

Search for Planets like Earth around Late-M Dwarfs: Precise Radial Velocity Survey with IRD

PI: Bun'ei Sato (Tokyo Institute of Technology)

Co-PI: Nagayoshi Ohashi (NAOJ, Subaru)

E. AKIYAMA¹, W. AOKI², C. BEICHMAN³, T. BRANDT⁴, G. CATALDI⁵, C. CLERGEON⁵, T. CURRIE⁵, R. DONG⁶, Y. FUJII^{7,8},
H. FUJIWARA⁵, A. FUKUI², H. GENDA^{7,8}, T. GROFF⁹, O. GUYON^{5,10,11}, D. HALL³¹, H. HARAKAWA², J. HASHIMOTO^{2,11},
Y. HAYANO², M. HAYASHI², K. G. HELMINIAK¹², T. HENNING¹³, T. HIRANO⁸, K. HODAPP³¹, Y. HORI^{2,11}, Y. IKEDA¹⁴,
S. INUTSUKA²⁴, H. T. ISHIKAWA²¹, M. ISHIZUKA¹⁵, H. IZUMIURA², S. JACOBSON³¹, M. JANSON¹⁷, N. JOVANOVIĆ²³,
E. KAMBE², H. KAWAHARA¹⁵, T. KODAMA¹⁵, Y. KOIZUMI⁸, E. KOKUBO², M. KONISHI^{2,11}, T. KOTANI^{2,11}, T. KUDO²,
T. KUROKAWA^{2,11}, N. KUSAKABE^{2,11}, M. KUZUHARA^{2,11}, J. KWON¹⁶, C. LEE⁵, J. LIVINGSTON¹⁵, M. MACHIDA²⁸,
T. MATSUO²⁷, D. MAWET²³, M. MCELWAIN⁹, V. MEADOWS²⁹, E. MIEDA⁵, T. MIZUKI¹⁶, J. MORINO², T. NAGATA²⁰,
T. NAKAGAWA¹⁶, T. NAKAJIMA^{2,11}, N. NARITA¹⁵, J. NISHIKAWA^{2,11,21}, S. NISHIYAMA¹⁸, H. NOMURA⁸, M. OGIHARA²,
D. OH²⁵, M. OMIYA^{2,11}, S. OSHINO², T. PYO⁵, E. SERABYN³, M. SITKO¹⁹, H. SUTO^{2,11}, R. SUZUKI², Y. TAKAGI⁵,
H. TAKAMI², T. TAKARADA⁸, N. TAKATO², M. TAMURA^{2,11,15}, Y. TANAKA³⁰, H. TERADA², R. A. TORRES¹⁷,
E. L. TURNER²², A. UEDA², T. USUDA², T. UYAMA¹⁵, S. VIEVARD⁵, J. WANG²³, J. WISNIEWSKI²⁶, AND Y. YANG²¹

1. Hokkaido University; 2. NAOJ; 3. JPL/Caltech; 4. UC Santa Barbara; 5. Subaru Telescope; 6. University of Victoria; 7. ELSI; 8. Tokyo Institute of Technology; 9. NASA Goddard; 10. University of Arizona; 11. Astrobiology Center, NINS; 12. Nicolaus Copernicus Astronomical Center; 13. MPIA; 14. Photocoding; 15. University of Tokyo; 16. ISAS/JAXA; 17. Stockholm University; 18. Miyagi University of Education; 19. Space Science Institute; 20. Kyoto University; 21. SOKENDAI; 22. Princeton University; 23. Caltech; 24. Nagoya University; 25. National Meteorological Satellite Center; 26. University of Oklahoma; 27. Osaka University; 28. Kyushu University; 29. University of Washington; 30. TUAT; 31. University of Hawaii, IfA

ABSTRACT

Small, rocky planets in the habitable zones of their host stars are the most promising places to search for alien life. With the discovery of thousands of extrasolar planets, we now know that small planets like Earth are also some of the most abundant. Detecting these planets around Sun-like stars is very challenging due to the low probability of transits (<1%), the infrequency of transits even if they do occur (~1/year), and the tiny radial velocity (RV) signal (~10 cm s⁻¹). Late-M dwarf stars (LMDs; spectral type later than M4) are much more promising targets for a Doppler search. Low stellar masses (~0.1–0.2 M_{\odot}) and short planetary orbital periods (~3–30 days) result in much larger RV signals (~1–5 m s⁻¹) for ~1–3 Earth mass planets (hereafter Earth-mass planets) in the habitable zone (HZ). However, LMDs are very faint at optical wavelengths, and there have as yet been no systematic planet searches around nearby LMDs.

We propose an extensive, precise near-infrared (NIR) RV survey focused on LMDs using the newly-commissioned InfraRed Doppler (IRD) instrument in the framework of the Subaru Strategic Program (SSP). IRD is ideally suited to measure precise RVs of LMDs because: (1) the stars have a flux peak in NIR, (2) there are many absorption lines in their NIR spectra, and (3) those lines are less affected by stellar activity than optical lines. IRD is an echelle spectrometer that covers the NIR wavelengths from 0.97 to 1.75 μm with a high spectral resolution (70,000); it uses a laser frequency comb as an extremely precise wavelength calibrator. IRD's high instrumental stability and its use of a laser frequency comb allow us to measure the RVs of LMDs with a precision of 2 m s⁻¹ or better.

The goals of our survey are: (1) to discover Earth-mass planets in the HZ around these low-mass stars, and (2) to uncover the distribution of planetary systems containing Earth-mass planets and more massive planets. To achieve these goals, we will carry out an extensive high-precision RV survey of 60 carefully selected low-mass stars using 175 observing nights of the Subaru telescope over 5 years. Our initial target samples are collected from the literature and from the results of our pre-selection observations with optical spectroscopy of H α lines. This will screen out active stars, leaving us with the targets best suited for very high precision RV measurements. According to a simulation of the RV observations based on both theoretical and empirical planet populations, our proposed observations can discover Earth-mass planets in close-in orbits with periods less than a few tens of days and super-Earths (~5 Earth-masses) in orbits with periods less than a few hundred days. So far, only two Earth-“mass” planets in HZ have been reported with

the RV method, and M-dwarf planetary companions with 1-100 Earth-masses beyond the snow line are known almost only from microlensing surveys. Although other infrared Doppler projects also aim to search for Earth-like planets around early-to-late M dwarfs, only IRD-SSP will have the telescope aperture and observing strategy to successfully target LMDs.

In our 5-year survey, we expect to discover more than 60 planets, including a few tens of Earth-like planets, with some of those in the HZ. The numbers are large enough to uncover the distribution of planetary systems around low-mass stars solely by our survey, as optical RV surveys have done for higher-mass planets around solar-type stars. This will impose several critical constraints on the less-explored planetary formation theories for low-mass stars. Our newly detected planets will also provide some of the best candidates for direct imaging characterization with the next generation 30-m class telescopes such as TMT/PSI. Even non-detections with IRD will provide crucial inputs such as M-dwarf planetary demographics and target selections for future planet searches around nearby M dwarfs.

1. OUTLINE

This proposal is organized as follows. In Section 2, we describe the background of our proposed RV survey for late-M dwarfs including the achievements of previous RV surveys, key questions that remain to be solved, and the necessity of IRD. In Section 3, we define the goals of the IRD-SSP survey and validate them from the viewpoints of expected planet yield, the number of requested nights, the advantages of IRD compared with the other explorations of extrasolar planets, and follow-up prospects for planets discovered by IRD-SSP. Section 4 is devoted to an instrumental description of IRD, which can now achieve an RV precision better than 2 m s^{-1} . In Section 5, we describe our survey plan including sample selection, observation plan, and expected results based on a simulation. In Section 6, we describe data analysis that consists of the reduction of echelle spectra, precise RV measurements, stellar activity corrections, and determination of the stellar parameters of our target stars. Section 7 contains ancillary science related to the IRD-SSP: chemical abundance studies of M dwarfs and transit follow-up observations. Section 8 and Section 9 are devoted to team organization, and data archive and release policy, respectively, and we conclude the proposal in Section 10.

2. BACKGROUND

2.1. Brief summary of previous extrasolar planet explorations: all roads leading to M dwarf planets

Since the first discovery of an extrasolar planet around 51 Peg (Mayor & Queloz 1995), more than 3500 extrasolar planets have been confirmed around various types of stars, and those planets show a remarkable diversity in their properties (e.g. Udry & Santos 2007). The discoveries have mostly been made for solar-type stars (i.e. FGK dwarfs) by the RV and transit methods. In contrast, planets around other types of stars, especially late-M type dwarfs (spectral type later than M4), have not been well explored yet.

Past RV surveys have shown that planets are common around solar-type stars. About 70% of nearby solar-type stars have at least one planet within an orbital period (P) of 10 yr (Mayor et al. 2011), and low-mass planets are more abundant than massive giant planets (Howard et al. 2010; Mayor et al. 2011). The occurrence rate of super-Earths and Neptune-mass planets on tight orbits ($P < 50$ days) is estimated to be about 20–30%, while that of giant planets ($\gtrsim 10 M_{\oplus}$; M_{\oplus} is the Earth mass) on tight orbits is $\lesssim 5\%$ (Howard et al. 2010; Mayor et al. 2011; Figure 1). The occurrence rate of giant planets increases with stellar mass from 2% around M-type dwarfs to 20% around A-type dwarfs (Johnson et al. 2010), which may correlate with the masses of the protoplanetary disks. Low-mass planets are very common around M dwarfs, and are present around at least 50% of such stars (Bonfils et al. 2013). Meanwhile, occurrence rates for very cool (low-mass) stars ($< 0.3 M_{\odot}$) are not estimated yet due to the small number of observations. Whether the occurrence rates of low-mass planets continue to increase even around the lowest-mass stars remains an open question.

Transit surveys, particularly with the *Kepler* spacecraft, have now discovered more than 3000 planets, including many Earth-sized ones (Borucki et al. 2011) and even planets in the Habitable Zone (HZ), in which they can retain liquid water on the surface, around solar-type stars. The census of the *Kepler* planets (within 0.25 au or 50 days) indicates that small planets are abundant around solar-type stars and that the occurrence rate rises substantially as the stellar mass decreases; Howard et al. (2012) reported that the occurrence rate of 2–4 R_{\oplus} planets around M0 dwarfs is about 25%, while that around F dwarfs is as low as 5%. This statement was later reinforced by Mulders et al. (2015), who reported that the occurrence rate of small planets (1–4 R_{\oplus}) is an order of magnitude higher around M dwarfs compared to solar-type stars for the entire range of orbital distances surveyed by *Kepler* (Figure 2). The *Kepler*-based planet occurrence rate was calculated to be 2.5 ± 0.2 planets per early M dwarf, for planets with radii of 1–4 R_{\oplus} and orbital periods shorter than 200 days (Dressing & Charbonneau 2015, hereafter DC15). In addition, they found that Earth-sized planets (1–1.5 R_{\oplus}) and super-Earths (1.5–2.0 R_{\oplus}) are present with the occurrence rates of $0.56^{+0.06}_{-0.05}$ and $0.46^{+0.07}_{-0.05}$ per early M dwarf, over orbital periods shorter than 50 days. Most interestingly, the occurrence rates of Earth-sized planets and super-Earths within the HZ around M dwarfs are $0.16^{+0.17}_{-0.07}$ and $0.12^{+0.10}_{-0.05}$, respectively, even adopting a conservative definition of the HZ (DC15). This implies that RV monitoring of dozens of nearby M dwarfs is

enough to detect Earth-size planets, provided the RV instrument has sufficient measurement precision to detect such small Doppler signals.

Earth-size planets, however, are not necessarily Earth-mass planets. Masses of Earth-size planets can actually range from ~ 1 to $10 M_{\oplus}$ depending on their composition; mass measurements are needed to use the distribution of the *Kepler* planets to test planet formation theories. In addition, the planets found by *Kepler* are mostly located around faint stars far away from Earth (~ 1 kpc), preventing us from following them up with ground-based facilities to investigate their properties, such as mass and atmosphere. Nearly all *Kepler* stars having planets in or near the HZ are too faint to conduct detailed follow-up observations.

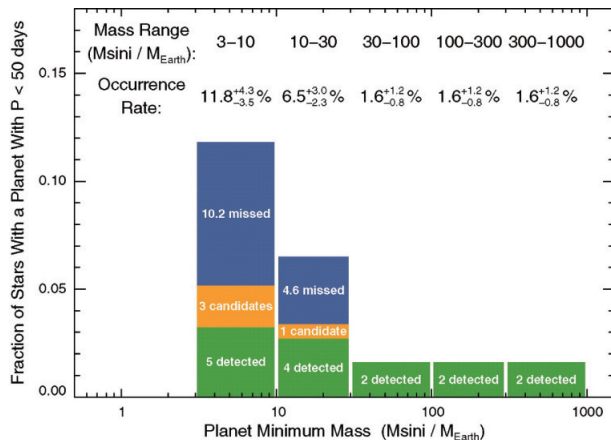


Figure 1. Fraction of solar-type stars hosting a planet with $P < 50$ days against planet minimum mass derived from a Doppler survey at Keck (Howard et al. 2010). Extrapolation of a power-law mass distribution fitted to the measurements predicts that 23% of solar-type stars harbor a close-in Earth-mass planet.

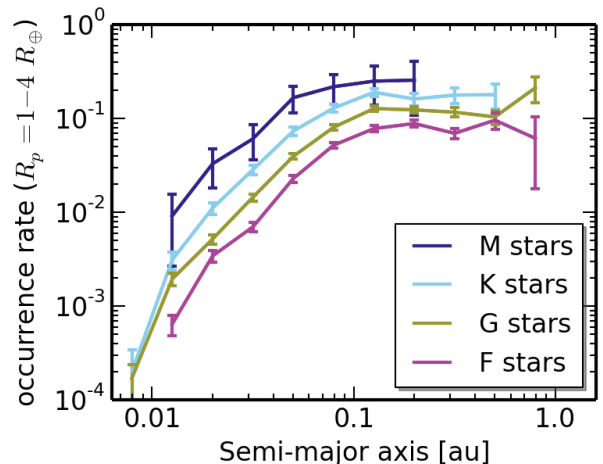


Figure 2. Planet occurrence rate as a function of semi-major axis for stars of different spectral type (Mulders et al. 2015). Close-in small planets are most common around M dwarfs, the lowest mass stars.

These studies of extrasolar planets have primarily focused on planets around solar-type stars. Planets around lower-mass stars are now known to be abundant, but the faintness of these stars at visible wavelengths has made follow-up difficult even with large-aperture telescopes like Subaru. Current planet searches therefore do not target stars beyond mid-M dwarfs ($\sim 0.3 M_{\odot}$) (Delfosse et al. 1998a; Butler et al. 2006; Endl et al. 2006; Bonfils et al. 2013). Apart from the transit or RV technique, gravitational microlensing surveys for extrasolar planets, in which the lensing objects tend to be low-mass stars, have revealed a population of super-Earth to Neptune-sized planets (Cassan et al. 2012; Suzuki et al. 2016). However, their sensitivity is limited to relatively large orbital distances (0.5–10 au) and follow-up observations for the planets are impossible because of the nature of the gravitational microlensing method. Hence, considering that M dwarfs are the most numerous stars (e.g., Covey et al. 2008), we almost certainly have yet to study the most “common” planetary systems in the Galaxy.

2.2. Planet formation and habitability around late-M dwarfs: the key science questions

M dwarfs are the most ubiquitous in the Galaxy. Dedicated planet surveys around M dwarfs are indispensable in order to generally understand planet formation and habitable planets in the Galaxy. Extrasolar planet exploration around late-M dwarfs provides several scientific merits that are essential to study planet formation and habitability, as well as biosignatures and life beyond Earth.

Search for Earth-mass Habitable Planets: M dwarfs, especially late-M dwarfs, are the most promising stars to search for habitable planets. The low luminosity of an M dwarf brings the HZ closer to the star (e.g. Kopparapu et al. 2013). Even planets with $P \lesssim 10$ days orbiting a late-M dwarf may be habitable. A standard theory of planet formation predicts that low-mass M dwarfs tend to preferentially harbor low-mass planets, and that many Earth-mass planets are likely rich in water (Ogihara & Ida 2009; Tian & Ida 2015). Planetary habitability is strongly related to orogeny, such as continental formation, plate tectonics, and planetary magnetic fields, as well as atmospheres and liquid water on the surface of a planet. Potentially habitable planets with known masses that were discovered by radial-velocity measurements on the ground are more massive than Earth; we know of only 11 potentially habitable planets around nearby stars ($\lesssim 30$ pc), which are well suited for future characterization. Note that almost all known potentially habitable planets around nearby stars are super-Earths. Super-Earths may have thick, stagnant-lid or sluggish tectonics rather than plate tectonics on Earth, which fails to initiate

internal convective motions and influences the lifetime of planetary magnetic fields. Whether super-Earths are geologically habitable is still a question under debate. Therefore, the discovery of Earth-mass planets in the habitable zone will provide some of the best candidates to search for signatures of life beyond Earth. Even a small number of Earth-mass habitable planets discovered by the IRD-SSP survey should open up a new frontier in planetary habitability outside the solar system.

The detection of RV signals as small as $< 2 \text{ ms}^{-1}$ (see Section 4 for IRD’s performance) can unveil planets in the HZ down to Earth in mass around late-M dwarfs. Figure 4 shows the RV variations of host stars induced by HZ planets with masses of 1, 2, 3, 5, and $10 M_{\oplus}$. An Earth-mass HZ planet around a $0.1 M_{\odot}$ star induces RV signals 20 times greater than the Earth induces on the Sun ($\sim 10 \text{ cm s}^{-1}$). For early-M dwarfs, the frequency of habitable planets was observationally estimated to be $\eta_{\oplus} = 0.41^{+0.54}_{-0.13}$ (super-Earths) using HARPS (Bonfils et al. 2013) and $\eta_{\oplus} = 0.16^{+0.17}_{-0.07}$ (Earth-sized planets with $1\text{--}1.5 R_{\oplus}$) using *Kepler* (Dressing & Charbonneau 2015). In contrast, a late M dwarf is usually too faint for RV monitoring and transit photometry at visible wavelengths. As a result, only a few planetary systems with Earth-mass planets have thus far been detected around late M dwarfs, such as Proxima Centauri (Anglada-Escudé et al. 2016), TRAPPIST-1 (Gillon et al. 2017), and Ross 128 (Bonfils et al. 2017). In addition, contamination from RV signals induced by intrinsic stellar activity, i.e., stellar “jitter,” can mimic the RV signals induced by low-mass planets, especially at visible wavelengths. Near-infrared (NIR) spectrographs are less sensitive to stellar jitter, and thus provide enhanced capabilities for detecting HZ Earth-mass planets orbiting late-M dwarfs with $\sim 0.1\text{--}0.2 M_{\odot}$ (see Section 2.3). If our newly-discovered planets transit their host stars, their atmospheric compositions, as well as the presence of clouds and haze, can be probed by multi-band photometric observations (see Section 7.2). Further characterization of terrestrial planets via direct imaging could be accomplished by follow-up observations with thirty-meter class telescopes in the 2020s, and space-based planet imagers in the 2020s and 2030s (see Section 3.4).

Unbiased distributions from terrestrial planets to giant planets beyond the snow line: How giant planets, i.e., planets with mass $\gtrsim 10 M_{\oplus}$, formed is one of the long-standing problems for planet formation theory. The core-accretion model predicts that once a planetary core reaches a critical mass, it triggers rapid gas accretion from the ambient disk and forms a massive gaseous envelope. Ice condensation beyond the snow line enriches the surface density of building blocks for planets, promoting giant planet formation in the outer region of a disk. In the solar system, the gas giants Jupiter and Saturn exist beyond the snow line, as predicted by the core-accretion model, whereas Uranus and Neptune are also beyond the snow line but failed to become gas giants (the so-called failed core of a gas giant). Why don’t Uranus and Neptune have massive envelopes? Why is Saturn less massive than Jupiter even though it has a more massive core? Even the origin of the four giant planets in the solar system still has many unresolved issues.

Giant planet populations outside the solar system should give us hints on the prevalence of planetary systems in which gas giants and ice giants coexist and the diversity in mass and orbital period of giant planets. Although giant planets that migrated inwards via disk-planet interactions were found in the proximity to central stars, they likely experienced post-formation processes such as photo-evaporative mass loss due to stellar X-ray and UV irradiation, which should obscure their formation histories. The demographics of giant planets in the locations where they formed, beyond the snow line, will provide direct and powerful tests of planet formation theory. In particular, the mass range of planets from ~ 10 to $100 M_{\oplus}$ should be closely related to the mass boundary separating ice giants and gas giants, the so-called critical core mass, according to the core-accretion model. Radial-velocity measurements of Sun-like stars over the past decades have unveiled a large population of long-period giant planets (e.g. Cumming et al. 2008). However, extrasolar planet populations outside of the snow line, especially giant planets with masses $\sim 10\text{--}100 M_{\oplus}$, remain poorly understood (see Figure 3) because of the lack of high-precision long-term ($\gtrsim 10$ yrs) RV monitoring observations. Because planets beyond the snow line and smaller than $\lesssim 100 M_{\oplus}$ are detected only by microlensing surveys, they are not amenable to follow-up observations. However, the condensation of water ice in the disk around a late-M dwarf occurs at $\lesssim 0.1 \text{ au}$; the compactness of planetary systems around late-M dwarfs enables us to search for planets beyond the snow line. This means that short-term RV monitoring ($\sim 1 \text{ yr}$) of late-M dwarfs can explore the mass-period distribution of planets from $\sim 10 M_{\oplus}$ to $13 M_{\text{Jup}}$ outside of the snow line, which tells us the information about the initial masses of planetary building blocks (solid material and disk gas) and the critical mass for runaway gas accretion that falls into $\sim 10\text{--}100 M_{\oplus}$. In addition, the mass-period distribution of giant planets beyond the snow line will give us a constraint on planetary migration, because inward migration of giant planet cores before the onset of rapid gas accretion inhibits giant planet formation (see the upper left and right panels in Figure 6 in Section 3).

The presence of giant planets in the outer region of a disk strongly affects planetary accretion in the terrestrial planet-forming region, as seen in the origin of the late heavy bombardment (Gomes et al. 2005) and Mars

formation predicted by the Grand Tack hypothesis (Walsh et al. 2011). Our IRD survey will be sensitive to (habitable) terrestrial planets and to giant planets beyond the snow line, all around the same stars. The mass-period distribution of terrestrial planets (for example, multiple-planet systems in a mean-motion resonance), can retain signatures of the late stages of planet formation, including giant impacts, disk-planet interactions, and planet migration. These unbiased planet distributions will help us better understand the whole picture of planet formation and orbital evolution of planets in a protoplanetary disk.

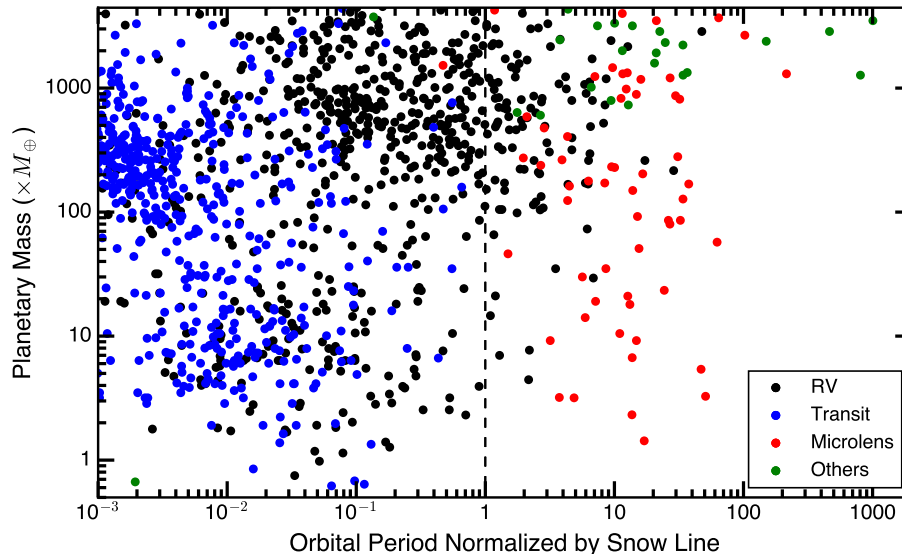


Figure 3. Population of extrasolar planets with measured masses and orbital periods/semi-major axes, taken from the NASA Exoplanet Archive (<https://exoplanetarchive.ipac.caltech.edu>). Their orbital periods are normalized by those at the location of the snow lines of the systems determined by $2.7(L_*/L_\odot)^{1/2}$ au using the (main-sequence) mass-luminosity relations given in Eker et al. (2015) and Cuntz & Wang (2018). Low-mass planets outside the snow line remain largely unexplored in previous and current planet surveys except for some microlensing detections.

Planet population in the radiation environment around late-M dwarfs: The habitability of small planets around low-mass stars depends on factors including volatile inventories, initial atmospheres, and mass loss by stellar X-ray and XUV photons. Late M-dwarfs are especially prolific in their X-ray and XUV flaring, making atmospheric loss and water retention a key question for habitability. In fact, photoevaporative mass loss (photoevaporation: Owen & Wu 2017) and core-powered mass loss (Ginzburg et al. 2018) have already been hypothesized to cause the deficit in the occurrence rate of small transiting planets with $1.5\text{--}2 R_\oplus$, as seen for Sun-like stars (e.g. Fulton et al. 2017) and early-M dwarfs (Hirano et al. 2018), albeit the survey completeness of Earth-sized planets ($1\text{--}1.5 R_\oplus$) is low. The planets discovered by our IRD survey will cover a broad mass range from $\sim 1 M_\oplus$ to $13 M_{\text{Jup}}$ (see Figure 6 in Section 3), and will reveal the effects of stellar activity on the mass distribution of small planets around M dwarfs. Although the late-M dwarfs that the IRD survey will observe are currently inactive stars (see Section 5.1), their stellar evolution in the early phase (for $\lesssim 0.1$ Gyrs) should influence both the thermal and atmospheric evolution of small planets. Therefore, these nearby planets detected by IRD will be the best testbeds to understand “habitability” under extreme conditions, and will be well-suited to future characterization with large ground- and space-based observatories.

2.3. NIR RV measurements for late-M dwarfs and the necessity of IRD

Planet searches around low-mass stars require high-precision RV measurements in NIR-bands. The flux distributions of late-M dwarfs peak at NIR wavelengths (Figure 5). These stars are very faint in the visible, where most existing instruments operate. For stars of spectral type later than $\sim M4$, which are the main targets of our IRD-SSP, the NIR RV measurements are improved in precision compared with visible ones because of the large flux difference that outweighs the richness of the spectral features in the visible (Reiners et al. 2010).

RVs measured from NIR spectra also likely offer the advantage that activity-induced *jitter* is reduced relative to measurements in the visible. Magnetic activity plays a role in limiting the effectiveness of RV planet searches because it results in inhomogeneities on the stellar surfaces (spots and plage). Visible spectra are thus modulated by stellar

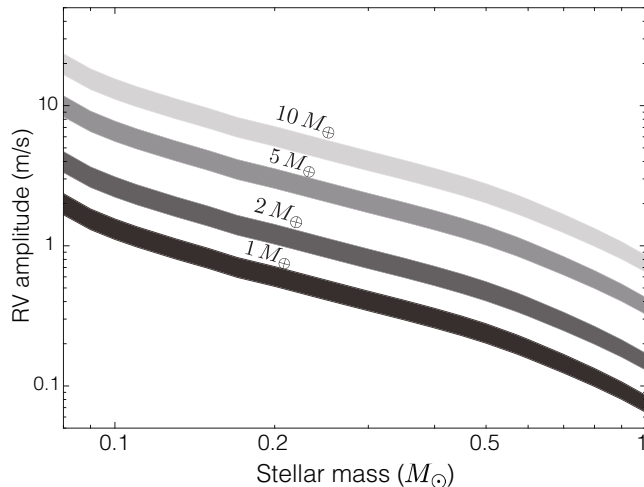


Figure 4. Stellar RV semi-amplitudes induced by a planet with mass of 1, 2, 5, and $10 M_{\oplus}$ in the HZ. The inner and outer edges of the HZ correspond to the moist greenhouse and maximum greenhouse models for one Earth-mass planet with the same atmosphere as that of Earth given in Kopparapu et al. (2013), respectively. The stellar evolution model adopts Baraffe et al. (1998).

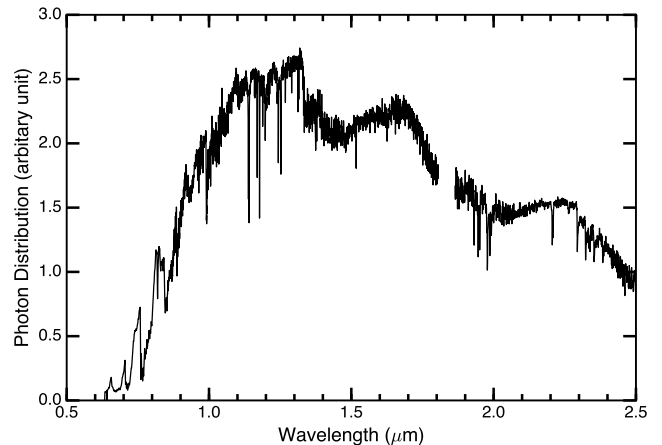


Figure 5. Normalized photon distribution of GJ 406 (M6 dwarf; Kirkpatrick et al. 1991). This figure plots the spectra archived in the DwarfArchives.org[†] for wavelengths $< 0.85 \mu\text{m}$, and the IRTF Spectral Library^{††} for $> 0.85 \mu\text{m}$ (Cushing et al. 2005; Rayner et al. 2009). The range of $\sim 1.81\text{--}1.86 \mu\text{m}$ is masked.

[†]<http://spider.ipac.caltech.edu/staff/davy/ARCHIVE/index.shtml>;

^{††} http://irtfweb.ifa.hawaii.edu/~spex/IRTF_Spectral_Library/index.html

rotation and activity variations, and these intrinsic variations mimic real RV signals and are very confusing for planet searches (e.g., Queloz et al. 2001). It is thought that RVs measured from NIR spectra will exhibit reduced *jitter* relative to the visible because of the smaller contrast, at the longer wavelengths, between the cool inhomogeneities caused by activity and the average stellar surface.

The success of NIR RV measurements, however, has so far been limited due to the absence of suitable instruments and the lack of a suitable spectral reference. Early works in NIR RV measurements that used telluric lines as a wavelength reference achieved a precision of $50\text{--}400 \text{ m s}^{-1}$ for late-M and L dwarfs (e.g. Martín et al. 2006; Zapatero Osorio et al. 2009; Blake et al. 2007; Prato et al. 2008; Blake et al. 2010; Huélamo et al. 2008; Seifahrt & Käufel 2008). Bean et al. (2010) first reached a precision better than 10 m s^{-1} using the CRIRES spectrograph on VLT. They used an ammonia gas absorption cell as a wavelength reference in *K* band demonstrating that precision of $\sim 3\text{--}5 \text{ m s}^{-1}$ is obtainable over timescales shorter than 6 months, and conducted a planet search targeting 30 southern mid- to late-M dwarfs using CRIRES. However, it is necessary to improve the RV precision to detect Earth-like planets in the HZs of late M dwarfs.

3. GOALS OF THE IRD-SSP SURVEY

In the previous sections, it has been emphasized that high precision NIR RV searches around late-M dwarfs with Subaru/IRD will play a critical role in the discovery of planets with $\sim 1\text{--}3 M_{\oplus}$ in the HZ. The detection of such small planets is still too difficult around earlier type stars, including Sun-like stars. Planet exploration for late-M dwarfs is also critical for understanding the formation of Earth-like planets and their “habitability” under the heavy flaring of a young M dwarf. Here, we will describe the science goals of the IRD survey more specifically and compare the IRD survey with other similar surveys.

3.1. Survey goals

Based on the scientific motivations described above, we propose a five-year RV survey for planets orbiting late M-type dwarfs using Subaru/IRD as a new SSP project. This survey will provide unprecedented depth in the exploration of low-mass extrasolar planets orbiting late-M dwarfs, leading to outstanding discoveries and significant statistics of planetary systems. The main purposes of our IRD-SSP are to:

- Discover Earth-mass planets in the HZ around nearby late-M dwarfs as promising targets for future space- and ground-based atmospheric characterization, and to
- Uncover the planet population from Earth-mass planets to giant planets both within and beyond the snow line around late-M dwarfs in order to understand the whole picture of formation and orbital evolution of planets.

Toward these two purposes, our IRD-SSP survey has the following two goals: to

- Derive the mass distribution of planets down to Earth-mass at $P \lesssim 10$ days, of planets down to super-Earths ($P \lesssim 300$ days), and of giant planets (in particular with 10-100 Earth-masses) beyond the snow line ($P \sim 100$ –150 days) and at large distance ($P \lesssim 1000$ days), and to
- Find Earth-mass planets in the HZ around nearby late-M dwarfs.

The advantages of our IRD RV survey over transit surveys such as *Kepler* are:

- High-precision RV measurements are the only way of determining (minimum) masses of planets reliably apart from transit timing variations of a few multi-planet systems. These masses are necessary to constrain the planets' formation histories and orbital evolution, as well as to infer their bulk compositions.
- Planet searches around late-M dwarfs via Doppler methods are more sensitive to planets with long orbital periods beyond the snow line. Such planets are beyond the reach of both ongoing high-precision RV surveys around solar-type stars (like HARPS) and upcoming transit surveys around M-type stars (like *K2* and *TESS*).

3.2. Validating the survey goals

We have used population synthesis to predict our survey's planet yield, and to verify that our IRD survey can achieve the goals that we outlined above. Since the expected planet yield depends on planet formation models, especially, the role of planetary migration, we have considered two cases of Type I migration, which affects the orbital evolution of planets with $\lesssim 100 M_{\oplus}$: $C = 0.01$ (slow Type I migration) and 0.1 (fast Type I migration), where C means that the expected Type I migration speed of a planet embedded in an isothermal disk is reduced by a factor, C ¹. Although we confirmed that our population synthesis models with $C = 0.01$ – 0.1 are consistent with the cumulative distribution of observed close-in super-Earths around Sun-like stars, mass-period distributions of small planets with $P > 10$ days and giant planets with $\lesssim 100 M_{\oplus}$ beyond a snow line, which remain to be unexplored around Sun-like stars because of observational difficulties as stated in Section 2, are sensitive to the efficiency of planetary migration (see Figure 6). In addition to theoretically predicting planet population around late-M dwarfs, we have also used an observationally-predicted planet population based on planet frequency around early-M dwarfs derived from the *Kepler* survey, albeit late-M dwarfs as the IRD-SSP targets would be more likely to harbor different planetary systems. Thus, we evaluate the total number of requested nights for the five-year IRD-SSP survey and then, scientific impacts of the IRD-SSP survey's yield on exoplanetary science.

3.2.1. Expected planet yield based on theoretically predicted planet population

Our theoretical models of planetary accretion (e.g., [Ida et al. 2013](#)) include giant impacts and gravitational scattering among planets and the effects of planetary migration. We considered two cases of Type I migration as mentioned above, a slow ($C = 0.01$) and fast ($C = 0.1$) Type I migration. We generate Monte Carlo realizations of 60 samples out of 150 mid-to-late M-type stars ($M_{\star} \sim 0.1$ – $0.25 M_{\odot}$ and $T_{\text{eff}} \sim 2700$ – 3400 K) in the IRD target list. Figure 6 shows the mass-period distribution of our expected extrasolar planets with mass $\gtrsim 0.1 M_{\oplus}$. We find that gas giants are rare in orbits around M-type stars. On the other hand, both Earth-mass planets and super-Earths are expected to be commonly detected around our IRD target stars, and most of them are water-rich as well. Many water-rich Earth-sized planets fall into the conventional definition of the HZ.

Based on the simulated planet distribution shown in Figure 6, we calculate the number of planets that the five-year Subaru/IRD survey can detect (see Section 5.2 for more detail). We consider three cases: five years of observations of 60, 80, or 100 sample stars. The number of detectable planets, n_p , around 60 M dwarfs in the IRD target list are shown inside each box in Figure 6. Our calculation predicts that the IRD survey can find 68, 86, and 98 planets including 3–4 habitable planets (1–5 Earth-mass ones) around 60, 80, and 100 sample stars, where we adopt *the HZ in which one Earth-mass planet with the same atmosphere as that of Earth can retain liquid water on the surface*. We confirm that our IRD survey has the potential to detect habitable Earth-mass planets around M-type stars and giant planets beyond the snow line ($P \gtrsim 100$ –150 days). We also find that the detection efficiency of planets orbiting late-M dwarfs by the five-year IRD survey is almost one planet per star regardless of sample selection.

3.2.2. Expected planet yield based on empirical planet population derived from the *Kepler* survey

¹ Type II migration can influence the frequency of long-period gas giants which are more massive than Saturn-mass planets. As seen in Figure 6, gas giant planets with $\gtrsim 100 M_{\oplus}$ are expected to rarely form. Thus, our population synthesis considered the efficiency of only Type I migration.

We have used the *Kepler*-based planet occurrence rate derived by DC15, which is a function of orbital periods and planetary radii. We extrapolate this empirical planet occurrence rate for early-M dwarfs to late-M dwarfs to expect the distribution of planets identified with our IRD survey. Then, we partially followed Cloutier et al. (2018) to populate planets and convert their radii to masses. Cloutier et al. (2018) adopts the empirical relationship of planetary mass and radius provided by Weiss & Marcy (2014). Note that the planet occurrence rates based on the *Kepler* observations for early-M dwarfs are available for planets with radii of $0.5\text{--}4 R_{\oplus}$ and orbital periods of $0.5\text{--}200$ days (DC15). Our calculation predicts that the IRD survey can find 83 ± 12 planets including 7 habitable planets (1-5 Earth-mass ones) around 60 sample stars.

3.2.3. The total number of requested nights and expected scientific outputs

Based on our survey simulations fully described in Section 5.2.3, here we summarize why we request 175 nights on Subaru and what scientific outputs are expected from this survey. We evaluate the requested nights for the IRD-SSP to achieve each science goal/motivation below, based on all of our adopted planet populations. Then, note that, as the number of allocated nights increases, the number of late-M dwarfs observable with sensitivity high enough for our science goals increases, resulting in increased planet detections.

The detection of at least “1” habitable Earth-mass planet (science goal A, see below) requires ~ 100 nights for two of three simulated planet populations. As well as occurrence rates of small planets, \sim the 100-night RV survey can reveal habitats of small planets in the proximity to host stars ($P \lesssim 10$ days) and those with $P \lesssim 10\text{--}100$ days (science goal C), in which habitable planets reside. In order to verify whether late-M dwarfs have planetary systems similar to those around early-M dwarfs unveiled by the *Kepler* survey (i.e., the *Kepler*’s empirical population) (science goal C), ~ 100 -night RV observations are necessary to determine occurrence rates of small planets with $1\text{--}10 M_{\oplus}$ within a snow line ($P \lesssim 100$ days) with 1σ -confidence. The detection of Earth-mass planets around late-M dwarfs that exist in a given bin of orbital periods, e.g. $P \lesssim 10$ days (currently, $8 \rightarrow 18\text{--}32$ ones after the 5yr IRD-SSP) and $P \sim 10\text{--}100$ days, can be ensured with 3σ -confidence by 175-nights observations; increasing telescope times allocated to the IRD-SSP survey from 100 to 175 nights allows us to discover more than 2 habitable Earth-mass planets.

In addition, the IRD-SSP survey needs ~ 120 nights in order to examine giant planet frequency with $\sim 10\text{--}100 M_{\oplus}$ beyond a snow line around M dwarfs as one of diagnostic tools of giant planet formation theory and planetary migration (science goal B).

Thus, based on the available nights for Subaru/SSP, the IRD survey will request 175 nights in total for observing at least 60 late-M dwarfs, which can maximize the IRD-SSP survey’s yield and scientific outputs. The expected scientific impacts of the five-year IRD-SSP using 175 nights on our understanding of planet formation/evolution theory are as follows.

Science goal A: The detection of habitable planets: We predict the number of detections of habitable planets is 1–3 based on the theoretical planet population models. The prediction depends on some of the key factors parameterized in the models; in other words, the survey result would constrain the theories (see below). Meanwhile, the detection yield of habitable planets is predicted to be 7 ± 3 if it is reasonable to extrapolate the empirical occurrence rate of small planets around early-M dwarfs revealed by *Kepler* to late-type ones. If the occurrence rate of small-planets around late-M dwarfs differs significantly from that around early-M stars, we will find more or less habitable planets than the prediction. As in Section 2.1, small planets are more abundant around lower-mass stars than higher-mass ones, allowing us to expect more detections of habitable planets around late-M dwarfs. Even if only one habitable planet is detected, that is extremely important to study extraterrestrial life in future. We quantitatively predict that it is possible to perform follow-up characterizations of a habitable planet detected by IRD using a future instrument on a large-aperture telescope like TMT (see Section 3.4).

Science goal B: Impacts of giant planet distributions on the planet formation/evolution theory: Figure 6 shows that the IRD-SSP survey can harvest giant planets beyond a snow line around M dwarfs, which are currently in a veil around Sun-like stars. Microlensing surveys predict that giant planets are ubiquitous outside a snow line around M dwarfs (Fukui et al. 2015). If the IRD-SSP survey yields a plentiful population of giant planets around M dwarfs as well as microlensing surveys, M dwarfs may initially have either a more massive disk than an empirical disk mass-stellar mass scaling relation predicted (Pascucci et al. 2016) or a high dust-to-gas ratio in a disk, as inferred from the TRAPPIST-1 system. Otherwise, a slow Type I migration may be favorable in a protoplanetary disk around M dwarfs, as demonstrated by our population synthesis models in Figure 6. In contrast, the planet frequency around early-M dwarfs inferred from the *Kepler* suggests the rarity of giant planets ($P \lesssim 100$ days) (Dressing & Charbonneau 2015). In fact, the giant planet frequency is thought of to be positively correlated with stellar mass and metallicity (e.g. Johnson et al. 2010). The IRD-SSP survey will be able to estimate the occurrence rate of giant planets in the proximity of M dwarfs and beyond the snow line (see

Figure 6) and will play a crucial role in resolving the apparent discrepancy between microlensing surveys and the *Kepler* results. In addition, the number of planetary systems in which giant planets with mass of $10\text{--}100 M_{\oplus}$ exist beyond a snow line will increase by a factor of ~ 2 after the five-year IRD-SSP survey, improving our understanding of the mass-period distribution of dozens of newly discovered giant planets will reveal how common long-period giant planets such as Uranus and Neptune are and will give us a clue to how massive planetary cores, i.e., a critical core mass, should be required to trigger rapid gas accretion.

Science goal C: Small planet populations around late-M dwarf systems: Several dozen small planets ($1\text{--}10 M_{\oplus}$) within a snow line are expected to be detected by the IRD-SSP (see Figure 6). Previous planet surveys found that the occurrence rate of small planets increases with decreasing stellar mass (F, G, K, early-M stars), as shown in Mulders et al. (2015) (see also Figure 2). The IRD-SSP can test a correlation between stellar mass and frequency of planets down to Earth in mass around late-M dwarfs, i.e., stars at the low-mass end of the initial mass function. The mass-period distribution of small planets observed by the IRD-SSP will also give some feedback to the theory of terrestrial planet formation. As seen in the recently discovered TRAPPIST-1 system (Gillon et al. 2017) around a late-M dwarf, solid material in a disk around M dwarfs tends to be centrally concentrated (Gaidos 2017). Such a compact and massive disk profile comes from efficient transport of building blocks from outside to an inner region via planetary migration. The whole distribution of small planets within a snow line around late-M dwarfs can examine how radial mass transport occurs in a disk around M dwarfs, namely, the efficiency of planetary migration (see Figure 6). In addition, small planet populations around M dwarfs obtained by the IRD-SSP can explore what determines the location of the innermost planet. Several different mechanisms (e.g., disk truncation by stellar magnetic fields and tidal torques at the stellar corotation radius), which strongly depend on stellar mass, have been proposed to make migrating planets get stalled. Compared to massive stars, fully-convective M dwarfs have stronger magnetic fields. A mass-period distribution of planets around M dwarfs, in particular, small planets which are supposed to undergo Type I migration, can be well suited for testing stopping mechanisms of planets. Currently, there is a lack of small planets discovered around low-mass stars. Therefore, the detection yields of small planets from the IRD-SSP survey will constrain orbital evolution of planets in an inner region during planet formation.

The IRD-SSP will find dozens of small planets with $P \lesssim 10$ days in the strong radiation environment of their host stars (see Figure 6). As stated in Section 2.1, the distribution of small planets is sensitive to post-formation processes. Based on follow-up spectroscopy of *Kepler* planet-hosting stars, Fulton et al. (2017) found a deficit of planet occurrence between $1.5\text{--}2.0 R_{\oplus}$ (known as the “radius gap”), which was confirmed by subsequent works (e.g., Van Eylen et al. 2017; Hirano et al. 2018; Berger et al. 2018). There are some theoretical scenarios to explain this radius gap, including the loss of the planet’s outer envelope due to high-energy photons (photoevaporation: Owen & Wu 2017) or by heat from an efficiently cooling core (Ginzburg et al. 2018). Large planet samples around a variety of stellar types would enable us to distinguish these possibilities (Ginzburg et al. 2018). Low-mass stars, especially late-M dwarfs, have a higher chance to show strong XUV activity, and thus dozens of planetary systems around those late-type stars revealed by the IRD-SSP would become a great testing bench for the post-formation processes. Although the radius gap is a concept revealed by “transit” observations, Doppler mass measurements of small planets by the IRD-SSP would also provide insight into not only the mass boundary between rocky planets and gas dwarfs but also small planet populations around M dwarfs, particularly ones with $T_{\text{eff}} < 3000$ K. These cool planet hosts will remain largely unexplored by current and upcoming transit missions such as *TESS* (see Section 3.3).

3.3. Advantages of IRD compared with other surveys

As described in Section 2.3, previous NIR-RV surveys achieved an RV precision of $\sim 3\text{--}5 \text{ m s}^{-1}$ at best. Although this is high enough to detect Neptune-mass planets around late-M dwarfs, it is not sufficient to detect super-Earths and Earth-mass planets in the HZs around late M stars. Next generation NIR spectrographs such as Subaru/IRD, which can achieve an RV precision of $< 2 \text{ m s}^{-1}$, can overcome the difficulty. Here we describe advantages of IRD in comparison with other similar surveys.

Currently, one intensive optical and IR Doppler survey for M dwarfs is ongoing (CARMENES; Reiniers et al. 2017), and two other projects besides IRD plan to start in the next 1–2 years: SPIRou (Cloutier et al. 2018), and HPF (Wright et al. 2018). Since HPF is still not in full specification including the detector and SPIRou is still under commissioning, it is difficult to compare their performance to one another and to IRD. For reference, the instrument performances expected or achieved and their target stars are listed in Table 1.

Due to the small apertures of the telescopes, both CARMENES and SPIRou are mainly targeting early to mid-M dwarfs, which would result in detecting more massive planets with a similar RV precision (see Figure 7 for the

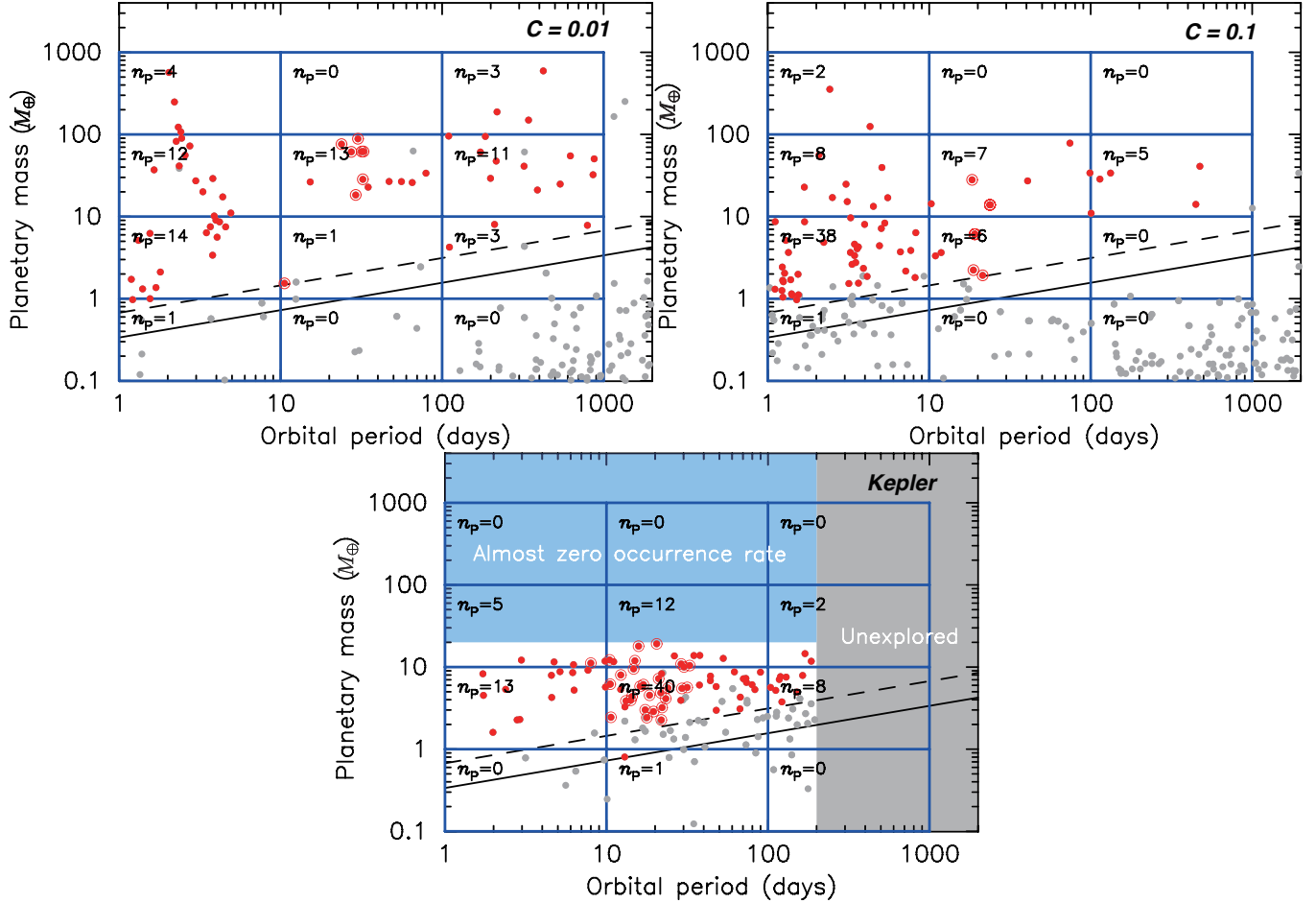


Figure 6. The total numbers of planets (n_p) detected by the five-year IRD survey. Solid and dashed lines indicate the dynamical mass of a planet around a $0.1 M_\odot$ star that causes the stellar RV amplitudes of 1 m s^{-1} and 2 m s^{-1} , which are comparable to the current RV stability of the IRD instrument. Red filled circles represent detectable planets out of simulated ones (non-detectable ones in gray), and habitable ones out of them are drawn by circle. The upper left, right, and bottom panel show planet populations in the cases of a slow ($C = 0.01$) and a fast ($C = 0.1$) Type I migration model, and planet population model based on planet occurrence rates derived from *Kepler* surveys, respectively. The *Kepler*’s empirical population is not applicable to the planetary mass-orbital period ranges that were not investigated by DC15. Note that large giant planets around early-M dwarfs are rare in the *Kepler* survey (e.g., DC15).

comparison of the sampled stars for various surveys). The target catalogs of these surveys include a fraction of M dwarfs later than M5 (e.g., [Reiners et al. 2017](#)), but those have to be nearby bright stars to secure good SNRs. Based on our pre-screening observations for the IRD sample selection, we found that such nearby stars tend to be relatively active stars, revealed by the $H\alpha$ emission, which are not suitable for precision RV measurements. Figure 8 compares the results of equivalent-width (EW) measurements for the late M dwarfs listed on the input catalogs of the IRD (purple) and CARMENES (green) surveys. While the IRD targets involve only quiet late-M dwarfs showing almost no $H\alpha$ emission, the CARMENES samples exhibit a wide variety of $H\alpha$ emissions and more than half of the sampled stars show some sign of stellar activities. Precise parameters and magnitude of activity for the SPIRou sample stars are not known, but our extensive screening observations in the past have revealed that nearby, bright late-M dwarfs are generally more active, with which we ended up with discarding a significant fraction of such bright M dwarfs from the input catalog. 8m-class telescopes like *Subaru* are required in order to uniformly observe quiet, inactive late-M dwarfs, which are best suitable for “blind” Doppler surveys. Regarding the survey with HPF, both IRD and HPF are targeting mainly late-M stars. However, IRD has several major advantages: (1) its wide NIR wavelength coverage including the H-band with the laser frequency comb, (2) smaller telluric absorption at *Subaru*’s higher elevation, (3) flexible telescope pointing, and (4) currently achieved RV precision.

It should also be stressed that the current and upcoming space-based transit surveys (i.e., *K2*, *TESS*) are not so sensitive to the types of stars and orbital parameters which we are focusing on. A majority of M stars that the *Kepler* prime mission and *K2* mission have been targeting are early-to-mid M dwarfs, and indeed almost no planet has been discovered around M dwarfs later than M5 by those missions. The *TESS* mission is going to observe a larger number

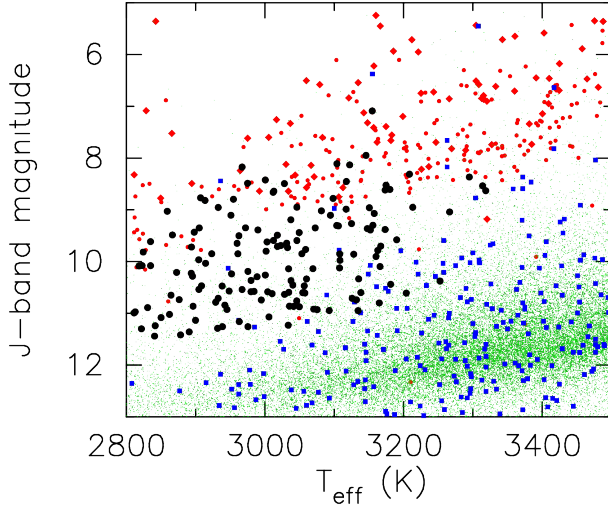


Figure 7. $T_{\text{eff}} - J$ mag distribution for M dwarfs in input catalogs of the past and currently planned exoplanet surveys. Sampled stars in the HARPS, IRD, CARMENES, *TESS* surveys are plotted by red diamonds, black circles, red circles, and green dots (TIC6.2-CTL), respectively. The blue squares are the *TESS* planets expected based on Barclay et al. (2018).

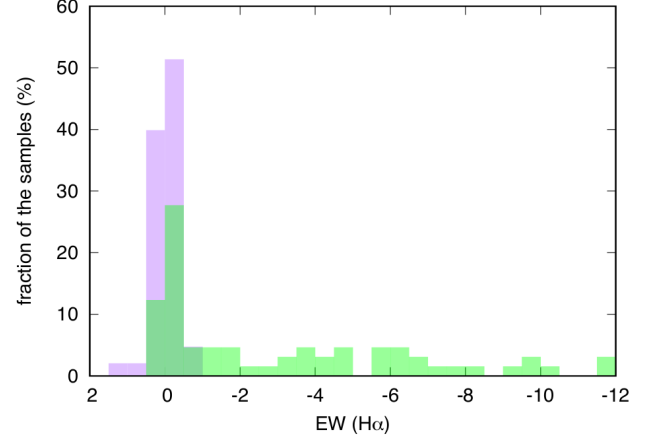


Figure 8. Histograms of target fractions (%) for EW of $H\alpha$ (656.28 nm) in Å. Negative EW values (to the right) indicate an “emission” at the line core; positive or near-zero EW values (to the left) suggest the stars are relatively inactive. The IRD targets (purple) include only quiet late-M dwarfs which were selected by our screening observations, while the CARMENES samples show a variety of activity levels.

of late M dwarfs with a more NIR-sensitive passband (Ricker et al. 2015), but according to the planet yield simulations for its entire survey (Sullivan et al. 2015; Barclay et al. 2018), only a handful of planets are expected to be detected around late-M dwarfs ($T_{\text{eff}} < 3000$ K), most of which are not bright enough for further follow-up observations such as mass measurements (see Figure 7). Moreover, the *K2* and *TESS* missions observe most of the sampled stars for a relatively short period (~ 80 days for *K2* and ~ 30 days for *TESS*) so that those missions basically will not be able to detect a large number of planets beyond the snowline, a region that is of great importance in the context of testing planet formation theories.

Table 1. Major IR Doppler surveys for M dwarfs

Spectrograph	IRD	CARMENES	SPIRou	HPF
Telescope	Subaru 8.2 m	Calar Alto 3.5 m	CFHT 3.6 m	HET 9.2 m
Elevation	4139 m	2168 m	4200 m	2026 m
Wavelength coverage	Y, J, H	Optical, Y, J, H	Y, J, H, K	z, Y, J
Wavelength resolution	70,000	90,000	70,000	50,000
Wavelength standard ^a	LFC	FP, Lamp	FP	LFC
First light	2017.8	2015.10	2018.3	2017.9
Survey start	2019A \sim (5 years)	2016.1 \sim (5 years)	2018B (3 years)	2018Q1 \sim
RV precision	$< 2 \text{ m s}^{-1}$	3–4 ($> M6$) m s^{-1}	1 m s^{-1} (goal)	3 m s^{-1} (1 m s^{-1} goal)
Targets (number of target)	M4–M7 (60)	M0–M7 (324)	M1–M8.5 (100)	M4–M9 (300)

NOTE—

^a “LFC” stands for Laser frequency comb, “FP” for Fabry-Perot calibrator, and “Lamp” for Hollow cathode lamp of UNE.

3.4. Near-future follow-up prospects of IRD planets

Non-transiting planets: Even if our detected planets are non-transiting, further follow-up studies are possible under the specific situations by several methods other than the precise RV measurements with IRD. In the case of a gas giant system whose semi-major axis is farther than 1 AU, the orbit will be partially unveiled by GAIA astrometry at the end-of-mission (2022). GAIA astrometry precision is expected to be about $40 \mu\text{as}$ for stars of GAIA G magnitude ~ 16 (de Bruijne et al. 2014), consistent with the magnitude of TRAPPIST-1 (M8, 12.5 pc), capable to detect the planetary orbit on $M_* = 0.1M_{\odot}$, $M_p = 0.1M_J$, 1 AU. By combining GAIA data with the RV measurements from IRD,

the complete orbit will be revealed including the orbital inclination and the longitude of the ascending node which are unable to be unveiled by a Doppler method only.

For a “very” hot Jupiter, whose semi-major axis is about 0.01 AU, it may be possible to perform spectroscopic observations of a dayside part of the planet irradiated by the central star (e.g., Birkby et al. 2017). Although such short-period giant planets around late-type stars may be exceptionally rare, this kind of observation makes it possible to estimate the orbit of the planet and atmospheric composition simultaneously. According to the Planetary Spectrum Generator (<https://psg.gsfc.nasa.gov/>), assuming a Saturn-mass giant planet of $M_p = 0.3M_J$, $R_p = 0.8R_J$ with an orbital radius of 0.01 AU orbiting the typical IRD-SSP targets ($R_* = 0.15R_\odot$, $T_{\text{eff}} = 3100$ K), the maximum star-planet flux ratio is expected to be 10^{-2} in $4\mu\text{m}$. At these contrasts, we can detect gases such as methane in the atmosphere with an effective SNR > 100 by using an instrument sensitive to thermal-infrared wavelength such as CRIRES ($1 - 5\mu\text{m}$) on VLT. The larger planet size relative to that of the central star will enable these atmospheric studies for planets as small as 0.3 Jupiter radii.

Multiple systems: The IRD-SSP survey will discover many close-in small planets including ones possibly in multiple-planet systems. We expect that follow-up observations of these planets by IRD and/or other large telescopes will find more planets in those systems, which increases the samples of multiple-planet systems around low-mass stars. The orbital architecture of multiple-planet systems can provide us with an important clue to understanding formation of those systems. There are two leading models for the formation of close-in terrestrial planets including super-Earths, namely in-situ accretion and orbital migration. In the in-situ accretion model (e.g., Chiang & Laughlin 2013), planets form from the massive disk in place, while in the orbital migration model (e.g., Snellgrove et al. 2001), planets form in the outer disk and then migrate inward due to the interaction with disk gas. The orbital architecture of multiple systems, especially the distribution of orbital periods (semi-major axes) of planets, can constrain the late stage of planet formation. In a multiple-planet system, if two adjacent planets are in a mean motion resonance, it is likely that they underwent some inward migration process to be captured in the resonance. On the other hand, if not, they may be formed in situ by giant impacts of planetary embryos, which usually leads to a non-resonant orbital configuration. With the observation of the multiple-planet systems around M dwarfs, we can clarify which model is more relevant to planet formation around low-mass stars.

One commonly adopted method to detect such multiple systems is to detect dominant signals in the RV data by calculating the recursive periodograms and identifying the most significant signal at every iteration (e.g., Lovis et al. 2011). In most cases, this method detects the most massive planet in the system. Intensive observations of the multiple-planet system candidates are essential to reveal the complete distribution of the planets including the less massive planets. We will conduct preferential observations of our candidate multiple-planet systems.

3.5. Future direct imaging and spectroscopy of IRD planets with ELTs

Our IRD survey will provide an important target list for future direct imaging and spectroscopy of planets with 30-m class ELTs (Extremely Large Telescopes) such as E-ELT, TMT, and GMT. Figure 9 shows the expected star-to-planet contrast in the J-band for detectable IRD planets and 5-sigma contrast curves for E-ELT/EPICS (Kasper et al. 2010) and for TMT/SEIT (Matsuo et al. 2015; Kawahara et al. 2012). Depending on the employed model, we found that reflected light from a few to 100 M_\oplus planets can be detected via direct imaging, and some of them could be characterized by spectroscopy. We emphasize that such cool, Earth to Jupiter mass planets have never been studied in detail so far. Characterization, especially by spectroscopy, will give us important information of their atmospheres and physical parameters. Furthermore, a combination of a high-contrast instrument and high-resolution spectroscopy will allow us to measure planet radial velocity which will yield information of planet kinematics such as rotation, spin, etc (Kawahara et al. 2014).

4. INSTRUMENT

4.1. Instrument description

The Infrared Doppler (IRD) instrument is a fiber-fed high-resolution near-infrared (NIR) spectrometer for the Subaru telescope covering the Y -, J -, and H -bands simultaneously with a maximum spectral resolution of 70,000. In this section, we give a general description and the expected performance of the instrument, which is now undergoing commissioning at the Subaru telescope. Commissioning at the summit started in March 2017 and first light observations were successfully carried out in August 2017. The laser frequency comb is fully available combined with IRD since January 2018. This section includes an overview of the spectrometer optics, fiber injection system, cryogenic system, scrambler, and laser frequency comb. The general schematic view of the instrument is shown in Figure 10 and the important characteristics of IRD are shown in Table 2.

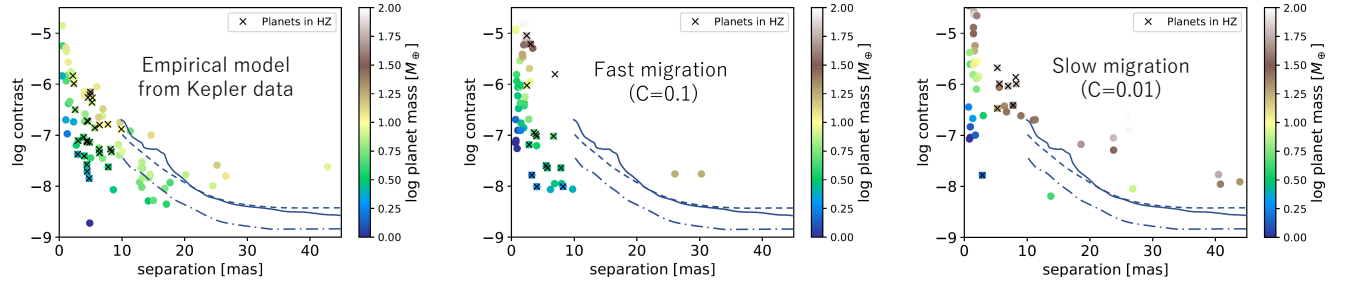


Figure 9. Star-to-planet contrast of detected IRD targets. The 5-sigma detection limits for 10 hour integration time for E-ELT/EPICS-EPOL, 1 hour and 10 hours integration time for TMT/SEIT are also shown by the solid, dashed, dash-dotted lines respectively. From left to right, the planet population synthesis is based on the Kepler data, two theoretical models with the fast and slow migration speed ($C = 0.1$ and 0.01 respectively). The planet albedo 0.3 is assumed to estimate the contrast.

Table 2. Main characteristics of IRD. Please note that the expected RV precision was derived from the laboratory experiment performed for roughly 2 weeks and preliminary on-sky data. The current overall throughput, 2.3%@~1000 nm, will be improved to 3% which is the value assumed in the simulation in this proposal.

Items	Current performance
Spectral resolution	$\sim 70,000$
Spectral coverage	970 – 1750 nm
Fiber field of view	0.48 arcsec
Instrumental RV measurement precision ^a	$< 2.0 \text{ m s}^{-1}$ (with a goal of 1.0 m s^{-1})
Laser frequency spectral coverage	1050 – 1750 nm
Overall throughput	2.3%@~1000nm, 3.3%@~1550nm

^aSee also Table 3.

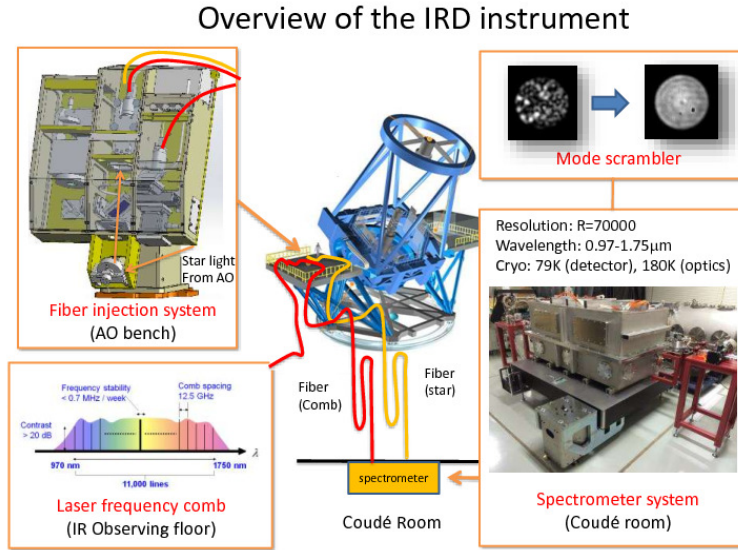


Figure 10. Overview of the instrument

Spectrometer: The spectrometer employs a new high-blazed echelle grating and a VPH cross-disperser to achieve a spectral resolution of 70,000 at maximum, covering the Y, J, H -bands (970 – 1750 nm) simultaneously with two $2.5 \mu\text{m}$ cutoff H2RG detectors. Two fibers can be used simultaneously to record a spectrum of a star and a wavelength reference source, a laser frequency comb. There are no moving parts in the spectrometer optics and the spectrometer is enclosed in a vacuum chamber located in the Coude room of the Subaru Telescope for good temperature stability. The camera lens and detector system are cooled down to 60 K and 79 K respectively; the other parts of the optical system are kept at ~ 180 K by radiative cooling from the radiation shield. Two low-vibration pulse tube cryocoolers

are used for the cooling of the camera-detector system and other optics. The total efficiency from the telescope to the detector and the spectral resolution measured at the summit are shown in Figure 11.

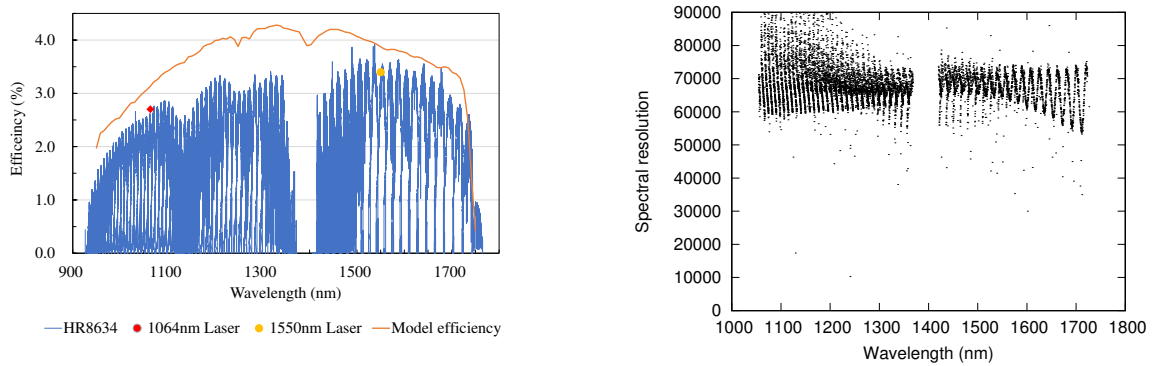


Figure 11. (Left) Total system efficiency from the telescope to the detector, estimated from the on-sky data of the rapid rotator HR 8634 (blue solid lines; [Abt et al. 2002](#)). The red and orange circles show the efficiency estimated from the laser light measurements of the efficiency of the spectrometer, scrambler, fiber and the telescope data from [Jovanovic et al. \(2015\)](#). The solid orange line means the total designed efficiency. (Right) The spectral resolution of the spectrometer measured by using the LFC.

Mode scrambler: We developed several types of mode scramblers ([Ishizuka et al. 2018](#)) to reduce modal noise. This is one of the major error sources in RV measurements, especially in NIR. IRD employs a fiber-based dynamic scrambler by twisting and bending a fiber, combined with a very long fiber (>200 m), which acts as a static scrambler. Scrambling the laser comb light is generally more difficult than scrambling starlight due to the laser’s high coherence. For the laser, we use a rotating ground glass scrambler in addition to our other techniques. We also use a polarization scrambler because the laser comb light is highly polarized and a spectrometer is very sensitive to variations of the polarization state of the input light.

Fiber injection module: The fiber injection module mounted on the Nasmyth platform injects starlight corrected by the adaptive optics system and light from an LFC into two optical fibers. Each fiber scrambles the light from its source and carries it to the spectrometer. A combination of 2 multi-mode fibers (MMF: OFS F8950, 60 μ m, 0’48 core diameter), 2 single-mode fibers (SMF: OFS BF05635-02), or an SMF and an MMF can be used depending on the user’s purpose. Two MMFs will be used as the baseline for the M-dwarf survey to ensure high coupling efficiency to the fiber.

Laser frequency comb: A newly developed laser frequency comb (LFC) covers 1050 – 1750 nm, almost the whole Y, J, H-bands, with a 12.5 GHz frequency interval. This provides the wavelength standard for IRD ([Kashiwagi et al. 2016](#)). Based on a pulse shaping technology using a frequency-stabilized laser, the broadband LFC with a wide mode spacing was directly generated using highly nonlinear fibers. Though the original comb spectrum had peaks and troughs, it was flattened to within about 5 dB across the range using a spectrum shaper consisting of a fiber Bragg grating, fixed optical filters, and a liquid crystal based programmable filter.

4.2. Expected performance

4.2.1. RV Error budget analysis from laboratory experiments

In order to predict IRD’s RV stability, we performed an error budget analysis by splitting the LFC’s light between two fibers. This enabled us to simulate relative RV measurements, which are similar to what we will actually obtain in an observation. The data for this error budget analysis was collected for ~ 2 weeks maximum and we will continue to obtain data to study more long term stability. The results of the error budget analysis are shown in Table 3 and the expected total instrumental error is 1.3–1.9 m s^{-1} . The RV error can be attributed to five main sources: 1) The wavelength stability of the LFC depends on the ambient temperature. The LFC system is enclosed in the temperature control room to ensure a maximum stability 2) Detector intra- and inter-pixel sensitivity variations introduce an apparent change of the Instrumental Profile (IP), when an absolute drift of the spectrum exists. Because the star and the LFC spectrum fall on different detector pixels and the pixel sensitivity varies across the chip, this will lead to a differential spectrum drift which cannot be corrected with the LFC. 3) and 4) Modal noise is one of the major error sources especially for a NIR spectrometer, and it is essentially an actual instability of IP caused by various reasons; telescope pointing variation, thermal deformation of the optics, temperature changes and mechanical deformation of an optical fiber, etc. The modal noise for the LFC and a star light should behave independently because the modal

noise tends to be high in the case of the very high coherent light like LFC, in contrast to partially coherent star light. Therefore we characterized the modal noises each for the LFC and the star light. 5) The spectrometer response, in terms of RV measurements, is very sensitive to the polarization state of the input light. Especially the LFC light is highly linearly polarized and it can cause $>10 \text{ m s}^{-1}$ drift at the worst case from our laboratory experiment. We employ a polarization scrambler to reduce the degree of polarization in a long-exposure case.

4.2.2. Throughput

We measured the current overall throughput of 2.3%@ $\sim 1000 \text{ nm}$ (3.3%@ $\sim 1550 \text{ nm}$) from the on-sky data (Figure 11). By the start of the survey, it will be improved to 3% which is the value assumed in the planet-yield simulation in this proposal (Section 5.2.3).

Table 3. Instrumental error budget analysis

Error source	Error budget (m s^{-1})
LFC stability	0.04
Detector intra & inter pixel sensitivity variation	0.96
Modal noise (LFC)	$\ll 0.7$
Modal noise (star)	0.5–1.2
Polarization noise	<0.76
Total instrumental noise	1.3–1.9 ^a

^aOur planet detection simulations (Section 5.2) conservatively adopt an instrumental RV instability of 2.0 m s^{-1} .

4.2.3. On-sky performance verification

We have undertaken observations of RV standard stars, Tau Ceti (G8V), and the famous M-dwarf Barnard’s star (M4V) to demonstrate the instrument stability on-sky since 2017/August. Note that we had to use a Thorium-Argon lamp (ThAr), a classical wavelength reference often used for the optical spectrometers in 2017, because of a malfunction of the LFC. The number of emission lines from the ThAr lamp is significantly smaller than that of the LFC and it certainly degrades the wavelength calibration. The LFC has been recovered in 2018/January and now it functions properly. Even with the ThAr lamp calibration, the RV error in the data of Tau Ceti and Barnard’s star is ~ 2.5 and 1.5 m s^{-1} over several-nights observations, respectively, which are close to the photon noise limits. We will continue observations of RV standard stars to characterize a long-term stability of the instrument.

5. SURVEY PLAN

In this section, we present our sample selection criteria and what sample stars are actually selected. We then describe the observation plan of the IRD Subaru Strategic Program (IRD-SSP) and how our plan is validated.

5.1. Sample selection

5.1.1. Sample selection criteria

Our RV survey samples a total of 60 mainly-late M dwarfs, which are required to have low masses (low effective temperatures), slow rotational velocities (i.e., slow $v \sin i$), and weak intrinsic surface activity, as well as to not be a double-lined spectroscopic binary. Stars with masses of $< 0.25 M_{\odot}$ are expected to show RV variations greater than 1 m s^{-1} caused by Earth-mass planets in the HZ; those stars are potential targets to detect Earth-mass planets in the HZ. Stellar masses are calculated based on the manner explained in Section 5.1.2. The rotational velocity ($v \sin i$) and effective temperature (T_{eff}) are essential to achieve an RV precision of less than 1 m s^{-1} (Reiners et al. 2010). As shown in Figure 12, we can achieve better RV precision for stars with slower rotations and lower T_{eff} since the RV precision depends on the sharpness and number of absorption lines (also see Section 6.2). M dwarfs with low T_{eff} and slow $v \sin i$ have a large number of sharp molecular absorption lines. We found from Figure 12 that it is important to select stars whose $v \sin i$ is smaller than 5 km s^{-1} to achieve the photon-limit RV precision of $\sim 1 \text{ m s}^{-1}$ with a realistic exposure time.

As is well known, stellar surface activities can disrupt the detection of small (Earth-like) planets, as stellar sunspots, plage, and other surface inhomogeneities can mimic planet-induced RV variations. In the planet search around M dwarfs, the selection of less-active stars, typically represented by low H α emission, provide two crucial advantages as described below.

1. We can mitigate the stellar RV *jitter* induced by activity. Although RV jitters are expected to be smaller in the near infrared, their amplitudes in active M dwarfs are much higher than the expected RV amplitude induced by Earth-mass planets (Reiners et al. 2010).
2. For M dwarfs, stellar activity is known to be strongly correlated with stellar rotation. Low-mass stars with long rotation periods are suitable for detecting planets in the HZ, because slowly-rotating M dwarfs ($v \sin i < 1 \text{ km s}^{-1}$) generally have a rotation period longer than 70 days (e.g., Newton et al. 2016, 2017) (also see Figure 13). For those slow rotators, periodic RV modulations associated with stellar rotation are not identified as planets in the HZ, which is crucial to search for planets in the HZ.

Given these advantages, we again prioritize only slowly-rotating M dwarfs (see Section 5.1.2 about the selection of slow rotators).

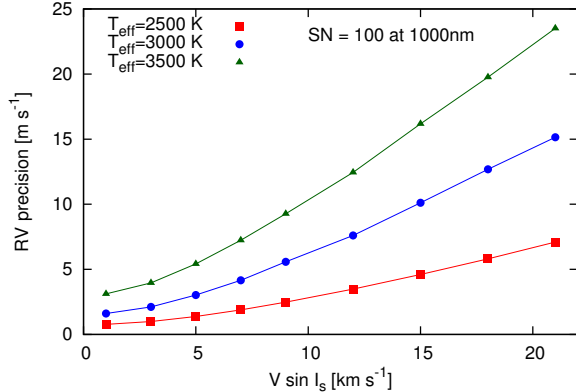


Figure 12. Achievable RV precisions as a function of stellar rotation velocity $v \sin i$. The RV precisions are estimated from simulated IRD spectra on the bases of theoretical NIR spectra of M dwarfs generated by the PHOENIX code (BT-SETTL; Allard et al. 2013) (see also Section 6.2.2). We assumed an SNR of 100 at 1000 nm for each simulated spectrum.

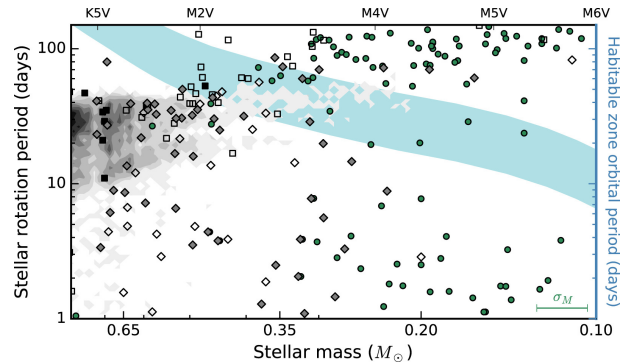


Figure 13. Stellar rotation periods of K–M type stars (Newton et al. 2016). Gray shading and other plots (diamonds, circles and squares) show the rotation periods from *Kepler* and other photometric monitoring observations (see Newton et al. 2016, and the references therein), respectively. The blue shade range indicates the periods comparable to the orbits of the Habitable zone around the stars.

We also exclude double-lined spectroscopic binaries from our sample stars since it is problematic to acquire precise RV measurements for them. High signal-to-noise ratio (SNR) spectra should be secured to achieve an RV precision of $\sim 1 \text{ m s}^{-1}$; SNR's of 100 and 50 are required for M dwarfs with effective temperatures of 3000 K and 2500 K, respectively (see Figure 12). *J*-band magnitudes for which we can achieve an SNR of > 50 with an exposure time of less than 30 minutes (the longest realistic exposure time in our survey) are expected to be $< 11.5 \text{ mag}$, provided that the total efficiency of IRD is $\sim 3\%$. Therefore, in summary, we select our sample stars from stellar catalogs and our pre-selection observations based on the following criteria: 1) stellar mass less than $0.25 M_{\odot}$, 2) rotational velocities ($v \sin i$) less than 5 km s^{-1} , 3) no significant $H\alpha$ emission, 4) measured/expected rotational periods longer than 70 days, and 5) no obvious multiple lines in the high resolution spectra.

5.1.2. Initial target selection

We have made an initial target list of late-M dwarfs (i.e., the IRD input catalog) well before the IRD-SSP exploration actually begins. The input catalog contains 150 late-M dwarfs that are well suited to our goals. We will then carefully select the 60 stars from this list most conducive to precise RV measurements. To achieve our science goals, we must first determine the stellar parameters of this initial target sample. We have therefore collected information from the literature and carried out extensive pre-selection observations.

The stellar-rotation velocity, $v \sin i$, significantly influences the precision of RV measurements. We therefore select M dwarfs that have as low a $v \sin i$ as possible. Because it is difficult to directly measure $v \sin i$ without high-resolution spectroscopy, we have instead constrained the $v \sin i$ of M dwarfs by measuring their $H\alpha$ emission, which is known to be a good tracer of stellar $v \sin i$ (e.g., Newton et al. 2017). Figure 14 shows the rotation velocities and $H\alpha$ -emission equivalent widths of M dwarfs, indicating that an M dwarf with the equivalent width of $H\alpha$ emission more than -0.75 \AA likely has $v \sin i$ slower than 5 km s^{-1} .

We have obtained the equivalent widths of $H\alpha$ emission of M dwarfs from Reid et al. (2003), Lépine et al. (2013), Alonso-Floriano et al. (2015), West et al. (2015) and Newton et al. (2017). Furthermore, we have measured

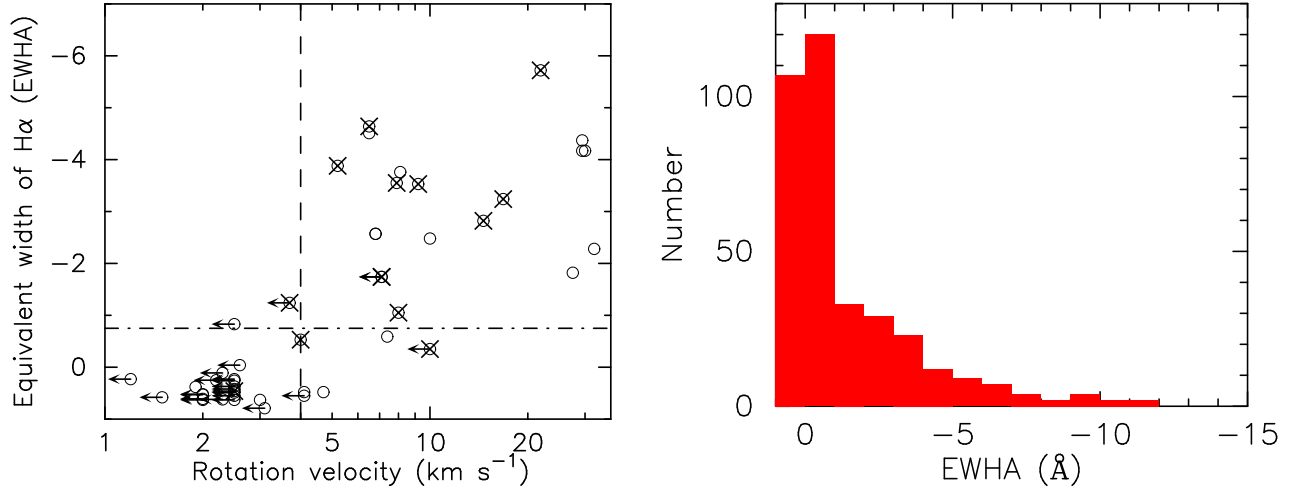


Figure 14. (*left*) Equivalent widths of H α emissions for late-M dwarfs ($0.1\text{--}0.3 M_{\odot}$) with measured rotation velocities. The crosses point out M dwarfs showing strong UV and/or X-ray activities. The rotation velocities of objects in this panel are from Jenkins et al. (2009), Browning et al. (2010), Deshpande et al. (2012), Delfosse et al. (1998b), and Mohanty & Basri (2003). See the main text for the manners and references to obtain the masses and UV/X-ray activities of M dwarfs, as well as the equivalent widths of H α emissions. (*right*) Histogram of equivalent width of H α emission in late M dwarfs considered for the IRD targets. Those with equivalent widths smaller than -0.75 \AA are listed in our full target list.

the H α equivalent widths of ~ 300 M dwarfs using several optical spectrometers: OAO/KOOLS, NHAO/MALLS, WIYN/Bench spectrograph, Calar Alto/CAFOS, and APO/DIS. As a result, we found that ~ 540 M dwarfs among all the M dwarfs that we checked for target selection have weak H α emission likely corresponding to $v \sin i$ slower than $\sim 5 \text{ km s}^{-1}$; the lowest-mass (i.e., late-M), and least active of these help form our initial target list of 150. Note that the exact $v \sin i$ of our target candidates will be derived using the data obtained in the IRD-SSP observations, excluding M dwarfs whose $v \sin i$ are too rapid for precise RV measurements (see below).

We have derived the other key parameters of M dwarfs: spectral type, T_{eff} , metallicity, and mass. The spectral types and T_{eff} are provided by using empirical relationships between $V - J$ color and those parameters (Lépine et al. 2013; Mann et al. 2015). In addition, we utilized the infrared flux method to determine T_{eff} of very-late-M dwarfs. We also obtain independent measurements of T_{eff} and spectral type using our own spectroscopy and literature measurements. We use the empirical formula of absolute K -band magnitude vs. stellar mass (Benedict et al. 2016), with parallaxes from Dittmann et al. (2014), or partly Hipparcos, van Leeuwen (2007). Finally, we determine the metallicities of M dwarfs from their colors using the empirical relationship of Johnson et al. (2012).

As in Section 5.1.1, we only select slowly-rotating M dwarfs, when their rotation periods are available by photometric monitoring. Otherwise, the rotation periods of M dwarfs without photometric rotation periods can be inferred by their H α emission, as described above. High resolution spectroscopy for sample stars taken by any other spectrographs and by IRD is also used for excluding rapidly rotating stars. As well as rotation information, X-ray (Voges et al. 1999, 2000) and UV databases (Bianchi et al. 2011) are used to rule out high-activity stars that produce large RV jitters. Using the Washington Double Star Catalog (Mason et al. 2001, 2014), we also remove close binaries that consist of components with projected separations less than 1000 au.

Figure 15 summarizes the stellar parameters of the M dwarfs in our input catalog. We list M dwarfs with $T_{\text{eff}} < 3400 \text{ K}$, masses $< 0.25 M_{\odot}$, and J -band magnitude < 11.5 . The target selection is further improved by prioritizing M dwarfs with weaker stellar activity. Accordingly, we list the best 150 stars as our IRD-SSP targets. We will vet their RV measurements and $v \sin i$ after observing those with IRD, enabling us to list the late-M dwarfs with slow rotation velocity and remove spectroscopic binaries. We will then establish the 60 targets that we will actually monitor with our IRD-SSP.

5.2. Observation plan to search for Earth-mass planets

In the IRD-SSP, we request 175 nights in total for 5 years for the use of IRD, and would like to be allocated observing runs with 5–6 half nights each in a month, to observe each sample star 80 times. This allocation would provide an expected yield of more than 60 planets and ~ 10 Earth-mass planets out of them based on our planet-yield simulations (see Section 5.2.3).

In our observations, we have two steps to confirm Earth-mass planets around our target stars. The first step is one to three observations for each star to rule out rapidly rotating stars and spectroscopic binaries (SB); see Section 5.2.1. Our second step will consist of intensive RV observations to verify candidate planets and to confirm the orbital

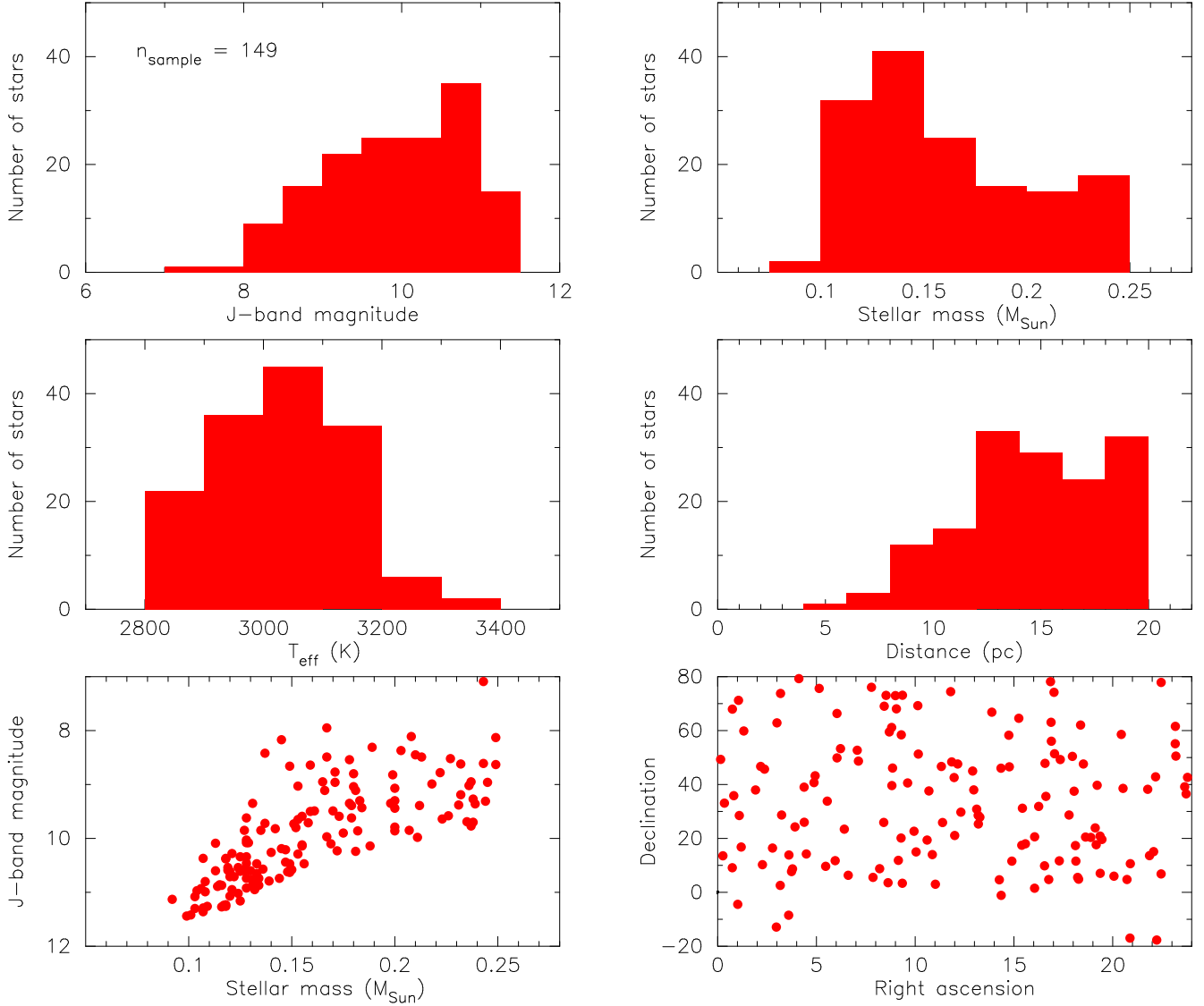


Figure 15. Parameters of our sample stars.

parameters of the planets (Section 5.2.2). We aim to confirm the existence of planets down to one Earth-mass, and also substellar companions up to brown dwarf masses in order to provide a complete picture of these planetary systems.

To decide a strategy suitable for our survey, we performed simulations with various survey plans. The requirements of the survey are summarized as follows: 1) Many Earth-mass planets (in the HZ) should be detected, 2) We should confirm the existence of planets with masses larger than one Earth-mass ($= M_{\oplus}$) and orbital periods less than 10 days, planets with masses of $> 5 M_{\oplus}$ and periods < 300 days, and planets with masses of $> 10 M_{\oplus}$ with periods < 1000 days. Furthermore, we expect to accumulate the 80 RV data points as soon as possible because we would like to reduce systematic variations on the RV surveys caused by intrinsic stellar variability and instrument stability, and to confirm planet candidates promptly. It is useful to be allocated onto mostly half nights as this would allow us to complete the RV survey of a star in two years rather than the four years required for full-night allocations² (Section 5.2.3). We also note that the queue mode observations combined with SExAO/CHARIS can make the observing schedule easier and the observations more efficient. The effective observing strategy for our goals is the plan shown in Table 4.

5.2.1. Screening to select targets suitable for detecting Earth-mass planets

In the first step of our IRD survey, we exclude stars from the target list that are not suitable for our planet search. Rapidly rotating stars, which disturb precise RV measurements due to their broadened absorption lines, are not fully

² Multiple half-night observations can sample the orbital phase of a star more efficiently than multiple full-night ones, avoiding instrument/stellar-activity instabilities and leading to a prompt identification of planets with various orbital periods.

Table 4. Parameters of our plan and the numbers of detectable planets based on theoretical and empirical planet populations

Total observing nights	175 (mainly bright nights)		
Allocation of the nights	5–6 half nights each in a month		
Survey period	5 years		
Observing overhead	480 s		
Instrument RV stability	$\sim 2 \text{ m s}^{-1}$		
RV precision	$\sim 2 \text{ m s}^{-1}$		
<hr/>			
Number of observations for each star	80		
Number of observations in one telescope visit	2		
Number of sample stars we can accumulate 80 RVs	60		
<hr/>			
<i>Number of planets expected in our survey</i>			
Planet population model	Theoretical (C=0.1)	Theoretical (C=0.01)	Empirical (Kepler)
Total number of detectable planets	68±8	64±6	83±12
Planets with >10 M_{\oplus} within 1000 days	23±8	42±5	18±5
Planets with >5 M_{\oplus} within 200 days	29±5	40±5	54±8
Earth-mass planets with orbital periods of < 10 days	21±6	9±3	6±3
Planets in habitable zone	7±4	7±2	21±4
Habitable Earth-like planets	3±1	1–2	7±3

identified by the pre-sample selection described in Section 5.1. Double-lined SBs have composite spectra and make it very difficult to determine a precise RV for the primary target. We exclude double-lined SBs by checking the absorption lines in the observed IRD spectra. First, we observe all of our 150 targets once and then, discard both rapidly rotating stars and double-lined SBs. We further check for the existence of short period (sub)stellar companions using several RV observations in a short observing period for the rest of our sample stars. We then make intensive observations for the targets that we have not excluded through this screening process. Thus, all of our intensively-monitored targets will be inactive single stars on the basis of their $H\alpha$ emission and our screening observations. Using the spectra obtained by the screening observations, we also determine the stellar parameters as described in Sections 6.4, and 7.1.

5.2.2. Survey to search for Earth-mass planets

In the second step of our IRD-SSP, we aim to detect the planetary systems and to measure their properties around inactive late-M dwarf stars. To achieve these goals, we perform intensive and uniform time-series surveys with a large number of RV observations for all samples. In order to detect Earth-mass planets, we need to distinguish small RV variations (1 to several m s^{-1}) caused by the planets from the stellar intrinsic variations. The intrinsic variations are mainly induced by stellar surface modulations associated with the rotational periods of the stars. The highest rate of false positives will therefore be for planets with orbital periods similar to the stellar rotation period. In order to follow the stellar activities, we will monitor stellar variability in our targets by simultaneously monitoring chromospheric lines, and analyzing the shapes of absorption lines to evaluate the effect of stellar spots on measured RVs (see also Section 6.3). Furthermore, an accumulation of many RV measurements can suppress the activity influences. According to simulations of high precision RV surveys of slowly rotating M-dwarfs with weak simulated surface activities (Barnes et al. 2011), it is possible to detect an Earth-mass planet in the HZ around inactive $0.1 M_{\odot}$ M-dwarf stars when 50 RV data are accumulated. Our survey allocates 80 RV observations per star in order to reach our survey goals described above, which is also enough to suppress the weak surface activities.

5.2.3. Simulation of the RV survey

We have simulated the 5-year IRD-SSP program to optimize our survey plan. In the simulations, we investigate the most suitable survey plan taking into account the following factors: the properties of our target stars, instrument conditions, and limitations to our project from the observatory, Subaru telescope. Also, we utilized theoretical simulations as the expected planet population (see also Section 3.2) and estimated the number of planets detected by our survey. In order to calculate the visibility of the sample stars, we took into account two constraints and one request on the use of the Subaru telescope to the IRD-SSP survey: allocated nights of less than 300 in total, survey periods to be less than 5 years, and maximum time allocation of half nights to be 70 per semester in nights with lunar phases of more than 0.5 (bright nights). We considered observation schedules for the RV measurements of all the sample stars under the following instrumental conditions: overhead time (telescope moving and fiber injection through AO) and

exposure times required for sufficient signal-to-noise ratios; for example, a 1200-second exposure time is necessary to achieve $\text{SNR} \sim 100$ for late-M dwarfs with $J = 10$ mag (see Section 5.1)³.

As shown in Table 4, we simulated RVs of all observations with their expected errors and stellar wobble on the RV time series assuming the RV variations induced by planets, stellar rotational modulations of $\sim 1 \text{ m s}^{-1}$, and instrumental instability of $\sim 2 \text{ m s}^{-1}$ (see Figure 16). We varied the number of allocated nights, the number of observations for each star, and the number of observations in one telescope visit to a target. As shown in Figure 16, we can accumulate a sufficient number of observation data efficiently when we are allocated the observing nights onto half night each. We carried out periodogram analyses of the RV data set and calculated the False Alarm Probability (FAP) of the orbital periods of the assumed planets. If the peak in the periodogram has a FAP below 0.01, we consider the planet to be detected. We then estimated detection probability of the planets in our observations with survey parameters described in Table 4. We find that our IRD survey can detect Earth-mass planets with orbital periods of $\lesssim 10$ days around target stars by 80 RV observations (see Figure 17). Thus, our simulations demonstrate that the IRD-SSP is capable of discovering one habitable Earth-like planet, although we note that the actual planet yield from the IRD survey depends on simulated planet populations.

We randomly assigned the planetary systems simulated by the population syntheses to sample stars. We then computed a periodogram analysis of each star with a given observing schedule and a simulated planet, and checked to see whether the false alarm probability of a signal of the planet had fallen below 0.01, as described above. Our simulations indicate that many Earth-mass planets with short orbital periods can be detected around low-mass stars in our samples. In order to estimate the expected number of planets detectable in our survey and their mass-period distribution, we ran a series of Monte Carlo simulations for 60 sample stars 10 times. We evaluated detectability of the planets and their detection rates for a given observing plan in each simulation.

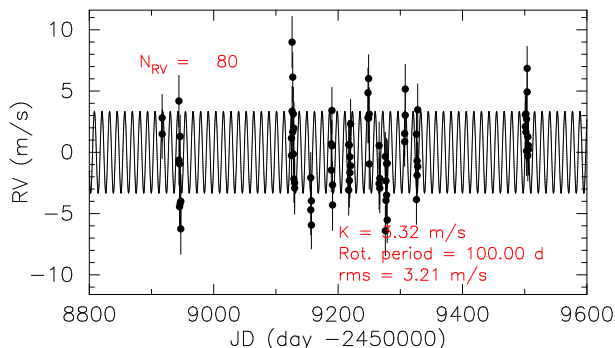


Figure 16. Simulated RV observations of a sample star. Dots indicate simulated RV data points. Solid lines represent orbital curves of a planet with a mass of $3 M_{\oplus}$ and a period of 12.2 days.

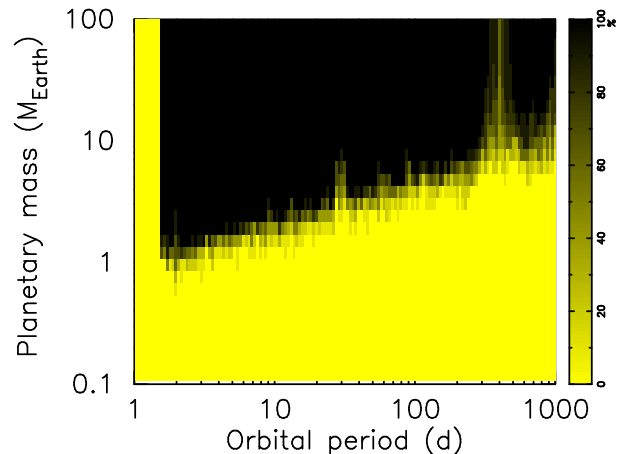


Figure 17. Detection limits of planets around $0.12 M_{\odot}$ star. The color represent detection probability. Black region indicate planets that can be detected in the IRD-SSP survey with 80 RV observations for each star. Planets with <1.5 days are excluded in this simulation.

6. DATA ANALYSIS

6.1. Data reduction

Our data reduction procedure largely follows the standard manner of the HDS (High Dispersion Spectrograph) data reduction. We have developed a pipeline that automatically extracts one-dimensional spectra from two-dimensional images taken with IRD. This software is coded in Python, and uses the PyRAF package that can handle IRAF as one of Python's modules. Since the wavelength range (that is, the echelle format) of IRD is fixed for stabilizing the instrument, it is easy to perform spectral extraction work regardless of the user's experience once we determine the format. Below is a brief description of the processing performed by this tool in order. In IRD, there are two fiber injections to simultaneously record the stellar spectrum and the wavelength calibration source; the following analysis is identically performed for each of them.

³ This estimated integration time includes the effect of fiber-injection degradation related to the Strehl ratio.

Bias Subtraction: First, we remove the bias, the count levels corresponding to zero intensity. The detectors of IRD are two H2RG detectors, each of which has 2048×2048 pixels, arranged in a mosaic pattern. In order to shorten the readout time to several seconds, the acquired image using H2RG is read out by dividing all the pixels into 32 channels along the detector’s column for IRD’s case. Each readout channel has different voltage for its own reference, thus their respective bias levels appear in different counts. This bias pattern is known as “stripe pattern” and is necessary to be subtracted from raw data for each channel. This “destriping” is enabled with a python script in the IRD-reduction pipeline, which works following a similar process as in HiCIAO’s destriping process (Brandt et al. 2013).

Scattered Light: Part of a light beam fed into a spectrograph can be scattered in an unexpected direction due to the effect of dust or scratches on the reflector’s surface. This causes uneven background lighting, which does not account for the true signal. In order to remove this scattered light from each object frame, two dimensional fitting of a polynomial function is applied to inter-order pixels of the echelle spectrum, which are not illuminated by stellar light but by scattered light. For this correction, we call the task “apscatter” of PyRAF.

Flat Fielding: Flat-fielding corrects the pixel-to-pixel sensitivity variations so that each pixel will respond equally to the same photon flux. The lamp flat spectra taken through the two fiber injections are used for this process. The pipeline calls the “apflatten” task for the correction, taking into account the aperture mask, which specifies the echelle order pixels.

1D Spectrum Extraction: After processing these corrections, the obtained two-dimensional frames are ready to be extracted into one-dimensional spectra. The task “apall” is applied for this process.

Wavelength Calibration: The extracted one-dimensional spectra are yielded as a function of pixel-intensity relation. A wavelength calibration is a task to obtain the conversion functions between pixel and wavelength, which are normally fitted as polynomial functions. This pipeline is capable to perform wavelength calibration with the thorium-argon hollow cathode lamp (ThAr lamp) as well as the laser frequency comb. The PyRAF tasks of “ecidentify”, “refspectra” and “dispcor” are executed for this process.

In a general usage, the Th-Ar lamp spectrum is obtained through the fiber injection for objects at start and end of the observation of each night, so the instrumental shift of spectrum can be estimated by linearly interpolating them. After obtaining the polynomial functions fitted for the Th-Ar spectra, the same functions are applied to the stellar spectra but the wavelength baseline shifted along the interpolation (the first calibration, hereafter). At this point, the pipeline provides one-dimensional spectra with satisfactory quality for most of the scientific usage except for precise RV measurements.

As for the case of precise RV measurements, further calibration is needed to attain much better precision. The observations require a reference spectrum, which has been simultaneously obtained for every exposure. Although the stellar spectra involving the first calibration will be obtained as mentioned, the reference spectra provide better calibration precision with the use of the first calibration as the initial state. Details are shown in Section 6.2.

In order to compare the RVs observed when the Earth is at different positions of orbit, the velocity of the Earth compared to the Sun, or the heliocentric velocity, needs to be corrected. As a second correction, the movement of the observer with respect to the center as the Earth rotates should also be corrected. These corrections are calculated with several cm s^{-1} precision by the TEMPO⁴ software which analyses pulsar timing variations.

6.2. RV measurement

In the “simultaneous reference” technique, precise RVs are usually determined in two steps. First, using the wavelength-calibrated 1D stellar spectrum, absolute RVs from the positions of absorption lines are measured. This is often done by cross-correlating the observed spectra with a template spectrum; the world’s most successful optical spectrographs (e.g., ESO/HARPS) adopt box-shaped numerical masks as template spectra (Pepe et al. 2002). Let us denote this measured RV by v_{obs} . The second step is the measurement of spectral “drift” due to variations of temperature and pressure within the spectrograph. The spectral drift of the order $1/1000$ pix translates to $\sim 1\text{--}2 \text{ m s}^{-1}$ in the RV measurement; the correction of this small pixel drift is essential to achieve extreme precision. To assess the spectral drift, the simultaneously observed wavelength-reference spectrum (the laser-comb spectrum in case of IRD) is cross-correlated against a template wavelength-reference spectrum. Assuming that the pixel (wavelength) drift of the stellar spectrum is exactly the same as that of the reference lamp spectrum, the RV drift is estimated by fitting

⁴ J. H. Taylor, R. N. Manchester, D. J. Nice, J. M. Weisberg, A. Irwin, N. Wex and others. Ephemeris routines by E. M. Standish, NASA/JPL.; <http://tempo.sourceforge.net>

the cross-correlation (CC) function by e.g., a single Gaussian and measuring the center v_{drf} of the CC function. The final stellar RV value is then measured as

$$v_{\text{RV}} = v_{\text{obs}} - v_{\text{drf}}. \quad (1)$$

The RV fitting procedure of stellar spectra has several steps: 1) estimation of IP's (for each segment) from the laser-comb spectrum by least-squares deconvolution (LSD; Donati et al. 1997), 2) extraction of the intrinsic stellar spectrum $S(\lambda)$ by deconvolution of IP's and removal of telluric lines, and 3) fitting of each spectral segment. We describe this procedure in detail below.

6.2.1. RV measurement from high resolution spectra

An important difference between the optical and near-infrared (NIR) RV measurements is that NIR spectra suffer a strong telluric absorption even in the “window” bands (Y , J , H). The telluric imprints significantly vary depending on the target's airmass and observing condition (e.g., humidity). In NIR spectroscopy, standard stars have been traditionally observed shortly before or after the scientific exposures to correct for the telluric lines, but it is unrealistic in the Doppler survey to take standard stars for each scientific exposure. Therefore, instead of employing the CC technique used in the optical RV measurement, we attempt to correct for the impact of telluric lines by modeling an observed NIR spectrum $f_{\text{obs}}(\lambda)$ as

$$f_{\text{obs}}(\lambda) = (k_1\lambda + k_0)[S(\lambda(1 - v_{\text{obs}}/c))T(\mathbf{A}; \lambda(1 - v_{\text{tell}}/c))] * \text{IP}, \quad (2)$$

where $S(\lambda)$ and $T(\mathbf{A}; \lambda)$ are the intrinsic stellar spectrum, and telluric absorption spectrum which is characterized by the relevant telluric parameters \mathbf{A} (i.e., precipitable water vapor amount, target airmass), respectively. The factor $(k_1\lambda + k_0)$ indicates the overall normalization, which we express by a linear function here. IP denotes the instrumental profile of the spectrograph (or almost equivalently, point-spread function of the 2D spectrum), which slightly differs for each spectral segment. We split the observed spectrum of each echelle order into small spectral segments (10–20), and the RV fitting is performed for each segment. Assuming that the spectrograph is ultra-stabilized and there is a negligible impact of modal noise, we fix the IP function in Equation (2) at the empirically estimated one⁵.

For each spectral segment (of ~ 1 nm), we fit the observed spectrum $f_{\text{obs}}(\lambda)$ by Equation (2). The basic fitting parameters here are k_0 , k_1 , \mathbf{A} , v_{obs} , and v_{tell} . The additional parameter v_{tell} is introduced so as to account for a possible small variation of telluric line positions, due to e.g., winds. This in turn implies that the telluric line positions are not used in the wavelength calibration and stellar RV measurements. There are in total ~ 1000 segments that the IRD spectrograph can simultaneously cover, from 970 nm to 1600 nm, although a significant fraction (20–60%) of those segments are not usable for RV measurements due to low blaze count level and/or strong telluric absorptions. The final RV value and its uncertainty are determined based on the distribution of v_{obs} measured for each segment and its internal error.

The other components in Equation (2) are described below:

Telluric Absorption Spectrum $T(\mathbf{A}; \lambda)$: For the telluric spectrum, we use theoretical transmission spectra generated by Line By Line Radiative Transfer Model (LBLRTM; Clough et al. 2005). For each grid point in the $\mathbf{A} = (W, A)$ plane, where W is the precipitable water vapor content and A is the target's airmass, we generated a synthetic telluric spectrum on the summit of Mauna Kea. To do so, the $T - p$ (and height H) profile and volume mixing ratio of each atmospheric molecule are required. We employed the averaged profiles based on the Global Data Assimilation System (GDAS)⁶ meteorological archive at the location of Mauna Kea for $H \leq 26$ km, and also downloaded the MIPAS⁷ mid-latitude night-time profiles for $H > 26$ km. We then input these profiles into the LBLRTM code, which generates the telluric transmission spectrum in the NIR with the input variables, (W, A) . In the RV fit (Equation 2), the telluric spectrum $T(\mathbf{A}; \lambda)$ is generated by interpolation from those template telluric spectra on the (W, A) grid, and \mathbf{A} is optimized simultaneously.

Intrinsic Stellar Spectrum $S(\lambda)$: One tricky issue to extract precise RVs by Equation (2) is that the intrinsic stellar spectrum $S(\lambda)$ is difficult to estimate since NIR spectra often suffer from strong telluric absorptions and night airglow. In the NIR, it is also reported that the molecular line lists (both line positions and intensities) are incomplete, and theoretically synthesized spectra often disagree with observed ones. This means that using a theoretical spectrum as $S(\lambda)$ could yield a possible systematic error in v_{obs} when fitted to the observed spectrum.

⁵ This process corresponds to using a fixed box-shaped numerical mask in the optical CC technique.

⁶ <https://ready.arl.noaa.gov/READYcmet.php>

⁷ The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) model atmosphere: <http://www-atm.physics.ox.ac.uk/>

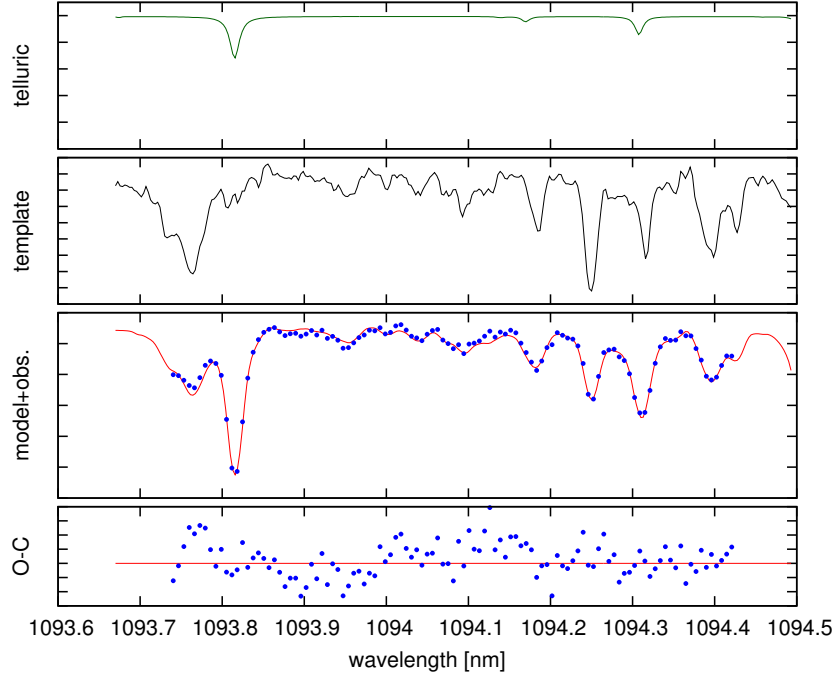


Figure 18. Blue points represent an observed spectrum by IRD for an M dwarf. The red line is its best-fit model using the intrinsic stellar template (black) and telluric absorption spectrum (green).

One solution to this issue is to “deconvolve” the multiple observed spectra taken at different epochs and extract $S(\lambda)$ from the observed spectra; the positions of stellar absorption lines vary ($\sim 30 \text{ km s}^{-1}$) with respect to the telluric absorption lines at each epoch, owing to the observer’s (Earth’s) relative motion along the target’s direction. This fact allows us to decouple the stellar lines from telluric lines by spectral deconvolution and telluric model fitting. For spectral deconvolution, we use the “iterative/recursive” deconvolution described in [Coggins et al. \(1994\)](#). We then apply the telluric model-fitting to the deconvolved spectrum to obtain a telluric-free deconvolved spectrum.

6.2.2. IRD mock data analysis: photon-noise limit

To estimate the RV precision that IRD can achieve, we performed a series of Monte Carlo simulations in the shot-noise limiting case. Although this sort of numerical simulations have been reported in the literature ([Reiners et al. 2010](#); [Rodler et al. 2011](#)), we repeat similar simulations to account for IRD’s specifications such as the wavelength coverage, spectral resolution⁸, and pixel sampling, etc. Below, we briefly describe our Monte Carlo simulation for estimation of IRD’s RV precision for each spectral type. The precision discussed here is the pure “stellar” RV precision, meaning the uncertainty in v_{obs} of Equation (1). The RV error arising from the wavelength calibration (i.e., the uncertainty in v_{drf}) will be discussed in Section 6.2.3.

First, we began with a theoretical NIR spectrum of M dwarfs generated by the PHOENIX code (BT-SETTL; [Allard et al. 2013](#)). We here assumed three types of M dwarfs: early M ($T_{\text{eff}} = 3500 \text{ K}$), mid-M ($T_{\text{eff}} = 3000 \text{ K}$), and late M ($T_{\text{eff}} = 2500 \text{ K}$) dwarfs⁹. Since the RV precision is highly dependent on the stellar rotation, we convolved each PHOENIX spectrum with the rotation plus macroturbulence broadening kernel, following [Hirano et al. \(2011\)](#). We then multiplied the broadened spectrum by the telluric transmission spectrum synthesized by LBLRTM for an arbitrary observing condition on Mauna Kea. To simulate the observed spectrum by IRD, this mock stellar spectrum is multiplied by the IRD throughput including the blaze function for each echelle order, and convolved by the IP for each spectral segment. Finally, the resulting spectral fluxes were converted into the photon counts and stored with the same pixel sampling. The photon-noise was given on each pixel based on the SNR of an arbitrary pixel, for which we adopted 1000 nm. We also took account of the readout noise by randomly adding a white noise with its standard deviation being 14 e^- .

⁸ The spectral resolution R depends on the position of the spectrum. The value of R can vary even within the same echelle order.

⁹ For simplicity, we adopted the surface gravity of $\log g = 5.0$ and metallicity of $[\text{Fe}/\text{H}] = 0.0$.

Figure 18 plots the fitting procedure for a segment of the actual observed IRD spectrum (blue points). From top to bottom, $T(\mathbf{A}; \lambda)$, $S(\lambda)$, $f_{\text{obs}}(\lambda)$, and $O - C$ (residuals) are displayed in the figure. In Figure 12, we plot the achievable RV precisions as a function of stellar rotation velocity $V \sin I_s$. We assumed an SNR of 100 at 1000 nm for each simulated spectrum¹⁰. As expected, rapid rotators exhibit worse RV precisions due to line-broadening which leads to less Doppler information, but later M dwarfs show better RV precisions even though we assumed the same SNR at 1000 nm. This is mainly because late M dwarfs generally have a larger number of molecular lines and richer features in their spectra, providing more intrinsic Doppler information in the spectra. For early M dwarfs, it is almost impossible to achieve $< 2 \text{ m s}^{-1}$ with SNR of ~ 100 , but we can expect to achieve an RV precision of $\sim 1 \text{ m s}^{-1}$ for relatively rapid rotators (up to $v \sin i = 5 \text{ km s}^{-1}$) if we focus on late M dwarfs. This is why our main targets are late M dwarfs in the IRD project: late M dwarfs are generally faint and only 8-m class telescopes are able to achieve this level of SNR with realistic exposure times. Considering the versatility of the telescope and uniqueness of the IRD spectrograph, Subaru/IRD is the best suited among other similar NIR spectrographs to search for planets around late M dwarfs.

6.2.3. Wavelength calibration error

We have described the RV precision extracted from the observed stellar NIR spectra in the shot-noise limiting case above. Another important factor relating to the RV precision is the wavelength calibration error. When we denote the “stellar” RV error by Δv_{obs} and calibration RV error by Δv_{drf} , the resulting RV precision Δv_{RV} in the absence of other systematic effects becomes

$$\Delta v_{\text{RV}} = \sqrt{\Delta v_{\text{obs}}^2 + \Delta v_{\text{drf}}^2}. \quad (3)$$

Since the magnitude of spectral drift differs in each echelle order, v_{drf} in Equation (1) should be evaluated for each order. This means that Δv_{drf} on every order should always be smaller than the best-achievable stellar RV precision Δv_{obs} (for each order), in order for the total RV precision to be dominated by the quality of the stellar spectrum. Based on the stability tests using IRD in the laboratory, we demonstrated that “relative” RV values are stable to $1 - 2 \text{ m s}^{-1}$ (see Section 4.2).

6.3. Mitigating effects of stellar activity in RV analysis

Stellar RV variations are caused by not only orbital motion of planets but also stellar activity, such as rotational modulation of spots or activity cycle of stars. They cause apparent variations in stellar RVs, which can masquerade a planetary signal. Thus to extract tiny RV variations due to low-mass planets, it is necessary to mitigate the effect of stellar activity on the orbital analysis. Infrared observations can be expected to reduce RV variations due to smaller flux contrast compared to optical wavelength. Even in the infrared, however, the spot-induced RV variations may not be fully eliminated. Therefore, we carefully select sample stars for IRD-SSP (see Section 5.1) whose activity is expected to be low enough, but we may still have to deal with RV jitter in order to extract the tiny planetary signal.

A Bayesian framework with Gaussian processes (Dumusque et al. 2017) is one of the widely used techniques for this purpose, which can model stellar activity noise based on correlations between RVs and inputs (e.g., observation times, activity indexes, line shape deformation, brightness variation, etc.). Figure 19 shows the validity of the technique we have developed. We applied the technique to an active solar-type star, and succeeded in reducing the effect of activity-induced RV jitter. This technique is also applicable to the analysis for late-M dwarfs of the IRD-SSP sample in order to further mitigate the effect of RV jitter after we select stars whose activity level is intrinsically low.

6.4. Determination of physical and atmospheric parameters of host stars

The effective temperatures (T_{eff}) are estimated in the sample selection for this program by temperature scales for the $V - J$ color (Section 5.1.2). The infrared flux method using the WISE W1 band is also useful to estimate T_{eff} for which sufficient photometry data are available. Calibration of the method is provided by a set of model atmospheres (e.g., Tsuji 2005).

High-resolution spectra obtained with IRD will enable us to examine the T_{eff} by the measurements of Fe abundances from neutral Fe lines and FeH molecular lines. In the near infrared range, in particular in J -band, clean Fe lines are identified for mid- and late-M dwarfs. These lines are useful to determine the Fe abundances because the line strengths are not very sensitive to assumed T_{eff} . On the other hand, FeH lines also found in this wavelength range are quite sensitive to the temperature. Measurements of FeH lines and neutral Fe lines will examine the consistency of abundance results, which provides an independent constraint on T_{eff} .

¹⁰ We found by repetitive simulations that the RV precision is inversely proportional to the SNR, which is fully consistent with previous works for optical RV measurements in the literature.

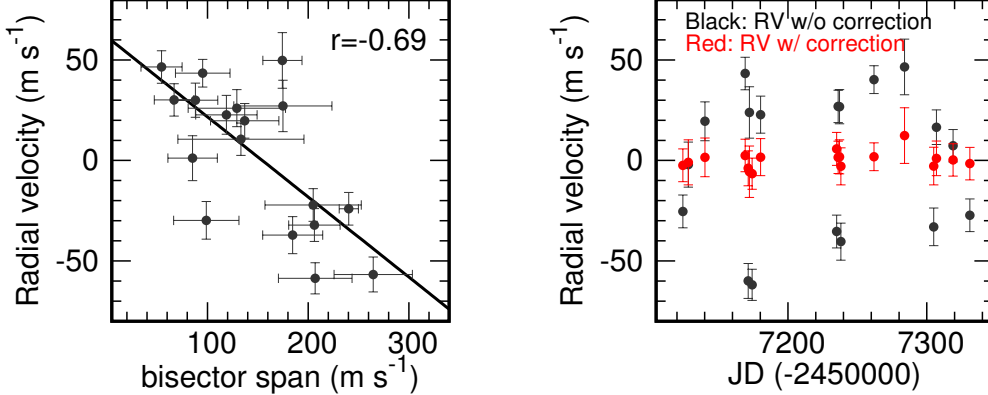


Figure 19. (Left): Bisector span of the CCF profile, which reflects asymmetry of stellar absorption-line profiles, versus RV of an active solar-type star. The solid line is the best linear fit to the data. The correlation coefficient, r , is given in the panel. (Right): RV as a function of JD (-2450000). The filled black circles represent RVs without the correction for stellar activity, and the filled red circles represent RVs with the correction. The RV scatter is reduced by a factor of ~ 5 with the activity correction for the case of this star.

Surface gravity is well estimated for objects for which distance is measured. Hipparcos parallaxes are available only for bright stars, but Gaia DR2 that will soon become available will provide new measurements for fainter ones. Profiles of strong lines in high-resolution IRD spectra might be useful to examine the surface gravity adopted in the analysis. The microturbulent velocity will be determined through the abundance analysis for neutral Fe lines and FeH molecular lines measured from IRD spectra.

The mass, luminosity, and radius of a star are deduced from the T_{eff} and surface gravity using standard models of stellar structure and evolution. When the bolometric flux and the distance from parallax are available, luminosity can be deduced. The radius is also obtained from the T_{eff} and luminosity. Hence, accurate estimates of T_{eff} and gravity (or distance) are key to obtaining physical parameters of host stars. Another key is to determine accurate metallicity and elemental abundance ratios, which will become available by IRD observations (Section 7.1).

7. ANCILLARY SCIENCE

7.1. Chemical abundances of nearby M dwarfs

The determination of chemical abundances of M dwarfs is important on its own even for stars that do not host planets. From the observational point of view, analyzing the spectra in the optical region is not suitable for M dwarfs because of the badly blended molecular features (TiO) along with the inefficiency due to the considerably low flux level. The merit of infrared spectroscopy has been recognized recently.

However, it should be noted that most of the near-infrared spectroscopic studies of M dwarfs are based on medium resolution ($R \sim 2,000$) spectra (Rojas-Ayala et al. 2012; Mann et al. 2013). These studies reveal several characteristic indices corresponding to specific line features, deriving relevant parameters (e.g. T_{eff} or metallicity) with the aid of appropriately calibrated formulae. These formulae are constructed from the infrared spectra of M-dwarf companions to FGK primaries whose abundances are obtained from high resolution optical spectroscopy. What these studies intended to derive is the “metallicity”, while determinations of chemical abundances of individual elements are out of the scope.

Based on the high resolution K band spectra ($R \sim 20,000$) of 42 M dwarfs obtained with IRCs/Subaru, carbon abundances were determined from the ro-vibrational lines of CO 2-0 band, and oxygen abundances were determined from ro-vibrational lines of H_2O (Tsuji & Nakajima 2014; Tsuji et al. 2015). These analyses are effective because in the cool atmosphere of M dwarfs, most carbon atoms are in CO molecules and most oxygen atoms left after the formation of CO are in H_2O molecules. The continuum level of the K band is heavily depressed by numerous H_2O absorption lines and the K band is not suitable for the abundance analysis of atomic species other than carbon and oxygen. Determinations of elemental abundances other than C and O are more promising in Y , J , and H bands where continuum depression by H_2O is absent.

In addition to being an excellent RV machine, IRD is a very good infrared spectrograph with simultaneous spectral coverage of the Y , J , and H bands. For instance, compared with a spectrograph of the previous generation, CRIRES/VLT, the simultaneous spectral coverage is more than one order of magnitude better. This wide spectral coverage combined with the high spectral resolution ($R \sim 70,000$) is a key to the metal abundance analysis of M

dwarfs. Although not as bad as H₂O lines in the K band, there are many FeH absorption lines in the Y , J , and H bands of M dwarfs. The Fe abundance estimated from FeH lines is sensitive to the effective temperature given as an input parameter for the spectral synthesis and FeH is not necessarily a good abundance indicator of Fe. Elemental abundances are better determined from neutral atomic lines that are not affected by FeH blending, which must be searched for with wide spectral coverage and high resolution.

The advantage of IRD over CRIRES/VLT was demonstrated by observing two mid-M dwarfs, Barnard’s star and LHS 1140 during the engineering runs in the summer of 2017 (Ishikawa, H. 2018, Master’s thesis, Sokendai). While from the CRIRES data, numerous FeH lines were primarily analyzed and only the metallicity [Fe/H] was derived (Önehag et al. 2012) because of the limited spectral coverage, the IRD data permitted the determination of abundances of neutral atoms through the equivalent width analysis of absorption lines. The abundances of several individual elements such as Fe, Cr, Ca and Ti were derived in mid-M dwarfs for the first time in the world. A near future plan is to extend the spectral analysis to late-M ($T_{\text{eff}} < 3000$ K) in which neutral atomic lines become less and molecular species become more. In the end, the latest M dwarfs ($T_{\text{eff}} < 2600$ K), will be studied by taking into account the effect of dust formation in the atmosphere.

From a large number of nearby M dwarfs observed through the IRD-SSP, a completely new parameter space of chemical abundances will be opened as a subfield of astronomy. We will certainly be able to investigate the planet-metallicity correlation in the M dwarf regime. To obtain the metallicities of more than 1000 M dwarfs, it is not feasible to observe individual M dwarfs with an 8-m class telescope. The metal abundances obtained by the IRD-SSP can be used to better calibrate the empirical formulae applied to medium resolution spectra obtained by a 4-m class telescope to derive “metallicity”.

One other potential study which will become possible is that of so-called an “M dwarf problem”. The number of low-metallicity M dwarfs may not be sufficient to match simple closed-box models of local Galactic chemical evolution. Metallicity of more than 4000 M dwarfs was estimated from the optical spectra from SDSS using a molecular band strength versus metallicity calibration developed using optical high-resolution spectra of nearby M dwarfs (Woolf & West 2012). High resolution infrared spectra obtained by the IRD-SSP will provide more reliable calibration for band strength-metallicity relation.

7.2. Transit follow-up

According to the IRD survey plan and its planet yield, about one or a few planets among all planets that will be discovered by our survey are expected to be transiting in front of their host star. Those transiting planets, even though the expected number is small, are of great interest as they provide rich information about the nature of the individual planets. First, we can determine the true mass of the RV-detected planets by combining the mass lower-limit (i.e., $m \sin i$) derived by the RV data with the orbital inclination measured by transit photometry. Second, from the transit photometry we can measure the planetary radius, which also gives the mean density of the planet by combining with the true mass, providing us the information about the bulk composition of the planet. Moreover, transiting planets also allow us to investigate the orbital obliquity of the planets (namely, the angle between the stellar spin axis and the planetary orbital axis) and their atmosphere. Thus, finding transiting planets and characterizing them with existing and future telescopes will be important to maximize the scientific impact of our RV survey.

To search for transits of the IRD-SSP planets, we can use MuSCAT (Multi-color Simultaneous Camera for studying Atmospheres of Transiting exoplanets: Narita et al. 2015) on the Okayama 188-cm telescope. MuSCAT is an optical three-band imager, which allows us to obtain high precision light curves of a star in g' , r' , and z_s bands simultaneously. Based on the past successful transit observations with MuSCAT, we expect that the transit of any planet that will be discovered by the IRD-SSP survey can be detected with MuSCAT, as far as they have a transiting orbit. According to the IRD-SSP survey simulation, an Earth-mass planet can be discovered around a late-M dwarf with a mass between 0.1 and 0.2 M_{\odot} , and a brightness of $J < 11.5$ ($z < 13.5$). Assuming a planetary radius of 1 R_{\oplus} , the transit depth of such a planet is expected to be $\sim 0.4\%$ ($\times (0.15 R_{\odot})^2 / R_{\text{star}}^2$), which is within reach of MuSCAT. In fact, we have succeeded in detecting transits of the two Earth-sized planets TRAPPIST-1d and e, the latter of which is located inside the habitable zone (Gillon et al. 2017), with MuSCAT in z_s band (see Figure 20). The host star is a nearby late-M dwarf with a mass of 0.089 M_{\odot} (Van Grootel et al. 2018) and a brightness of $z_s \sim 13.4$ mag (Ahn et al. 2012), which is almost within the range of our survey sample.

The operation of the Okayama 188-cm telescope has been taken over by a researcher group centered on the Exoplanet Observation Research Center of the School of Science of Tokyo Institute of Technology from NAOJ since April 2018. Members of the researcher group can access to the telescope at their own expense. Some of us are members of the group and have access to MuSCAT. Although depending on the availability of research budget, we may pay for a few tens of nights for the 188-cm telescope/MuSCAT per year during 2019-2024 and conduct photometric followup observations of the IRD-SSP planets to investigate if the individual planets have a transiting orbit or not.

Once we find transiting planets, we will investigate their atmosphere by means of transmission spectroscopy (i.e., measuring the transit depth as a function of wavelengths) using various ground- and space-based facilities. For the planets larger than $\sim 2 R_{\oplus}$, we can use the ground-based 6-8 m class telescopes such as 8.2m Subaru and 6.5m TAO. For example, from optical photometric observations with Subaru/FOCAS or HSC, we will be able to observe the Rayleigh scattering feature, from which we can measure the effective atmospheric scale height; assuming a hydrogen-dominated atmosphere, we can distinguish between clear and cloudy (hazy) atmospheres (e.g., [Narita et al. 2013](#)). From near-infrared spectro-photometry using Subaru/MOIRCS or TAO/SWIMS, we can probe molecules such as water and methane. The abundance of these molecules may preserve the information about their origin, i.e., inside or outside the snow line, and therefore provides an important clue for the planetary formation history (e.g., [Brewer et al. 2017](#)). For the planets smaller than $\sim 2 R_{\oplus}$, use of ELTs like TMT or space-based telescopes like HST or JWST is required unless the host star is very small ($R_s \lesssim 0.1 R_{\odot}$) and bright ($J \lesssim 9$).

Needless to say, it would be invaluable if we fortunately find a rocky planet in the habitable zone that is also transiting. We could test for its habitability by investigating the atmospheric composition with the future facilities; some potentially biotic molecules such as water, methane, and even oxygen could be probed in the era of TMT (e.g. [Rodler & López-Morales 2014](#)).

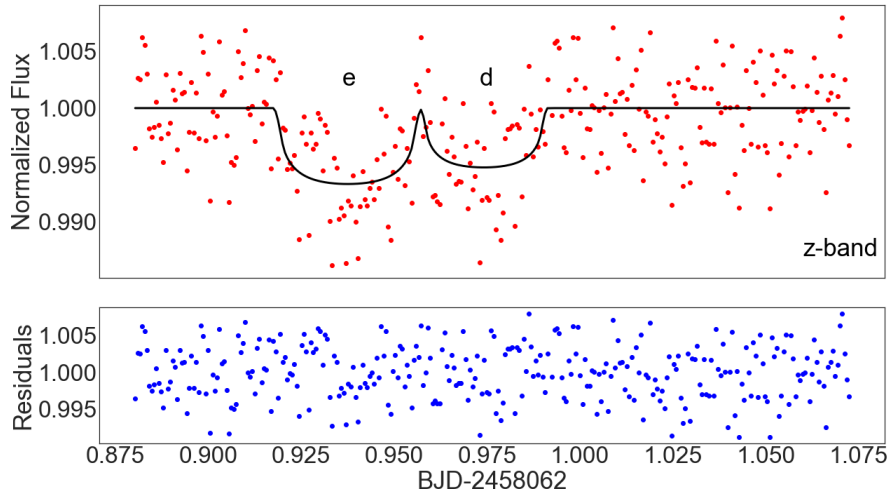


Figure 20. (*Top*) MuSCAT z_s -band light curve of TRAPPIST-1 that captured the consecutively occurred transits of planet e and d. The red points and black line are the observed data and the best-fit transit model, respectively. (*Bottom*) Residuals from the best-fit model.

8. TEAM ORGANIZATION

Our team is an excellent combination of the IRD instrument builders, optical RV extrasolar planet experts, and observatory members. The team also involves observational, instrumental, and theoretical experts of extrasolar planet and planet formation science in Japan and established foreign collaborators being involved with prominent facilities such as Keck, VLT, Gemini, Palomar, AKARI, SST, JWST, and ALMA. The extensive IRD-SSP observations will be conducted with at least 2 resident postdocs and several visiting observers. During appropriate intervals of the observing runs, both spectrometer and comb are semi-regularly maintained by the visiting instrument builders and another resident postdoc. We also have one postdoc dedicated for the comb maintenance. We also have strong data reduction members who are very familiar with both IR array data and RV reduction.

We expect that the experiences we acquired through the previous extrasolar planet SSP survey, SEEDS, will significantly help to perform the IRD-SSP survey. Like all other SSPs, the IRD-SSP is open to the whole Japanese community at any time, and we also consider international collaborations to be critical in order to maximize the scientific outputs from Subaru.

9. DATA ARCHIVE AND RELEASE POLICY

9.1. Data archive

Data archiving is automatically made via SOSS (Subaru Observation Software System) and we will follow the Subaru standard. Raw fits files except for “Fowler data”, and time series of photometric data derived from the Fowler data

will be archived. In addition to this, we will prepare a local data archive system in each Hilo and Mitaka for data retrieval and data analysis for SSP members.

9.2. Data release policy

Data release of raw data files will be done with SMOKA. Since our survey requires a relatively long span to accumulate an enough amount of data, we would like to request to extend the data-release policy from 18 months to 24 months. Reduced RV data and one-dimensional spectrum will be released after the data are published.

10. CONCLUDING REMARKS

IRD is a newly developed near-infrared high-dispersion spectrometer for the Subaru 8.2m telescope equipped with a very stable and wide-band laser frequency comb. IRD enables RV measurement precision of better than 2 m s^{-1} for late-M dwarfs, which allows us to detect small planets, including 1-3 Earth-mass planets in their habitable zone. With a discovery number of at least 60 planets expected from our simulations, the proposed 5-year IRD survey will for the first time explore the population of small rocky planets, including those habitable planets around late-M dwarfs, and for the first time explore both low-mass and giant planets beyond the snow line, which will be critical to testing the planet formation and evolution theories. The proposed science is one of the hottest topics in the extrasolar planet explorations and Subaru/IRD is the most suitable instrument to achieve this goal. The IRD survey will also provide precious targets for future direct explorations of small planets with TMT/E-ELT/GMT and the WFIRST-CGI coronagraph instrument as well as with JWST if they are transiting. Our survey is complementary to the ones targeting earlier M dwarfs such as optical/NIR RV campaigns planned on 8-m class telescopes and the TESS mission, also filling the gap toward further future missions for much more challenging small planet hunting and characterization around solar-type stars such as HabEx (Habitable Exoplanet Imaging Mission) and LUVOIR (Large UV/Optical/IR Surveyor).

REFERENCES

- Abt, H. A., Levato, H., & Grosso, M. 2002, *ApJ*, 573, 359
- Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, *ApJS*, 203, 21
- Allard, F., Homeier, D., Freytag, B., et al. 2013, *Memorie della Societa Astronomica Italiana Supplementi*, 24, 128
- Alonso-Floriano, F. J., Morales, J. C., Caballero, J. A., et al. 2015, *A&A*, 577, A128
- Anglada-Escudé, G., Amado, P. J., Barnes, J., et al. 2016, *Nature*, 536, 437
- Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 1998, *A&A*, 337, 403
- Barclay, T., Pepper, J., & Quintana, E. V. 2018, *ArXiv e-prints*, arXiv:1804.05050
- Barnes, J. R., Jeffers, S. V., & Jones, H. R. A. 2011, *MNRAS*, 412, 1599
- Bean, J. L., Seifahrt, A., Hartman, H., et al. 2010, *ApJ*, 713, 410
- Benedict, G. F., Henry, T. J., Franz, O. G., et al. 2016, *AJ*, 152, 141
- Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L. 2018, *ArXiv e-prints*, arXiv:1805.00231
- Bianchi, L., Herald, J., Efremova, B., et al. 2011, *Ap&SS*, 335, 161
- Birkby, J. L., de Kok, R. J., Brogi, M., Schwarz, H., & Snellen, I. A. G. 2017, *AJ*, 153, 138
- Blake, C. H., Charbonneau, D., & White, R. J. 2010, *ApJ*, 723, 684
- Blake, C. H., Charbonneau, D., White, R. J., Marley, M. S., & Saumon, D. 2007, *ApJ*, 666, 1198
- Bonfils, X., Delfosse, X., Udry, S., et al. 2013, *A&A*, 549, A109
- Bonfils, X., Astudillo-Defru, N., Díaz, R., et al. 2017, *ArXiv e-prints*, arXiv:1711.06177
- Borucki, W. J., Koch, D. G., Basri, G., et al. 2011, *ApJ*, 736, 19
- Brandt, T. D., McElwain, M. W., Turner, E. L., et al. 2013, *ApJ*, 764, 183
- Brewer, J. M., Fischer, D. A., & Madhusudhan, N. 2017, *AJ*, 153, 83
- Browning, M. K., Basri, G., Marcy, G. W., West, A. A., & Zhang, J. 2010, *AJ*, 139, 504
- Butler, R. P., Johnson, J. A., Marcy, G. W., et al. 2006, *PASP*, 118, 1685
- Cassan, A., Kubas, D., Beaulieu, J.-P., et al. 2012, *Nature*, 481, 167
- Chiang, E., & Laughlin, G. 2013, *MNRAS*, 431, 3444
- Clough, S. A., Shephard, M. W., Mlawer, E. J., et al. 2005, *JQSRT*, 91, 233
- Cloutier, R., Artigau, É., Delfosse, X., et al. 2018, *AJ*, 155, 93
- Coggins, J. M., Fullton, L. K., & Carney, B. W. 1994, in *The Restoration of HST Images and Spectra - II*, ed. R. J. Hanisch & R. L. White, 24
- Covey, K. R., Hawley, S. L., Bochanski, J. J., et al. 2008, *AJ*, 136, 1778

- Cumming, A., Butler, R. P., Marcy, G. W., et al. 2008, *PASP*, 120, 531
- Cuntz, M., & Wang, Z. 2018, *Research Notes of the American Astronomical Society*, 2, 19
- Cushing, M. C., Rayner, J. T., & Vacca, W. D. 2005, *ApJ*, 623, 1115
- de Bruijne, J. H. J., Rygl, K. L. J., & Antoja, T. 2014, in *EAS Publications Series*, Vol. 67, EAS Publications Series, 23–29
- Delfosse, X., Forveille, T., Mayor, M., et al. 1998a, *A&A*, 338, L67
- Delfosse, X., Forveille, T., Perrier, C., & Mayor, M. 1998b, *A&A*, 331, 581
- Deshpande, R., Martín, E. L., Montgomery, M. M., et al. 2012, *AJ*, 144, 99
- Dittmann, J. A., Irwin, J. M., Charbonneau, D., & Berta-Thompson, Z. K. 2014, *ApJ*, 784, 156
- Donati, J.-F., Semel, M., Carter, B. D., Rees, D. E., & Collier Cameron, A. 1997, *MNRAS*, 291, 658
- Dressing, C. D., & Charbonneau, D. 2015, *ApJ*, 807, 45
- Dumusque, X., Borsa, F., Damasso, M., et al. 2017, *A&A*, 598, A133
- Eker, Z., Soyduğan, F., Soyduğan, E., et al. 2015, *AJ*, 149, 131
- Endl, M., Cochran, W. D., Kürster, M., et al. 2006, *ApJ*, 649, 436
- Fukui, A., Gould, A., Sumi, T., et al. 2015, *ApJ*, 809, 74
- Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, *AJ*, 154, 109
- Gaidos, E. 2017, *MNRAS*, 470, L1
- Gillon, M., Demory, B.-O., Van Grootel, V., et al. 2017, *Nature Astronomy*, 1, 0056
- Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, *MNRAS*, 476, 759
- Gomes, R., Levison, H. F., Tsiganis, K., & Morbidelli, A. 2005, *Nature*, 435, 466
- Hirano, T., Suto, Y., Winn, J. N., et al. 2011, *ApJ*, 742, 69
- Hirano, T., Dai, F., Gandolfi, D., et al. 2018, *AJ*, 155, 127
- Howard, A. W., Marcy, G. W., Johnson, J. A., et al. 2010, *Science*, 330, 653
- Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, *ApJS*, 201, 15
- Huélamo, N., Figueira, P., Bonfils, X., et al. 2008, *A&A*, 489, L9
- Ida, S., Lin, D. N. C., & Nagasawa, M. 2013, *ApJ*, 775, 42
- Ishizuka, M., Kotani, T., Nishikawa, J., et al. 2018, Accepted for publications to *PASP*
- Jenkins, J. S., Ramsey, L. W., Jones, H. R. A., et al. 2009, *ApJ*, 704, 975
- Johnson, J. A., Aller, K. M., Howard, A. W., & Crepp, J. R. 2010, *PASP*, 122, 905
- Johnson, J. A., Gazak, J. Z., Apps, K., et al. 2012, *AJ*, 143, 111
- Jovanovic, N., Martinache, F., Guyon, O., et al. 2015, *PASP*, 127, 890
- Kashiwagi, K., Kurokawa, T., Okuyama, Y., et al. 2016, *Optics Express*, 24, 8120
- Kasper, M., Beuzit, J.-L., Verinaud, C., et al. 2010, in *Proc. SPIE*, Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III, 77352E–77352E–9
- Kawahara, H., Matsuo, T., Takami, M., et al. 2012, *ApJ*, 758, 13
- Kawahara, H., Murakami, N., Matsuo, T., & Kotani, T. 2014, *ApJS*, 212, 27
- Kirkpatrick, J. D., Henry, T. J., & McCarthy, Jr., D. W. 1991, *ApJS*, 77, 417
- Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, *ApJ*, 765, 131
- Lépine, S., Hilton, E. J., Mann, A. W., et al. 2013, *AJ*, 145, 102
- Lovis, C., Ségransan, D., Mayor, M., et al. 2011, *A&A*, 528, A112
- Mann, A. W., Brewer, J. M., Gaidos, E., Lépine, S., & Hilton, E. J. 2013, *AJ*, 145, 52
- Mann, A. W., Feiden, G. A., Gaidos, E., Boyajian, T., & von Braun, K. 2015, *ApJ*, 804, 64
- Martín, E. L., Guenther, E., Zapatero Osorio, M. R., Bouy, H., & Wainscoat, R. 2006, *ApJL*, 644, L75
- Mason, B. D., Wycoff, G. L., Hartkopf, W. I., Douglass, G. G., & Worley, C. E. 2001, *AJ*, 122, 3466
- . 2014, *VizieR Online Data Catalog*, 1
- Matsuo, T., Murakami, N., Kawahara, H., Kotani, T., & Tamura, M. 2015, *Second Earth Imager for TMT (SEIT) white paper*
- Mayor, M., & Queloz, D. 1995, *Nature*, 378, 355
- Mayor, M., Marmier, M., Lovis, C., et al. 2011, *ArXiv e-prints*, arXiv:1109.2497
- Mohanty, S., & Basri, G. 2003, *ApJ*, 583, 451
- Mulders, G. D., Pascucci, I., & Apai, D. 2015, *ApJ*, 798, 112
- Narita, N., Fukui, A., Ikoma, M., et al. 2013, *ApJ*, 773, 144
- Narita, N., Fukui, A., Kusakabe, N., et al. 2015, *Journal of Astronomical Telescopes, Instruments, and Systems*, 1, 045001
- Newton, E. R., Irwin, J., Charbonneau, D., et al. 2017, *ApJ*, 834, 85
- Newton, E. R., Irwin, J., Charbonneau, D., Berta-Thompson, Z. K., & Dittmann, J. A. 2016, *ApJL*, 821, L19
- Ogihara, M., & Ida, S. 2009, *ApJ*, 699, 824
- Önehag, A., Heiter, U., Gustafsson, B., et al. 2012, *A&A*, 542, A33
- Owen, J. E., & Wu, Y. 2017, *ApJ*, 847, 29
- Pascucci, I., Testi, L., Herczeg, G. J., et al. 2016, *ApJ*, 831, 125
- Pepe, F., Mayor, M., Galland, F., et al. 2002, *A&A*, 388, 632

- Prato, L., Huerta, M., Johns-Krull, C. M., et al. 2008, *ApJL*, 687, L103
- Queloz, D., Henry, G. W., Sivan, J. P., et al. 2001, *A&A*, 379, 279
- Rayner, J. T., Cushing, M. C., & Vacca, W. D. 2009, *ApJS*, 185, 289
- Reid, I. N., Cruz, K. L., Allen, P., et al. 2003, *AJ*, 126, 3007
- Reiners, A., Bean, J. L., Huber, K. F., et al. 2010, *ApJ*, 710, 432
- Reiners, A., Zechmeister, M., Caballero, J. A., et al. 2017, *ArXiv e-prints*, arXiv:1711.06576
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, *Journal of Astronomical Telescopes, Instruments, and Systems*, 1, 014003
- Rodler, F., Del Burgo, C., Witte, S., et al. 2011, *A&A*, 532, A31
- Rodler, F., & López-Morales, M. 2014, *ApJ*, 781, 54
- Rojas-Ayala, B., Covey, K. R., Muirhead, P. S., & Lloyd, J. P. 2012, *ApJ*, 748, 93
- Seifahrt, A., & Käufl, H. U. 2008, *A&A*, 491, 929
- Snellgrove, M. D., Papaloizou, J. C. B., & Nelson, R. P. 2001, *A&A*, 374, 1092
- Sullivan, P. W., Winn, J. N., Berta-Thompson, Z. K., et al. 2015, *ApJ*, 809, 77
- Suzuki, D., Bennett, D. P., Sumi, T., et al. 2016, *ApJ*, 833, 145
- Tian, F., & Ida, S. 2015, *Nature Geoscience*, 8, 177
- Tsuji, T. 2005, *ApJ*, 621, 1033
- Tsuji, T., & Nakajima, T. 2014, *PASJ*, 66, 98
- Tsuji, T., Nakajima, T., & Takeda, Y. 2015, *PASJ*, 67, 26
- Udry, S., & Santos, N. C. 2007, *ARA&A*, 45, 397
- Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2017, *ArXiv e-prints*, arXiv:1710.05398
- Van Grootel, V., Fernandes, C. S., Gillon, M., et al. 2018, *ApJ*, 853, 30
- van Leeuwen, F. 2007, *A&A*, 474, 653
- Voges, W., Aschenbach, B., Boller, T., et al. 1999, *A&A*, 349, 389
- . 2000, *IAUC*, 7432
- Walsh, K. J., Morbidelli, A., Raymond, S. N., O’Brien, D. P., & Mandell, A. M. 2011, *Nature*, 475, 206
- Weiss, L. M., & Marcy, G. W. 2014, *ApJL*, 783, L6
- West, A. A., Weisenburger, K. L., Irwin, J., et al. 2015, *ApJ*, 812, 3
- Woolf, V. M., & West, A. A. 2012, *MNRAS*, 422, 1489
- Wright, J. T., Mahadevan, S., Hearty, F., et al. 2018, in *American Astronomical Society Meeting Abstracts*, Vol. 231, *American Astronomical Society Meeting Abstracts* #231, #246.45
- Zapatero Osorio, M. R., Martín, E. L., del Burgo, C., et al. 2009, *A&A*, 505, L5