.35''$  for  $maggiaia < 17$  mag, and  $neargaia < 0.20''$  for  $maggiaia < 18$  mag. Sources with  $neargaia < 0.35''$  and  $maggiaia < 19$  mag are also excluded but only if the alert is  $m > 18.5$  mag and  $dt > 300$  days. A historical alert can be generated for spurious regions (especially near galaxy centers where bad subtractions are common) so caution is necessary in applying this filter: the criteria based on the reference catalog are only applied if the light curve is not at a local maximum and there are already several  $m < 19$  mag detections in the history (an indicator that the source has passed the filter before and not been saved).

The above summary is slightly simplified and the associated changes were not all made simultaneously. Prior to 2019 June, selection was performed using the basic filter described by Fremling et al. (2020); after 2020 June, the filter above was updated to decrease the exclusion radius around bright stars but cut more strictly on  $drb$ . This further-improved filter passed all of the TNS-cataloged transients that we missed on account of star-galaxy confusion (Section 3) with the exception of SN 2019gcc, a nuclear SN Ia that had a few detections several months prior to explosion that could be due to activity from a coincident weak AGN. Python code for all three versions (2018, 2019, and 2020) is available online.<sup>34</sup>

Note that  $drb$  became available in alert packets only in summer 2019, and asteroid and Gaia matching were also not available for packets early in the survey. When running the filter retroactively (for the purposes of our quality cut and verification checks) any criteria associated with missing fields are not applied.

## Appendix B Cross-Match Associations for Purity Filter

Any transient candidate saved to the program that passes the quality cuts and that has an ‘‘SN-like’’ timescale will be

included in our sample. To avoid excluding any potentially very fast or very slow transients, we also pass all transients with credible host-galaxy associations even if their timescale is not SN-like. We perform two host-galaxy association checks using cross-matches of different catalogs.

### B.1. Pan-STARRS Cross-match

The first check involves the Pan-STARRS1 catalog. The nearest three Pan-STARRS matches, with  $sgscore$  values, are located in the Avro packet data; we generally take the nearest cross-match of the three, although if the nearest source has  $sgscore > 0.75$  we will use one of the other two sources if it has  $sgscore < 0.75$  and is within  $5''$ . A plot of the magnitude versus offset of PS1 matches is shown in Figure 11, for both classified and unclassified events including all non-transient false positives saved to the program. We only show cross-matches with  $sgscore < 0.8$ . For moderate offsets the plot is overwhelmingly dominated by genuine transients. For very small offsets the transient population is contaminated by two other populations: AGNs at bright magnitudes and CVs at faint magnitudes; these are generally cases where  $sgscore$  has miscategorized the source. It is also contaminated by CVs at very large offsets and faint magnitudes (in this case due to chance occurrence near a galaxy).

The diagram is subdivided by lines into three regions: coincident cross-matches (at left), offset but highly probable cross-matches (center), and likely spurious cross-matches (lower right). The equation defining a coincident match is

$$\theta < 0.1 + 50 \times 10^{-0.2(m-15.8)}(0.3 + 0.7(1 + e^{15.8-m})^{-1}). \quad (B1)$$

The equation defining a spurious match is

$$\theta > 0.1 + 0.4(1 + 0.5 \times 10^{-0.8(m-21.5)})^{-1} + (1 + 10^{+0.2(m-15.5)})^{-1}. \quad (B2)$$

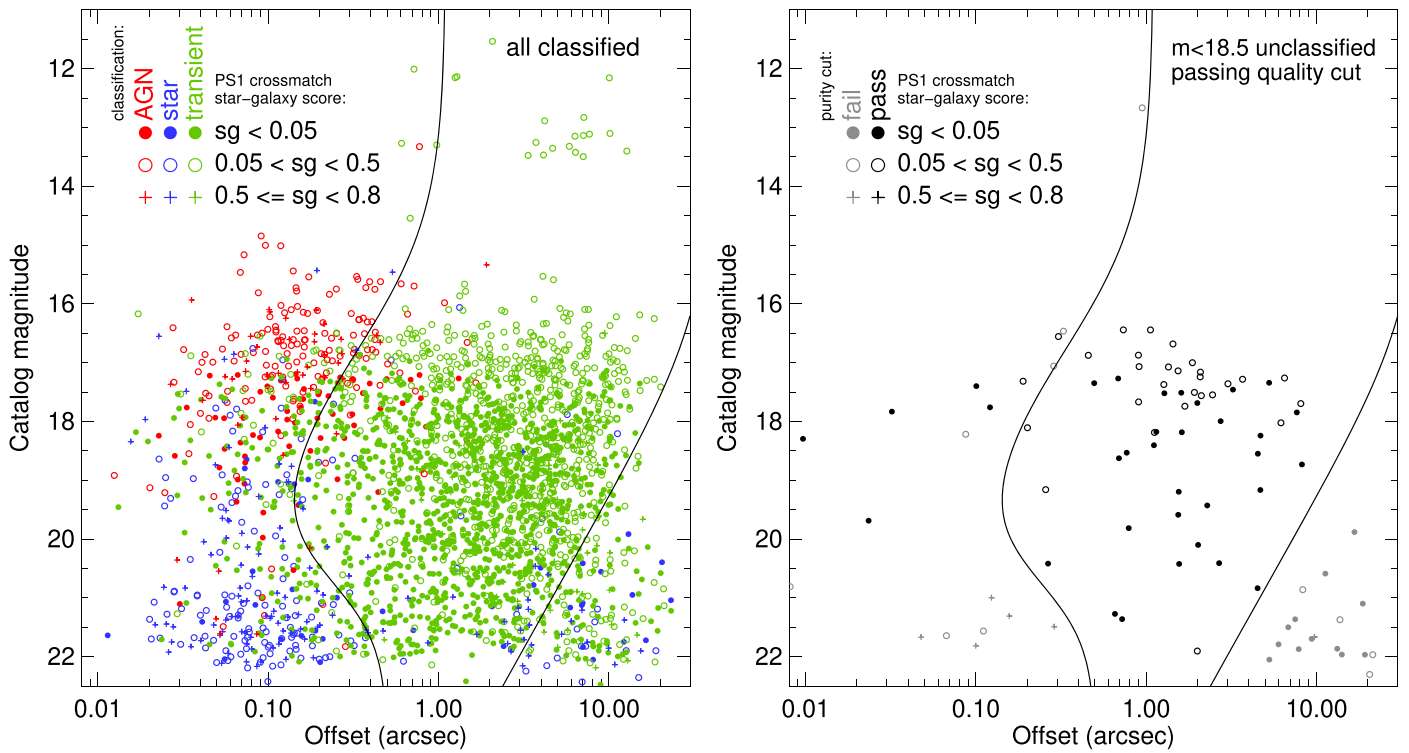
Here,  $\theta$  is the offset in arcseconds and  $m$  is the catalog magnitude of the candidate cross-match ( $r$  magnitude when available but another band is used if no  $r$  photometry exists). The exact form of these equations is arbitrary and was chosen largely using trial and error in order to avoid as much of the non-transient populations as possible while still including the vast majority of real transients. The slope term of  $-0.2$  is motivated by the assumption that galaxies on average have constant surface brightness on the sky (in general, dimmer galaxies are proportionally smaller), while the increasing offset at faint magnitudes reflects increasing positional measurement uncertainties for faint cross-matches.

Transient candidates within the middle region of the diagram (offset, probable matches) pass the purity cut if the cross-match has  $sgscore < 0.5$ . Transient candidates in the left of the diagram (coincident matches consistent with no offset) pass if the cross-match has  $sgscore < 0.05$ . Transients in the bottom right do not automatically pass the purity cut.

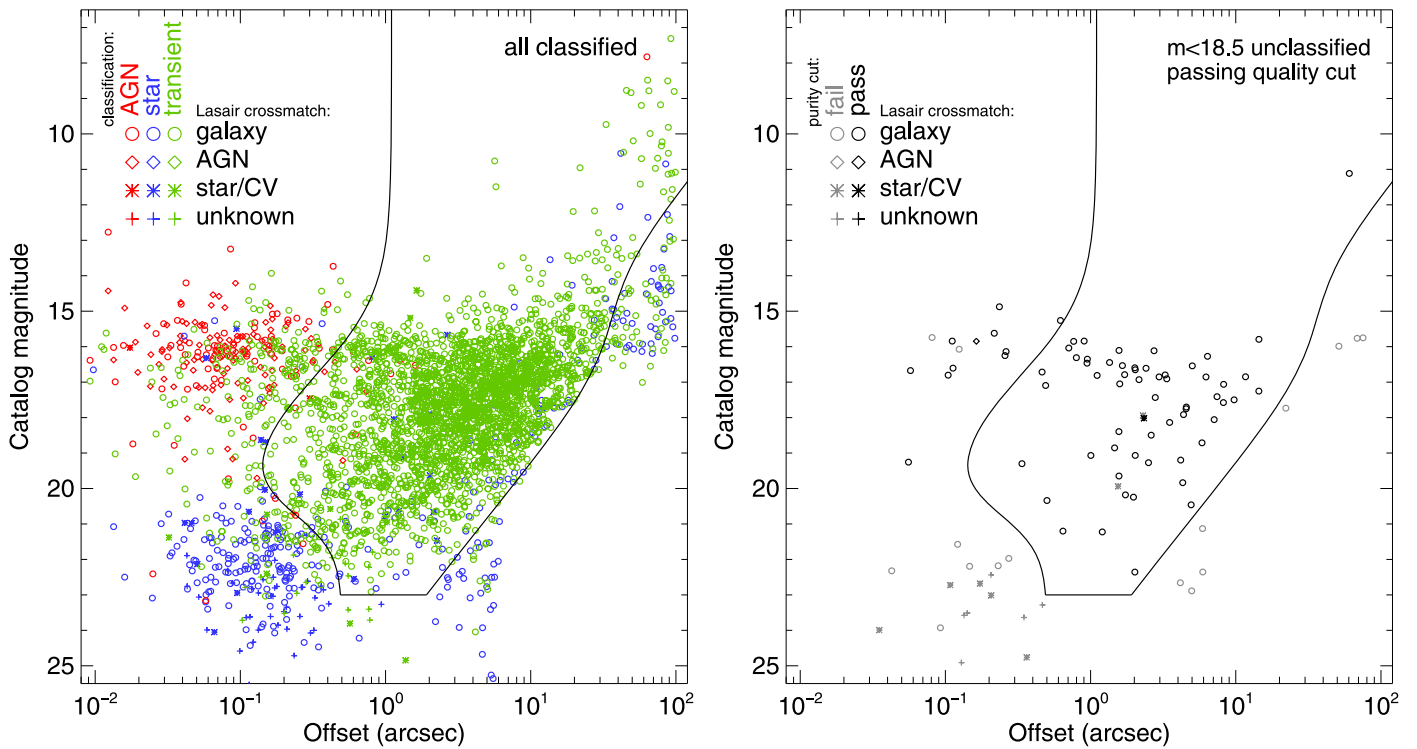
### B.2. Lasair Cross-match

The Pan-STARRS matches within the Avro packets only include the nearest three sources within  $30''$ , making them not particularly useful for transients in nearby and large galaxies. A useful  $sgscore$  value is also not always available (Section 2.1). We therefore also perform a second cross-match, relying on the cross-match tool provided by Lasair, which

<sup>34</sup> <https://github.com/dperley/ztf-bts-filters> (Perley 2020).



**Figure 11.** Angular offset vs. magnitude for Pan-STARRS catalog cross-matches. The left panel shows classified events saved to the program, color-coded by type (genuine transients, stars/CVs, or AGNs). The symbol indicates the star-galaxy score (Tachibana & Miller 2018; higher values are more star-like). The right panel shows unclassified transients passing our quality cut (Section 2.3). Events between the two lines automatically pass the purity cut as long as  $sg_{score} < 0.5$ . Events further to the left automatically pass if  $sg_{score} < 0.05$ . Events in the bottom right corner are likely to be chance associations.



**Figure 12.** Angular offset vs. magnitude for the most likely cross-matched source provided by Lasair. Panels and colors are as in Figure 11, but here the symbol indicates the type designation in the Lasair catalog. The lines are the same as in Figure 11. In this case, only events in the middle region associated with galaxies or “faint/unknown” automatically pass the purity cut.

searches for galaxy catalog associations out to much larger radii. For each candidate in our program, the list of potential cross-matches is obtained from this service, along with their  $g$

and  $r$  magnitudes and “type” (which can be “star,” “galaxy,” “agn,” and occasionally other types such as “cv.”) We reassign “agn” types to “galaxy” if the offset is more than  $1''$ .

A plot of offsets versus magnitudes from Lasair is shown in Figure 12. The general appearance is quite similar to the previous figure but extends to larger offsets and brighter galaxies. The spurious region of the diagram is mostly empty, since these are generally not cross-matched by Lasair in the first place.

We use the same equations to identify physical, non-coincident cross-matches as in Appendix B.1, although we also add a magnitude cut of  $m < 23$ . Transients with cross-matches in the central region of the diagram bounded by these three lines, for which the cross-match is a “galaxy,” automatically pass the purity cut.

### Appendix C

#### Choosing between Multiple Host Cross-Matches

It is frequently the case that there are several candidate host-galaxy cross-matches. In these cases it is important to determine which (if any) is the most credible host galaxy.

We use a simple least-likelihood method by calculating the probability that a given position, had it been randomly chosen across the entire sky, would be located as close to or closer to a galaxy as bright as or brighter than the host galaxy candidate under consideration. The general equation for this probability is

$$p = 1 - \exp(-\pi\theta^2\rho), \quad (\text{C1})$$

where  $\theta$  is the angular offset and  $\rho$  is the sky density of galaxies at least as bright (in apparent magnitude) as the putative host in the given filter band.

Because our cross-matching is automatic and must deal with shredded galaxies (which may have components very close to the transient), we employ a few approximations and modifications to

this basic approach. We use a simplified single power law of  $\rho = 220 \times 10^{0.55(m-18)} \text{ deg}^{-2}$ , which provides a reasonable approximation to the  $r$ -band number counts in Yasuda et al. (2001) at the bright end. Since we are only interested in highly probable cross-matches, we can also safely approximate  $1 - e^{-x}$  as  $x$ . Finally, because matches closer than  $1''$  are not meaningful for galaxies within our distance limit (ZTF pixels are  $1''$  in size and few galaxies are smaller than  $1''$  in size), we de-weight cross-matches of order  $1''$  by substituting  $\theta$  by  $\theta + 1''$ . Therefore, the actual equation used in practice is

$$p = \pi \left( \frac{\theta + 1}{3600} \right)^2 220 \times 10^{-0.55(m-18)}. \quad (\text{C2})$$

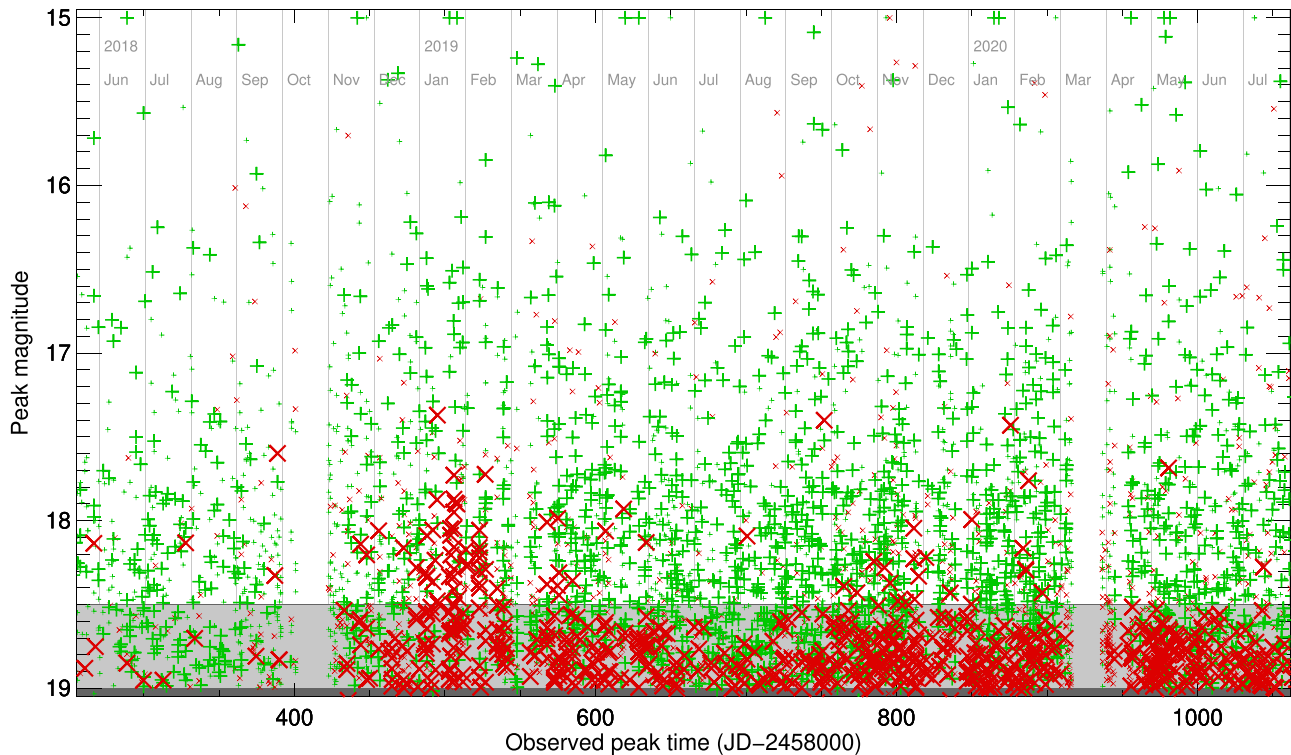
The galaxy with the lowest  $p$ -value is chosen as the association.

For the purposes of sample selection, only the relative value of  $p$  is meaningful; chance associations are rejected on a purely empirical basis as described in Appendix B.1. For host-galaxy assignment, we designate the transient as hostless if there is no galaxy within SDSS with a value of  $p < 0.1$  or, if SDSS is not available, no PS1 galaxy with  $p < 0.05$ . (The stricter cut for PS1 is due to the higher incidence of spurious sources in this catalog.) All SDSS host associations used in the host-galaxy analysis of this paper (Section 4.4) were manually vetted and the host reassigned if the automatically determined host was assessed to be inaccurate.

### Appendix D

#### Seasonal Dependence of Classification Rate

In Figure 13 we plot all candidate transients saved to the program by the time and magnitude of peak, color-coded by whether the event was classified or unclassified. This shows



**Figure 13.** Classification success or failure by time and magnitude of observed light-curve peak. Green “+” symbols indicate successful classifications and red “x” symbols indicate missing classifications. Large symbols indicate candidate transients that passed the quality and purity cuts; small symbols indicate candidate transients that did not pass these cuts. The primary determining factor governing classification success (at  $m < 18.5$  mag) is the impact of weather on spectroscopic follow-up runs: poor conditions affecting classical runs and P60 operations in early 2019 (and to a much lesser extent late 2019) led to a larger fraction of missing classifications in these periods. The gap in 2018 October is due to maintenance and the gap in 2020 March is due to extended bad weather. Transients with  $m < 15$  mag are fixed at  $m = 15$  mag in this plot.

**Table 3**  
Non-TNS Classifications

ZTF ID	IAU ID	Classification	Redshift	Reference/note
ZTF18aamfrvy	SN 2018ahe	...	0.01564	Added redshift from NED
ZTF18aazgfkq	SN 2018cmk	...	0.025724	Added redshift from NED
ZTF18abcfcoo	AT 2018cow	other	0.014145	New/peculiar transient class
ZTF18actuhrs	SN 2018evt	SN Ia-CSM	...	This work
ZTF19aadnwvc	AT 2019ye	SN Ia	0.077	ATEL 12426
ZTF19aagqkrq	AT 2019ahd	ILRT	...	This work
ZTF19aaniqrr	AT 2019cmw	other	0.519	New/peculiar transient class; paper in prep.
ZTF19aaplpaa	SN 2019cxx	...	0.025	Added redshift from NED
ZTF19aatubsj	SN 2019fdr	none	...	Possibly an AGN/NLSy1 based on late-time spectra
ZTF19aatevrp	SN 2019dke	...	0.010637	Added redshift from NED
ZTF19aavxfib	AT 2019gte	none	...	TDE classification is uncertain
ZTF19acdsqir	SN 2019sxd	...	0.066	Added redshift from NED
ZTF19acnfsij	ST 2019uiz	nova	M31	ATEL 13317
ZTF19acoaiub	AT 2019udc	ILRT	...	This work
ZTF19adakuos	AT 2019wvf	nova	M31	ATEL 13384
ZTF20aaertpj	AT 2020pv	SN Ib	0.02875	GCN 26703
ZTF20aaeuxqk	SN 2020ut	...	0.035	Revised redshift
ZTF20aakdppm	AT 2020ber	nova	M31	Recurrent nova M31N 1926-07c
ZTF20aatwonv	SN 2020euz	...	0.0226	Added redshift from NED
ZTF20abijfqq	SN 2020nlb	...	0.002432	Added redshift from NED
ZTF20abfhyil	SN 2020mrf	none	none	TNS classification based on featureless spectrum; probable CV

**Note.** An empty field (...) indicates that we retain the existing TNS classification or TNS redshift.

clearly the seasonal dependence of the success of our spectroscopic follow-up observations: during most of the year (summer and fall, especially) we are almost 100% complete to  $m < 18.5$  mag, but during the winter months when SEDM cannot operate for long stretches and classical follow-up runs may be weathered out, there are occasional periods where significant numbers of brighter transients are also missed.

### Appendix E Reclassifications

In almost all cases the classification we have used is the most recent classification associated with that object on TNS, although we remove some subtype information because (for spectra reported by us) we are not yet able to uniformly distinguish classical SN Ia subtypes or to separate SN IIP from SN IIL.

In a few cases we have used a different classification from what is reported on TNS—either because the classification was reported in a public reference other than TNS or because the TNS classification appears to be in error. We list these in Table 3.

### Appendix F The BTS Sample Explorer

To facilitate public use of our transient sample we have developed a web-based interface to display the sample (including objects that are not transients, which fail our cuts, etc.) and further filter it in various ways: by date of maximum light, by classification, by redshift, by peak magnitude (apparent or absolute), and by many other properties. It also provides P48 postage-stamp images of the transient and/or reference images, three-color PS1 images combined according to the method of Lupton et al. (2004), and light-curve plots.

The website front-end is built in basic PHP. A Python back-end is used to build and update the database; this back-end is the same as the one we have used to calculate timescales, cross-matches, and all other properties discussed in this paper. The back-end scripts to calculate these properties are executed automatically via `cron` every 3 hr. The scripts update the underlying data files (alert data, PS1 FITS images, TNS classification tables, etc.) by downloading from the relevant sources on a regular basis—daily for recent transients, less often for transients with a long history—or anytime a new event is saved to the program. Light-curve plots are also regenerated if new data points appear.

User queries execute extremely fast (within 1 s), although if images are requested these can take somewhat longer to load in their entirety for very large queries.

The interface is currently relatively basic: users can select from a range of options and enter start and end values for the purposes of filtering on various properties. More complex SQL-style queries are not yet possible although this is planned for the future.

Data are normally displayed as a table, but a grid mode can also be selected to specifically display the images (PS1 cutouts, light curves, or both).

Currently available data columns include the ZTF, IAU, and discoverer identifiers, the peak time and magnitude, the coordinates ( $\alpha$ ,  $\delta$ ), the half-peak-to-peak rise and fade times (and the sum of these values, the “duration”), the classification and redshift, the absolute magnitude of the transient, host absolute magnitude and color, Galactic latitude and Galactic extinction, and sample selection flags.

Further documentation can be found on the Explorer section of the BTS website.<sup>35</sup> An example screenshot is presented in Figure 14.

<sup>35</sup> <https://sites.astro.caltech.edu/ztf/bts/explorer.php>



## ORCID iDs

Daniel A. Perley  <https://orcid.org/0000-0001-8472-1996>  
 Christoffer Fremling  <https://orcid.org/0000-0002-4223-103X>  
 Jesper Sollerman  <https://orcid.org/0000-0003-1546-6615>  
 Adam A. Miller  <https://orcid.org/0000-0001-9515-478X>  
 Yashvi Sharma  <https://orcid.org/0000-0003-4531-1745>  
 Eric C. Bellm  <https://orcid.org/0000-0001-8018-5348>  
 Rahul Biswas  <https://orcid.org/0000-0002-5741-7195>  
 Thomas G. Brink  <https://orcid.org/0000-0001-5955-2502>  
 Rachel J. Bruch  <https://orcid.org/0000-0002-0786-7307>  
 Kishalay De  <https://orcid.org/0000-0002-8989-0542>  
 Richard Dekany  <https://orcid.org/0000-0002-5884-7867>  
 Dmitry A. Duvvuri  <https://orcid.org/0000-0001-5060-8733>  
 Alexei V. Filippenko  <https://orcid.org/0000-0003-3460-0103>  
 Avishay Gal-Yam  <https://orcid.org/0000-0002-3653-5598>  
 Ariel Goobar  <https://orcid.org/0000-0002-4163-4996>  
 Melissa L. Graham  <https://orcid.org/0000-0002-9154-3136>  
 Anna Y. Q. Ho  <https://orcid.org/0000-0002-9017-3567>  
 Ido Irani  <https://orcid.org/0000-0002-7996-8780>  
 Mansi M. Kasliwal  <https://orcid.org/0000-0002-5619-4938>  
 Young-Lo Kim  <https://orcid.org/0000-0002-1031-0796>  
 S. R. Kulkarni  <https://orcid.org/0000-0001-5390-8563>  
 Ashish Mahabal  <https://orcid.org/0000-0003-2242-0244>  
 Frank J. Masci  <https://orcid.org/0000-0002-8532-9395>  
 Shaunak Modak  <https://orcid.org/0000-0002-8532-827X>  
 James D. Neill  <https://orcid.org/0000-0002-0466-1119>  
 Jakob Nordin  <https://orcid.org/0000-0001-8342-6274>  
 Reed L. Riddle  <https://orcid.org/0000-0002-0387-370X>  
 Maayane T. Soumagnac  <https://orcid.org/0000-0001-6753-1488>  
 Nora L. Strotjohann  <https://orcid.org/0000-0002-4667-6730>  
 Steve Schulze  <https://orcid.org/0000-0001-6797-1889>  
 Kirsty Taggart  <https://orcid.org/0000-0002-5748-4558>  
 Anastasios Tzanidakis  <https://orcid.org/0000-0003-0484-3331>  
 Lin Yan  <https://orcid.org/0000-0003-1710-9339>

## References

Anderson, J. P., González-Gaitán, S., Hamuy, M., et al. 2014, *ApJ*, 786, 67  
 Anderson, J. P., & James, P. A. 2008, *MNRAS*, 390, 1527  
 Anderson, J. P., James, P. A., Habergham, S. M., Galbany, L., & Kuncarayakti, H. 2015, *PASA*, 32, e019  
 Andreoni, I., Kool, E. C., Sagues Carracedo, A., et al. 2020, arXiv:2008.00008  
 Arcavi, I., Gal-Yam, A., Kasliwal, M. M., et al. 2010, *ApJ*, 721, 777  
 Arcavi, I., Wolf, W. M., Howell, D. A., et al. 2016, *ApJ*, 819, 35  
 Arnett, W. D., Bahcall, J. N., Kirshner, R. P., & Woosley, S. E. 1989, *ARA&A*, 27, 629  
 Barris, B. J., & Tonry, J. L. 2006, *ApJ*, 637, 427  
 Bellm, E. C., Kulkarni, S. R., Barlow, T., et al. 2019b, *PASP*, 131, 068003  
 Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019a, *PASP*, 131, 018002  
 Berger, E., Soderberg, A. M., Chevalier, R. A., et al. 2009, *ApJ*, 699, 1850  
 Blagorodnova, N., Gezari, S., Hung, T., et al. 2017, *ApJ*, 844, 46  
 Blagorodnova, N., Karambelkar, V., Adams, S. M., et al. 2020, *MNRAS*, 496, 5503  
 Blagorodnova, N., Neill, J. D., Walters, R., et al. 2018, *PASP*, 130, 035003  
 Blanton, M. R., Kazin, E., Muna, D., Weaver, B. A., & Price-Whelan, A. 2011, *AJ*, 142, 31  
 Boissier, S., & Prantzos, N. 2009, *A&A*, 503, 137  
 Bond, H. E., Bedin, L. R., Bonanos, A. Z., et al. 2009, *ApJL*, 695, L154  
 Botticella, M. T., Pastorello, A., Smartt, S. J., et al. 2009, *MNRAS*, 398, 1041  
 Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, *MNRAS*, 351, 1151  
 Cai, Y. Z., Pastorello, A., Fraser, M., et al. 2019, *A&A*, 632, L6  
 Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560

Coppejans, D. L., Margutti, R., Terreran, G., et al. 2020, *ApJL*, 895, L23  
 Dahlen, T., Strolger, L.-G., Riess, A. G., et al. 2004, *ApJ*, 613, 189  
 Dahlen, T., Strolger, L.-G., & Riess, A. G. 2008, *ApJ*, 681, 462  
 Dauphin, F., Hosseinzadeh, G., Villar, V., Berger, E., & Gomez, S. 2020, AAS Meeting, 235, 276.18  
 De Cia, A., Gal-Yam, A., Rubin, A., et al. 2018, *ApJ*, 860, 100  
 de Jaeger, T., Zheng, W., Stahl, B. E., et al. 2019, *MNRAS*, 490, 2799  
 De, K., Kasliwal, M. M., Tzanidakis, A., et al. 2020, arXiv:2004.09029  
 De, K., Tzanidakis, A., Kasliwal, M. M., et al. 2019, Transient Name Server AstroNote, 124, 1  
 Dekany, R., Smith, R. M., Riddle, R., et al. 2020, *PASP*, 132, 038001  
 Dilday, B., Smith, M., Bassett, B., et al. 2010, *ApJ*, 713, 1026  
 Drout, M. R., Chornock, R., Soderberg, A. M., et al. 2014, *ApJ*, 794, 23  
 Duvvuri, D. A., Mahabal, A., Masci, F. J., et al. 2019, *MNRAS*, 489, 3582  
 Dyer, M. J., Dhillon, V. S., Littlefair, S., et al. 2018, *Proc. SPIE*, 10704, 107040C  
 Filippenko, A. V. 1997, *ARA&A*, 35, 309  
 Filippenko, A. V., Chornock, R., Swift, B., et al. 2003, *IAUC*, 8159, 2  
 Foley, R. J., Challis, P. J., Chornock, R., et al. 2013, *ApJ*, 767, 57  
 Foley, R. J., Chornock, R., Filippenko, A. V., et al. 2009, *AJ*, 138, 376  
 Fremling, C., Miller, A. A., Sharma, Y., et al. 2020, *ApJ*, 895, 32  
 Frohmaier, C., Sullivan, M., Maguire, K., & Nugent, P. 2018, *ApJ*, 858, 50  
 Frohmaier, C., Sullivan, M., Nugent, P. E., et al. 2019, *MNRAS*, 486, 2308  
 Galbany, L., Hamuy, M., Phillips, M. M., et al. 2016a, *AJ*, 151, 33  
 Galbany, L., Stanishev, V., Mourão, A. M., et al. 2014, *A&A*, 572, A38  
 Galbany, L., Stanishev, V., Mourão, A. M., et al. 2016b, *A&A*, 591, A48  
 Gal-Yam, A. 2012, *Sci*, 337, 927  
 Gal-Yam, A. 2019, *ARA&A*, 57, 305  
 Gal-Yam, A., Kasliwal, M. M., Arcavi, I., et al. 2011, *ApJ*, 736, 159  
 Gal-Yam, A., Mazzali, P. A., Manulis, I., & Bishop, D. 2013, *PASP*, 125, 749  
 Gillanders, J. H., Sim, S. A., & Smartt, S. J. 2020, *MNRAS*, 497, 246  
 Graham, M. J., Kulkarni, S. R., Bellm, E. C., et al. 2019, *PASP*, 131, 078001  
 Graham, M. L., Sand, D. J., Bildfell, C. J., et al. 2012, *ApJ*, 753, 68  
 Graur, O., Bianco, F. B., Huang, S., et al. 2017a, *ApJ*, 837, 120  
 Graur, O., Bianco, F. B., Modjaz, M., et al. 2017b, *ApJ*, 837, 121  
 Graur, O., Poznanski, D., Maoz, D., et al. 2011, *MNRAS*, 417, 916  
 Groot, P. J. 2019, *NatAs*, 3, 1160  
 Gutiérrez, C. M., Arnold, D., Copley, D., et al. 2019, *AN*, 340, 40  
 Hakobyan, A. A., Adibekyan, V. Z., Aramyan, L. S., et al. 2012, *A&A*, 544, A81  
 Hakobyan, A. A., Petrosian, A. R., McLean, B., et al. 2008, *A&A*, 488, 523  
 Ho, A. Y. Q., Goldstein, D. A., Schulze, S., et al. 2019b, *ApJ*, 887, 169  
 Ho, A. Y. Q., Kulkarni, S. R., Perley, D. A., et al. 2020c, *ApJ*, 902, 86  
 Ho, A. Y. Q., Perley, D. A., Beniamini, P., et al. 2020a, arXiv:2006.10761  
 Ho, A. Y. Q., Perley, D. A., Kulkarni, S. R., et al. 2020b, *ApJ*, 895, 49  
 Ho, A. Y. Q., Phinney, E. S., Ravi, V., et al. 2019a, *ApJ*, 871, 73  
 Horiuchi, S., Beacom, J. F., Kochanek, C. S., et al. 2011, *ApJ*, 738, 154  
 Hosseinzadeh, G., Dauphin, F., Villar, V. A., et al. 2020, arXiv:2008.04912  
 Irani, I., Schulze, S., Gal-Yam, A., et al. 2019, *ApJ*, 887, 127  
 Jencson, J. E., Kasliwal, M. M., Adams, S. M., et al. 2019, *ApJ*, 886, 40  
 Jiang, H., Hu, Z., Xu, M., et al. 2018, *Proc. SPIE*, 10702, 107022L  
 Jzezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, 873, 111  
 Kaiser, N., Aussel, H., Burke, B. E., et al. 2002, *Proc. SPIE*, 4836, 154  
 Kasliwal, M. M. 2011, PhD thesis, California Institute of Technology  
 Kasliwal, M. M., Cannella, C., Bagdasaryan, A., et al. 2019, *PASP*, 131, 038003  
 Kasliwal, M. M., Kulkarni, S. R., Gal-Yam, A., et al. 2012, *ApJ*, 755, 161  
 Kelly, P. L., & Kirshner, R. P. 2012, *ApJ*, 759, 107  
 Kulkarni, S. R. 2020, arXiv:2004.03511  
 Kulkarni, S. R., Ofek, E. O., Rau, A., et al. 2007, *Natur*, 447, 458  
 Kuncarayakti, H., Anderson, J. P., Galbany, L., et al. 2018, *A&A*, 613, A35  
 Li, W., Filippenko, A. V., Chornock, R., et al. 2003, *PASP*, 115, 453  
 Li, W., Leaman, J., Chornock, R., et al. 2011, *MNRAS*, 412, 1441  
 Lunnan, R., Yan, L., Perley, D. A., et al. 2020, *ApJ*, 901, 61  
 Lupton, R., Blanton, M. R., Fekete, G., et al. 2004, *PASP*, 116, 133  
 Madau, P., & Dickinson, M. 2014, *ARA&A*, 52, 415  
 Maeder, A., & Meynet, G. 2000, *ARA&A*, 38, 143  
 Mahabal, A., Rebbapragada, U., Walters, R., et al. 2019, *PASP*, 131, 038002  
 Mannucci, F., Della Valle, M., & Panagia, N. 2006, *MNRAS*, 370, 773  
 Margutti, R., Metzger, B. D., Chornock, R., et al. 2019, *ApJ*, 872, 18  
 Masci, F. J., Laher, R. R., Rusholme, B., et al. 2019, *PASP*, 131, 018003  
 Mattila, S., Dahlen, T., Efstathiou, A., et al. 2012, *ApJ*, 756, 111  
 Maund, J. R. 2018, *MNRAS*, 476, 2629  
 Maund, J. R., Smartt, S. J., Kudritzki, R. P., et al. 2006, *MNRAS*, 369, 390  
 McBrien, O. R., Smartt, S. J., Chen, T.-W., et al. 2019, *ApJL*, 885, L23  
 Melinder, J., Dahlen, T., Mencia Trinchant, L., et al. 2012, *A&A*, 545, A96

- Miller, J. S., & Stone, R. P. S. 1993, The KAST Double Spectrograph, Tech. Rep. 66, Lick Observatory
- Moriya, T. J., Tominaga, N., Langer, N., et al. 2014, *A&A*, **569**, A57
- Muthukrishna, D., Narayan, G., Mandel, K. S., Biswas, R., & Hložek, R. 2019, *PASP*, **131**, 118002
- Nordin, J., Brinnel, V., van Santen, J., et al. 2019, *A&A*, **631**, A147
- Nyholm, A., Sollerman, J., Tartaglia, L., et al. 2020, *A&A*, **637**, A73
- Ofek, E. O., Arcavi, I., Tal, D., et al. 2014, *ApJ*, **788**, 154
- Ofek, E. O., Rabinak, I., Neill, J. D., et al. 2010, *ApJ*, **724**, 1396
- Oke, J. B., Cohen, J. G., Carr, M., et al. 1995, *PASP*, **107**, 375
- Oke, J. B., & Gunn, J. E. 1982, *PASP*, **94**, 586
- Pastorello, A., Della Valle, M., Smartt, S. J., et al. 2007, *Natur*, **449**, 1
- Pastorello, A., & Fraser, M. 2019, *NatAs*, **3**, 676
- Pastorello, A., Mason, E., Taubenberger, S., et al. 2019, *A&A*, **630**, A75
- Pastorello, A., Valenti, S., Zampieri, L., et al. 2009, *MNRAS*, **394**, 2266
- Pastorello, A., Zampieri, L., Turatto, M., et al. 2004, *MNRAS*, **347**, 74
- Patterson, M. T., Bellm, E. C., Rusholme, B., et al. 2019, *PASP*, **131**, 018001
- Pejcha, O., Metzger, B. D., & Tomida, K. 2016, *MNRAS*, **455**, 4351
- Perets, H. B., Gal-Yam, A., Mazzali, P. A., et al. 2010, *Natur*, **465**, 322
- Perley, D. A. 2020, ZTF BTS Alert Filter Code, Zenodo doi:10.5281/zenodo.4054129
- Perley, D. A., Mazzali, P. A., Yan, L., et al. 2019, *MNRAS*, **484**, 1031
- Phillips, M. M. 1993, *ApJL*, **413**, L105
- Piascik, A. S., Steele, I. A., Bates, S. D., et al. 2014, *Proc. SPIE*, **9147**, 91478H
- Prajs, S., Sullivan, M., Smith, M., et al. 2017, *MNRAS*, **464**, 3568
- Prentice, S. J., Maguire, K., Smartt, S. J., et al. 2018, *ApJL*, **865**, L3
- Pursiainen, M., Childress, M., Smith, M., et al. 2018, *MNRAS*, **481**, 894
- Quimby, R. M., Yuan, F., Akerlof, C., & Wheeler, J. C. 2013, *MNRAS*, **431**, 912
- Rest, A., Garnavich, P. M., Khatami, D., et al. 2018, *NatAs*, **2**, 307
- Rigault, M., Neill, J. D., Blagorodnova, N., et al. 2019, *A&A*, **627**, A115
- Rubin, A., & Gal-Yam, A. 2016, *ApJ*, **828**, 111
- Sanders, N. E., Soderberg, A. M., Gezari, S., et al. 2015, *ApJ*, **799**, 208
- Schipani, P., Campana, S., Claudi, R., et al. 2018, *Proc. SPIE*, **10702**, 107020F
- Schlafly, E. F., & Finkbeiner, D. P. 2011, *ApJ*, **737**, 103
- Schulze, S., Yaron, O., Sollerman, J., et al. 2020, arXiv:2008.05988
- Shappee, B., Prieto, J., Stanek, K. Z., et al. 2014, AAS Meeting, **223**, 236.03
- Smith, K. W., Williams, R. D., Young, D. R., et al. 2019, *RNAAS*, **3**, 26
- Smith, N., Li, W., Silverman, J. M., Ganeshalingam, M., & Filippenko, A. V. 2011, *MNRAS*, **415**, 773
- Steele, I. A., Smith, R. J., Rees, P. C., et al. 2004, *Proc. SPIE*, **5489**, 679
- Strolger, L.-G., Dahlen, T., Rodney, S. A., et al. 2015, *ApJ*, **813**, 93
- Strolger, L. G., Smith, R. C., Suntzeff, N. B., et al. 2002, *AJ*, **124**, 2905
- Sullivan, M., Le Borgne, D., Pritchett, C. J., et al. 2006, *ApJ*, **648**, 868
- Tachibana, Y., & Miller, A. A. 2018, *PASP*, **130**, 128001
- Taggart, K., & Perley, D. 2019, arXiv:1911.09112
- Taylor, M., Cinabro, D., Dilday, B., et al. 2014, *ApJ*, **792**, 135
- Thompson, T. A., Prieto, J. L., Stanek, K. Z., et al. 2009, *ApJ*, **705**, 1364
- Tonry, J. L., Denneau, L., Heinze, A. N., et al. 2018, *PASP*, **130**, 064505
- Tsuna, D., Ishii, A., Kuriyama, N., Kashiyama, K., & Shigeyama, T. 2020, *ApJL*, **897**, L44
- Valenti, S., Howell, D. A., Stritzinger, M. D., et al. 2016, *MNRAS*, **459**, 3939
- van Velzen, S., Gezari, S., Cenko, S. B., et al. 2019, *ApJ*, **872**, 198
- van Velzen, S., Gezari, S., Hammerstein, E., et al. 2020, arXiv:2001.01409
- Villar, V. A., Berger, E., Metzger, B. D., & Guillochon, J. 2017, *ApJ*, **849**, 70
- Villar, V. A., Berger, E., Miller, G., et al. 2019, *ApJ*, **884**, 83
- Villar, V. A., Hosseinzadeh, G., Berger, E., et al. 2020, arXiv:2008.04921
- Wanderman, D., & Piran, T. 2010, *MNRAS*, **406**, 1944
- Wiseman, P., Pursiainen, M., Childress, M., et al. 2020, *MNRAS*, **498**, 2575
- Xiao, L., Galbany, L., Eldridge, J. J., & Stanway, E. R. 2019, *MNRAS*, **482**, 384
- Yasuda, N., Fukugita, M., Narayanan, V. K., et al. 2001, *AJ*, **122**, 1104
- Yoon, S. C., & Langer, N. 2005, *A&A*, **443**, 643
- Zinn, P. C., Stritzinger, M., Braithwaite, J., et al. 2012, *A&A*, **538**, A30