# Toward Reliable Planet Occurrence Rates with Kepler

Natalie Batalha
NASA ARC

# Requirements for Reliable $\eta_{earth}$

# Planet Occurrence Rates: Basic Framework

| Observed Distribution | Catalog Reliability | Detection Probability | Intrinsic Distribution |
| --- | --- | --- | --- |

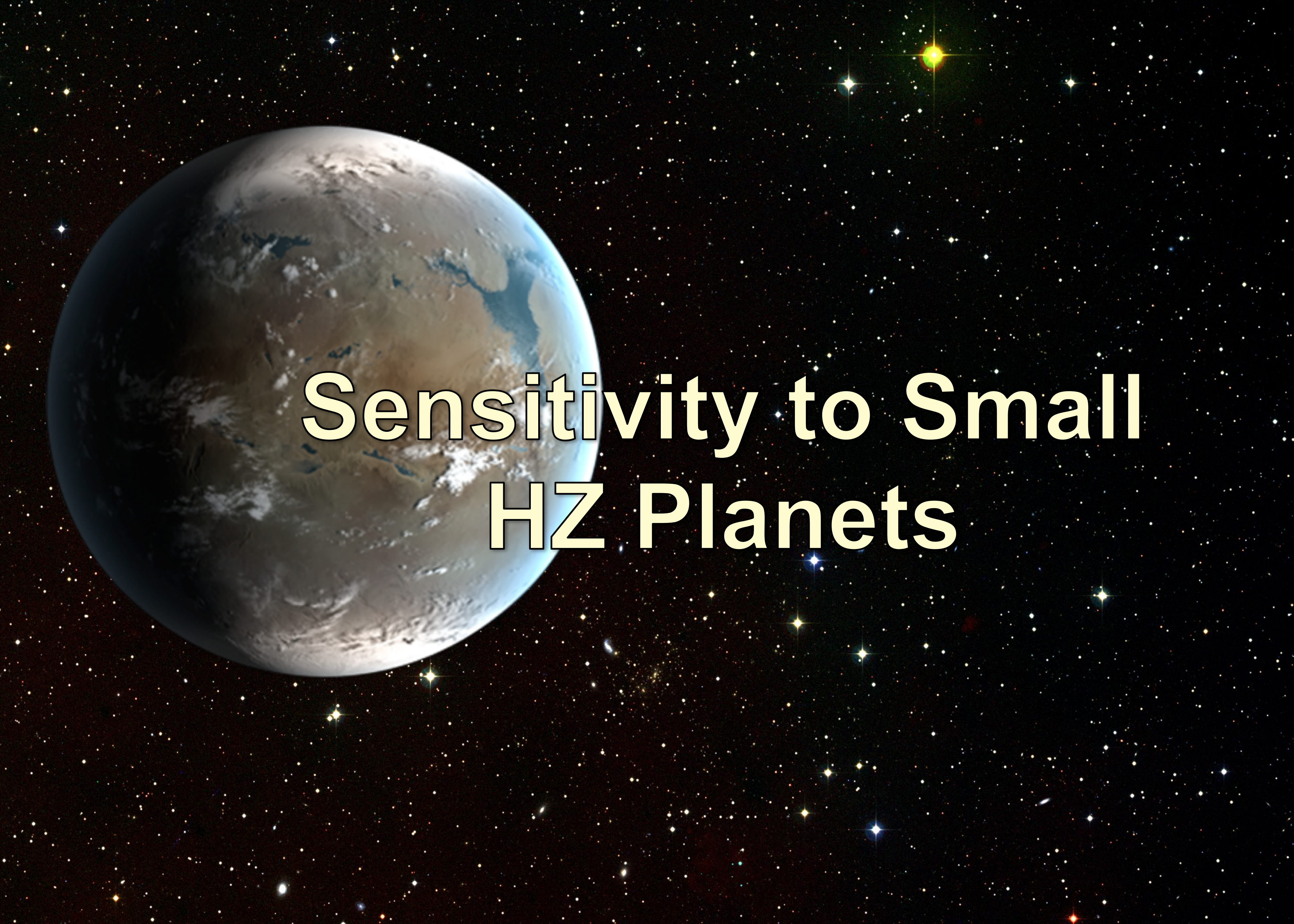
$$O(R_p, P) * R(R_p, P) = P_{tot}(R_p, P) * I(R_p, P)$$

$$P_{tot}(R_p, P) = \sum_{i=1}^{Nstars} P_{i,1}(R_p, P) * P_{i,2}(R_p, P) * P_{i,3}(R_p, P) \ldots$$

Occurrence rates are expressed as a distribution over planet size, period, insolation flux, and/or star type with corrections for both catalog reliability and catalog incompleteness.

# Requirements for a reliable eta-Earth determination

1) Sensitivity to earth-size planets in the HZ of G (K & M) stars.
2) Uniform detection catalog with posterior distributions on the planet properties.
3) Knowledge of Kepler's detection efficiency (completeness)
4) Knowledge of the catalog reliability
5) Well-documented and accessible data products for post-mission analysis by the community.


Sensitivity to Small
HZ Planets

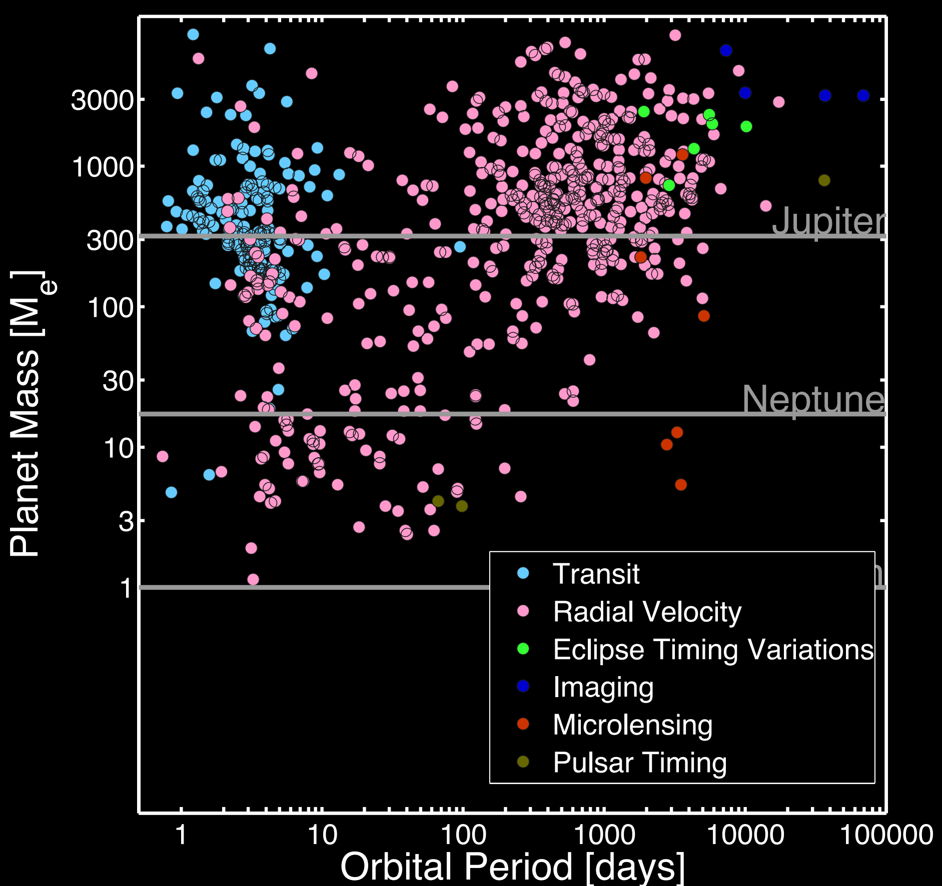
# Catalog Status

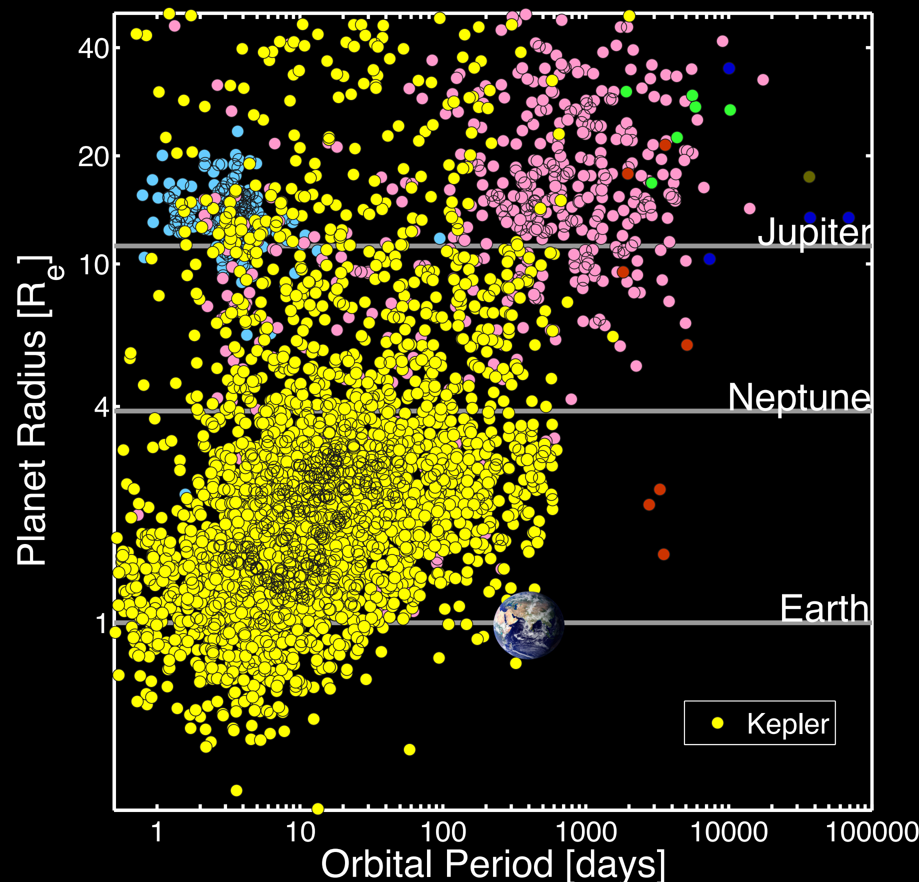
The Kepler prime mission has a three-year close-out period (10/1/2014 to 9/30/2014) to finalize pipeline development, the exoplanet catalog, and occurrence rate products.

| Data | Author | KOI # | Candidates | Status | Date |
|---|---|---|---|---|---|
| Q0 – Q1 | Borucki 1 | 956 | 312 | published | Feb, 2011 |
| Q1 – Q5 | Borucki 2 | 1610 | 1235 | published | July, 2011 |
| Q1 – Q6 | Batalha | 2668 | 2338 | published | Feb, 2012 |
| Q1 – Q8 | Burke | 3149 | 2738 | published | Dec, 2013 |
| Q1 – Q12 | Rowe | 4914 | 3579 | completing | Nov, 2014 |
| Q1 – Q16 | Mullally | 6251 | 4168 | completing | Nov, 2014 |
| Q1 – Q17 | | SOC 9.2: uniform processing | | | 2015 |
| Q1 – Q17 | | SOC 9.3: improved photometric apertures* | | | 2016 |
| Q1 – Q17 | | SOC 10: FINAL CATALOG | | | 2017 |

* See poster by Jeff Smith: significant improvements to SNR reported

# Sensitivity to small HZ planets

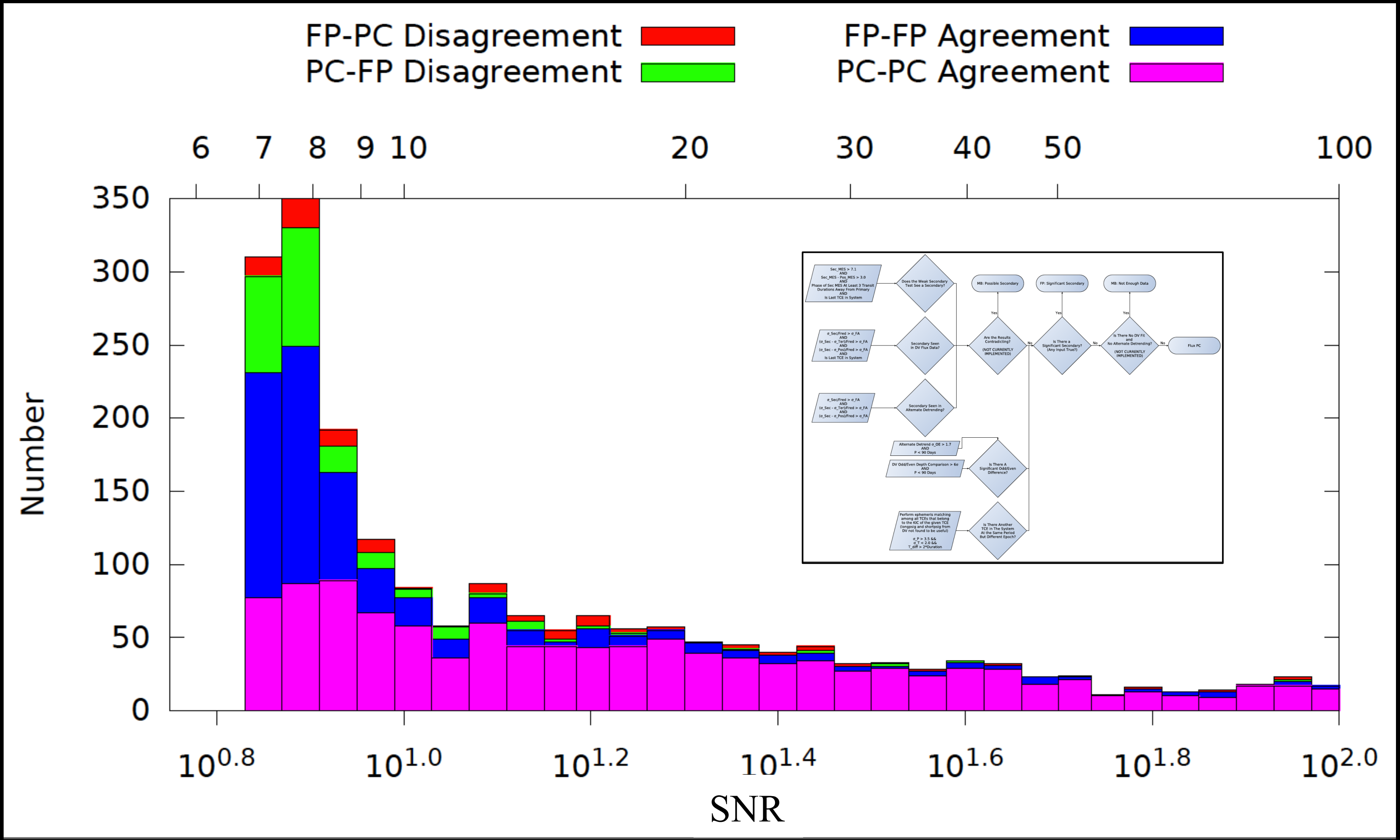



697 non-Kepler planets
90% <u>larger</u> than Neptune
583 unique stars
16% are multi-planet systems

3850 Kepler planet candidates
84% <u>smaller</u> than Neptune
2939 unique stars
22% are multi-planet systems

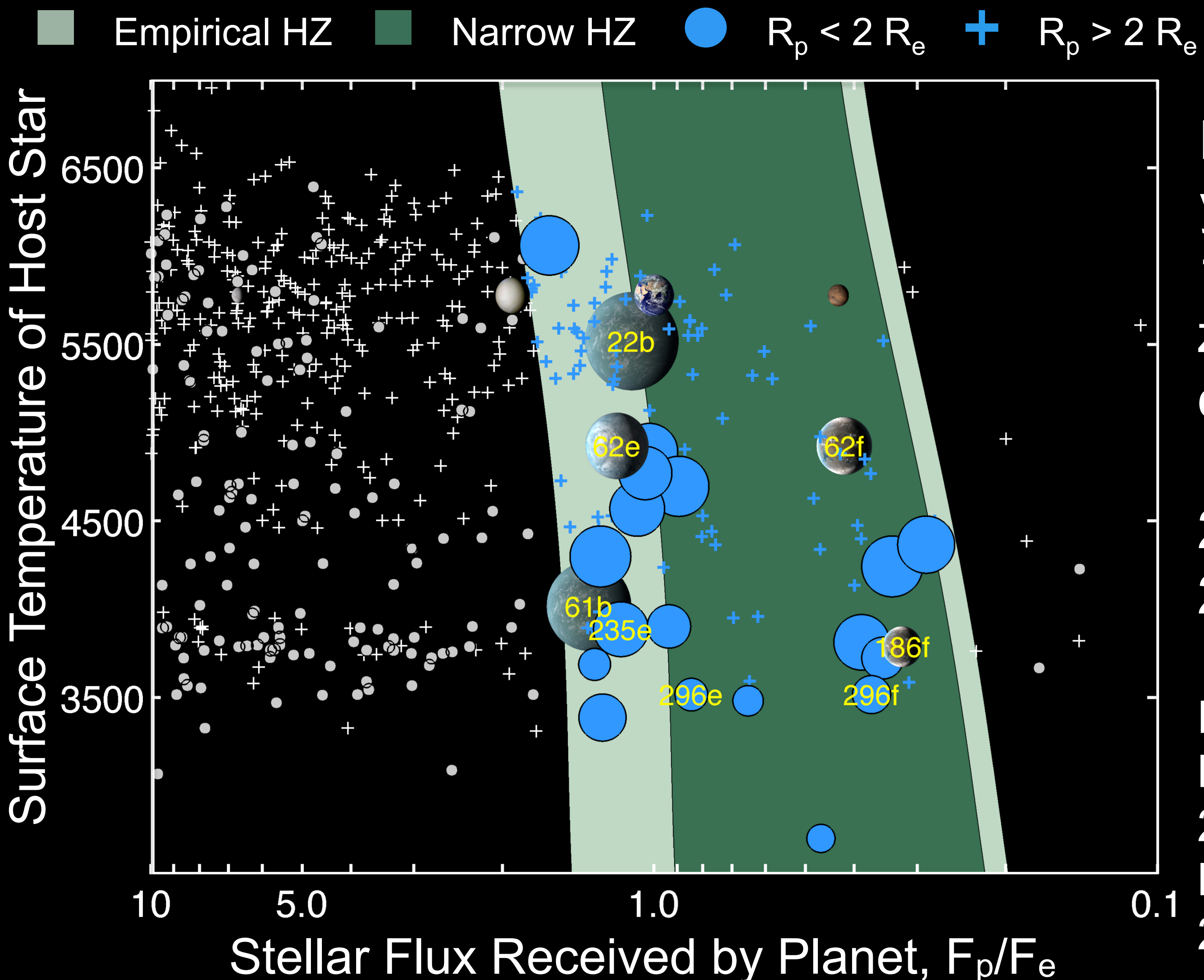
# Automating the Catalog Generation




95.9%, 93.4%, 86.6% agreement with humans at SNR > 20, 10, 7.1

Credit: Jeff Coughlin

# Sensitivity to small HZ planets

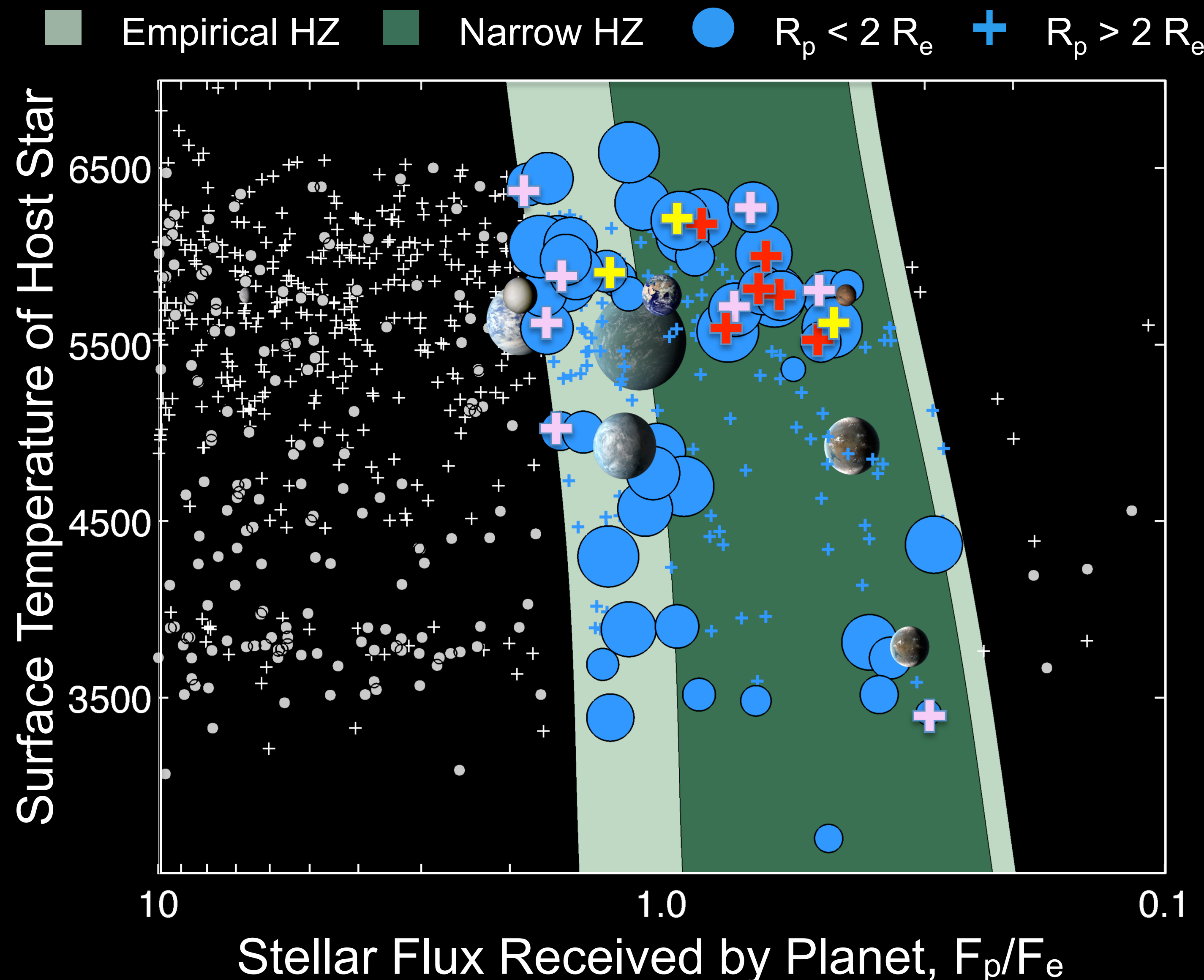



Based on 3 years of data 101 habitable zone candidates

21 smaller than 2 $R_e$.

HZ def'n: Kopparapu et al 2013, updated by Leconte et al 2014

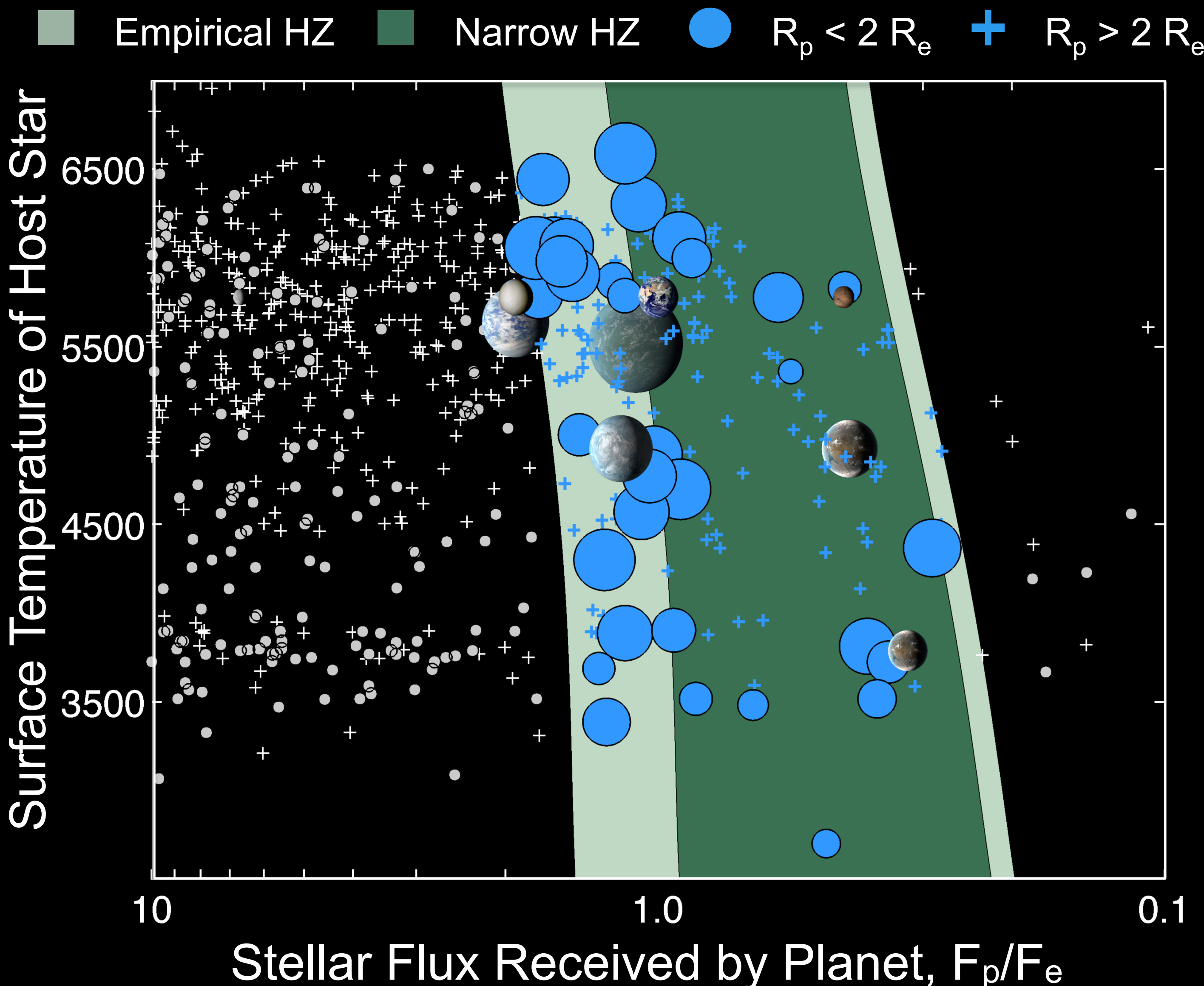
# Sensitivity to small HZ planets




Based on 4 years of data (in progress)

Crosses: suspected false alarms

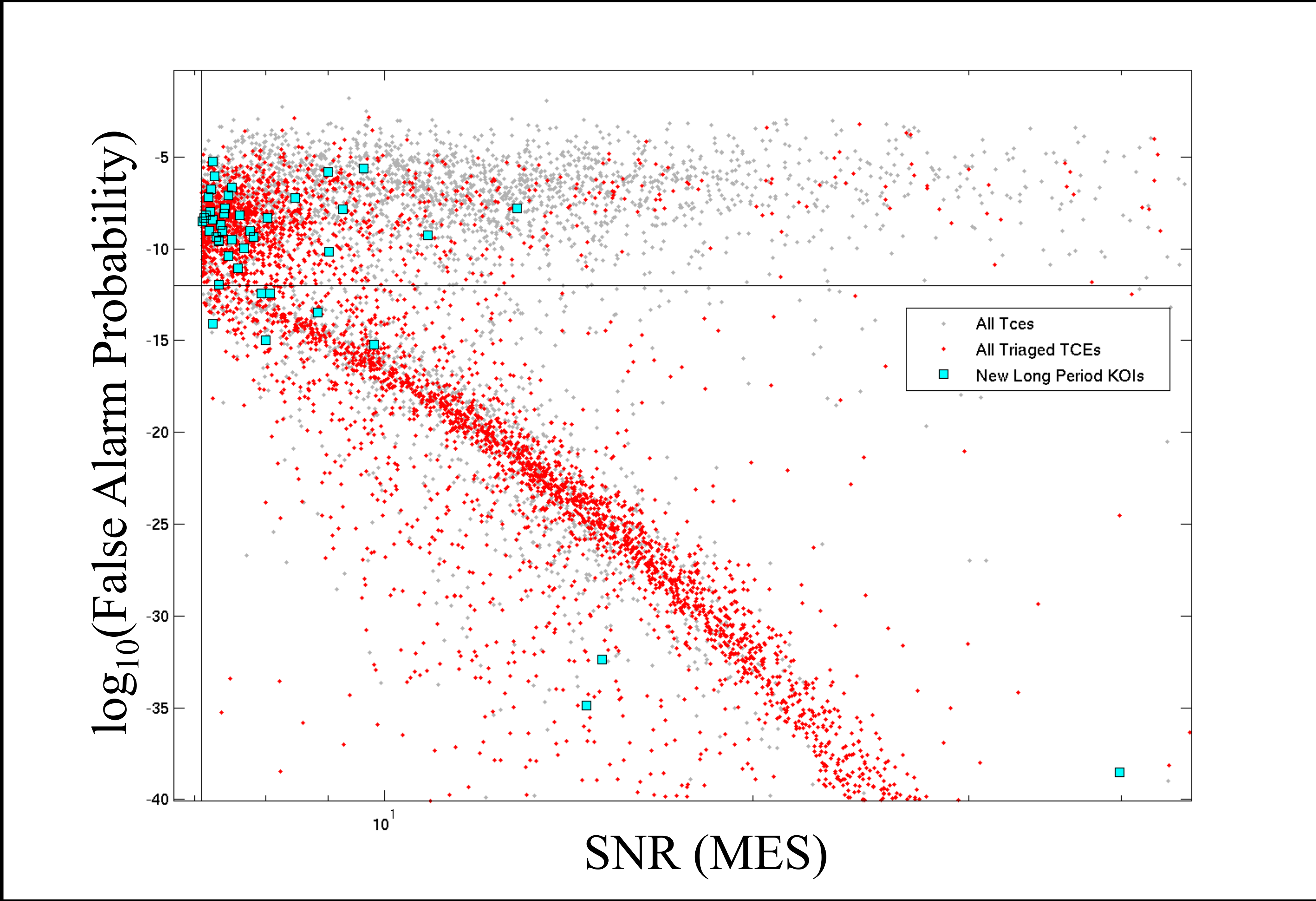
# Sensitivity to small HZ planets




Based on 4 years of data (in progress)

Removing crosses: 36 HZ candidates smaller than 2 $R_e$

# Bootstrap Statistic




Credit: Shawn Seader, Fergal Mullally


Well characterized
planet properties

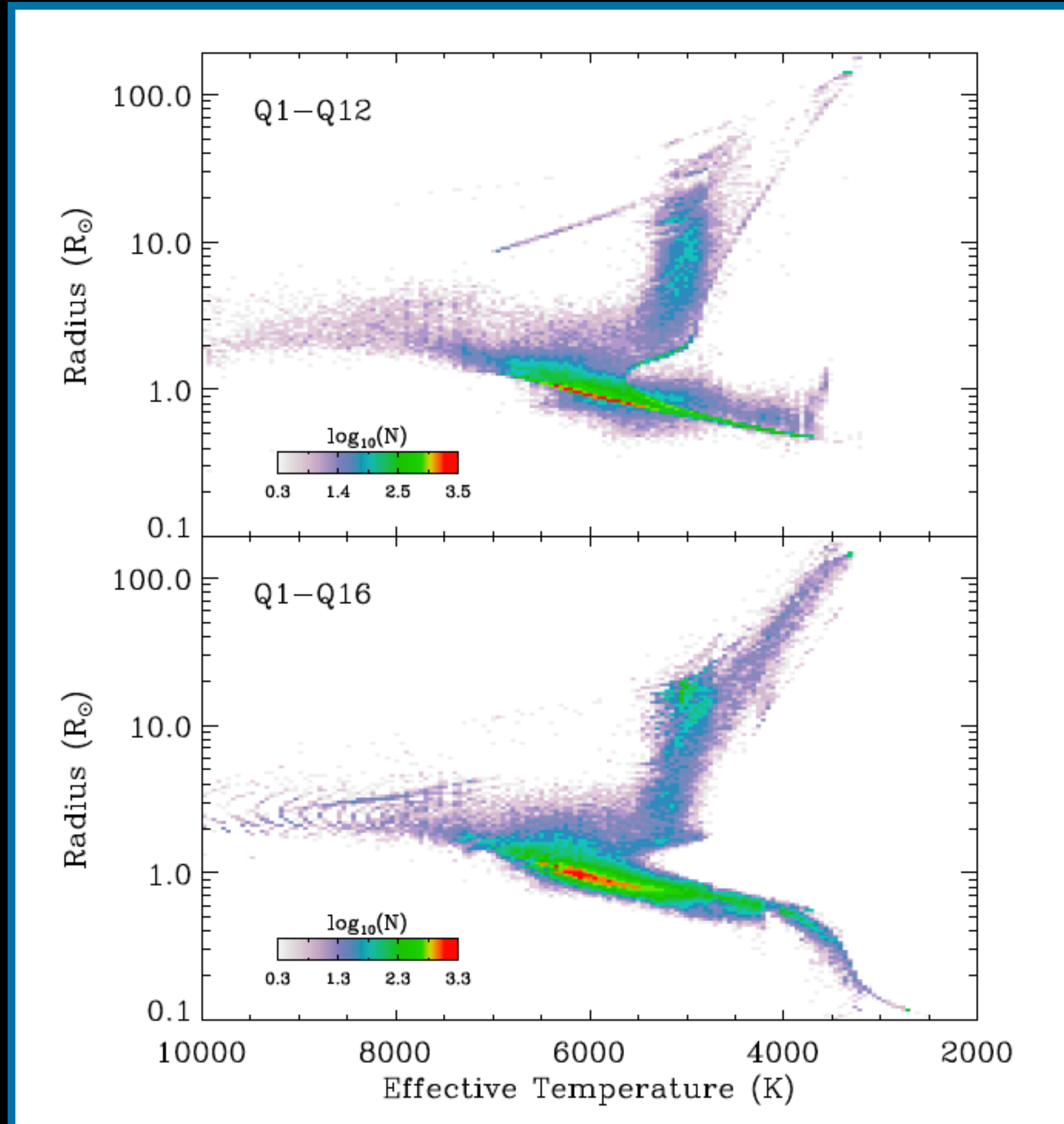
# Well-understood star & planet properties

Radius versus effective temperature of the ~190,000 stars observed by Kepler during the prime mission.

Upper panel: properties used for the 12 quarter pipeline run.

Lower panel: properties used for the 16 quarter pipeline run.

Figure taken from Huber et al. 2014, ApJS, 211, 2.



# Catalog Reliability: Star and Planet Properties

Star Properties Working Group charter includes:

- Deliver catalog of reliable star properties ($T_{eff}$, logg, $R_*$, $M_*$) for the entire Kepler target sample. (chair: D. Huber, **PSP**)

Resources:

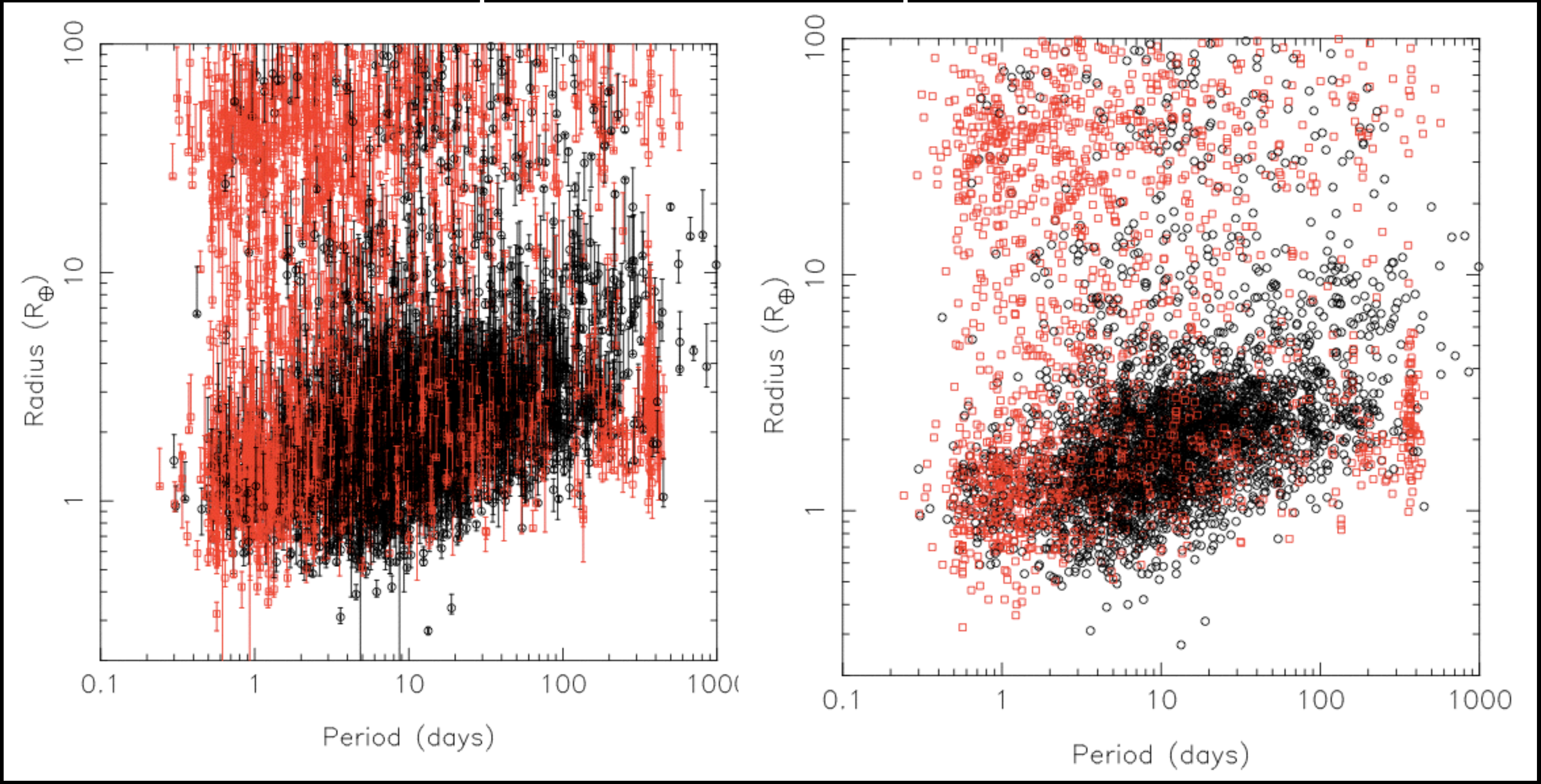
- KIC photometry (Brown et al. 2011)
- Published systematic errors (e.g. Pinsonneault et al. 2012)
- Properties Derived from Kepler Photometry (e.g. Bastien et al. 2013)
- Asteroseismology (**KASC**; e.g. Chaplin ’11, Silva Aguirre ‘11, Bedding ‘10)
- Spectroscopic Surveys: **FOP**, APOGEE, LAMOST, VIRUS (Endl **PSP**)
- Photometric Surveys: UBV, Panstarrs z, Strömgren, Kepler-INT (Ugri)
- Imaging surveys: **FOP** (Keck, Palomar, Gemini), RoboAO

Deliveries (to coincide with pipeline runs):

- Intermediate catalogs that provide corrections to KIC, superceded by properties in the published literature where available
- Final catalog based on analysis of all the available photometry and calibrated against a control sample.

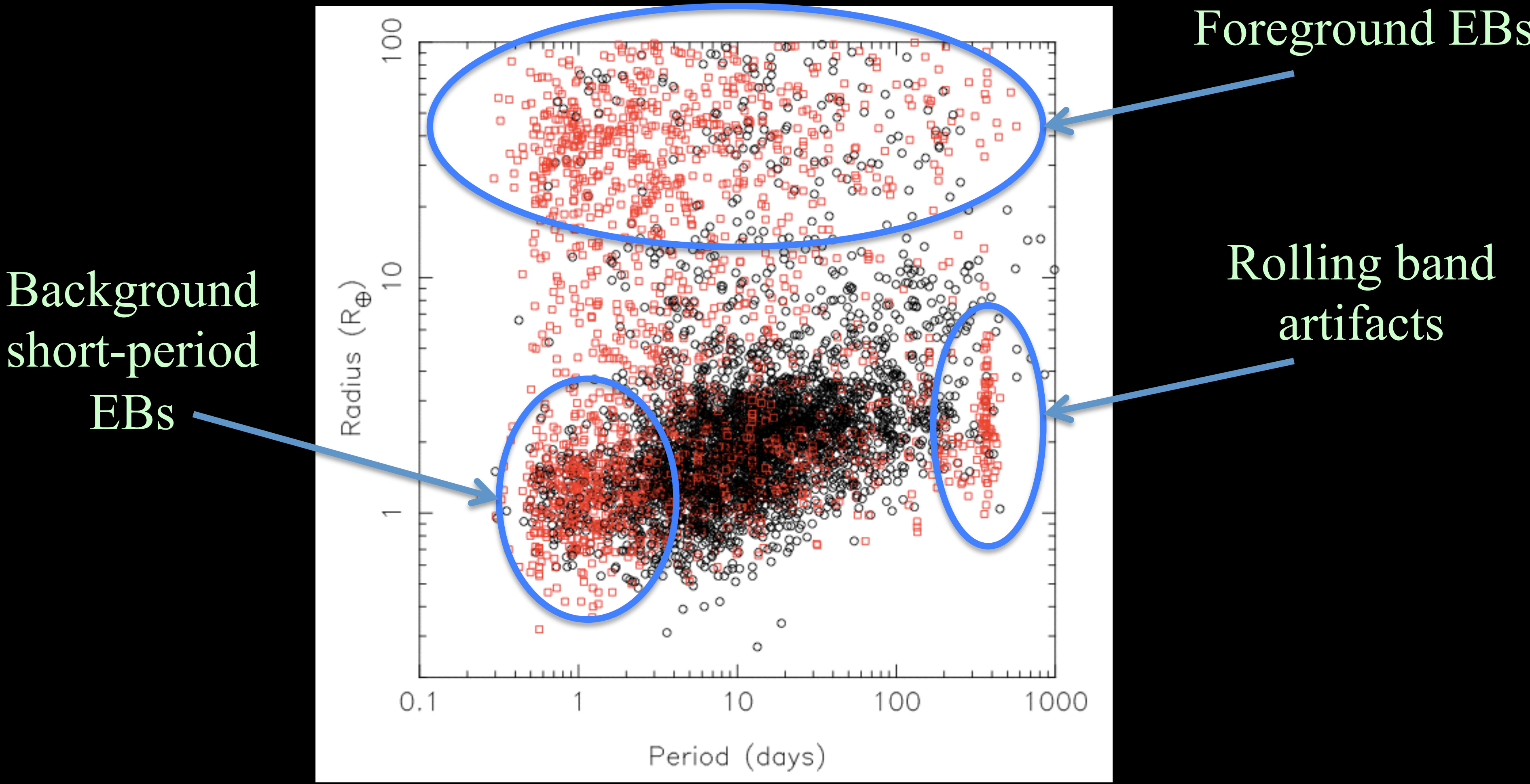
# Understanding Systematics

Red: false positives    Black: planet candidates



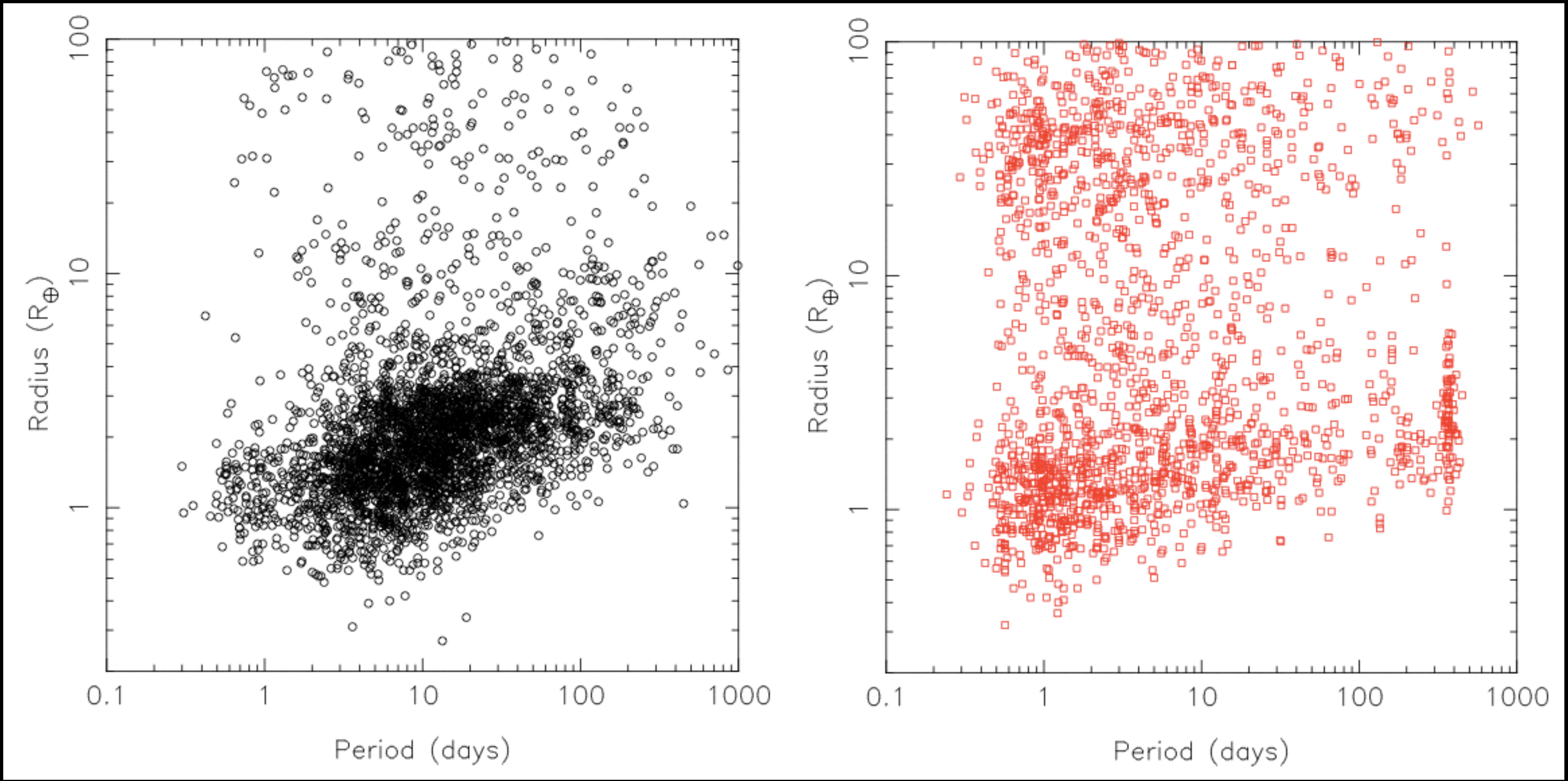

Credit: Jason Rowe

# Understanding Systematics




Credit: Jason Rowe

# Planet vs. EB Period Distribution




Credit: Jason Rowe

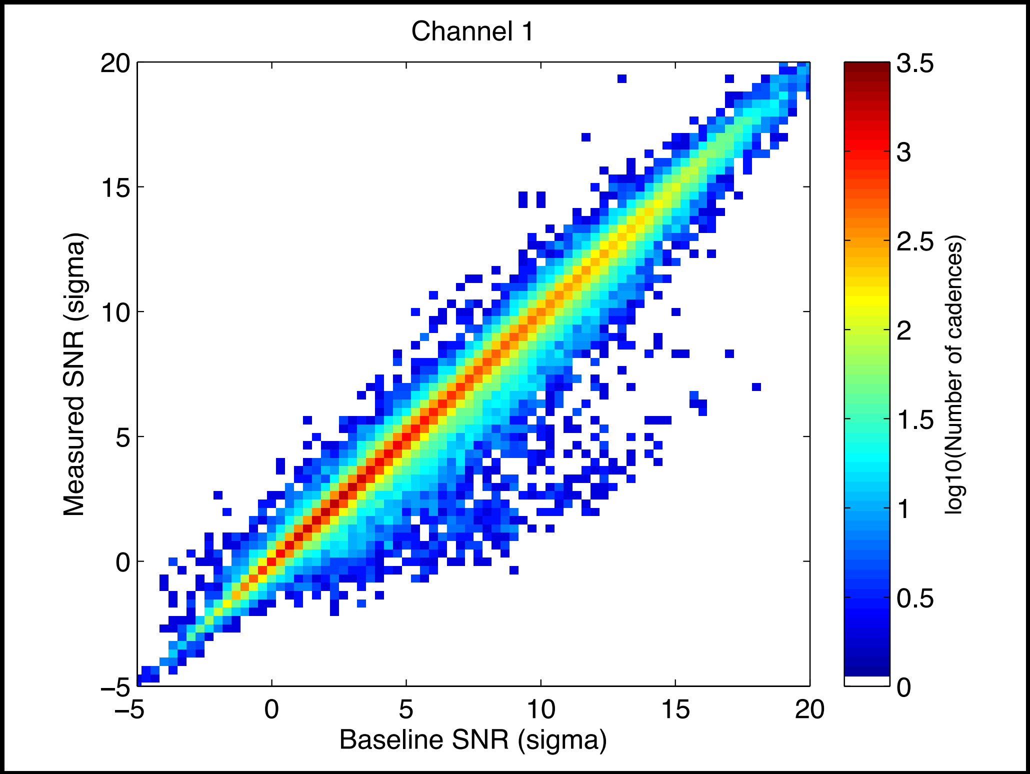

Quantitative
Knowledge of the
Detection Efficiency

# Transit Injection Tests

CAL — PA — PDC — TPS — DV — Tri-age — Dis-posi-tion

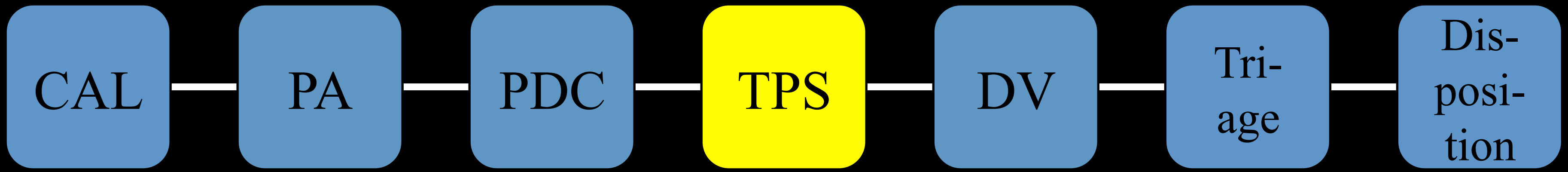
Front end of the pipeline does a very good job preserving transit signals.

Only 0.9% of targets experience significant suppression.
98% fidelity preserving single-transit SNR (3% scatter in SNR)




MS = 0.9973 (+/- 0.0011) x BS - 0.0151 (+/- 0.0049)
Christiansen et al 2013, ApJS, 207, 35 (PSP)

# Transit Injection Tests



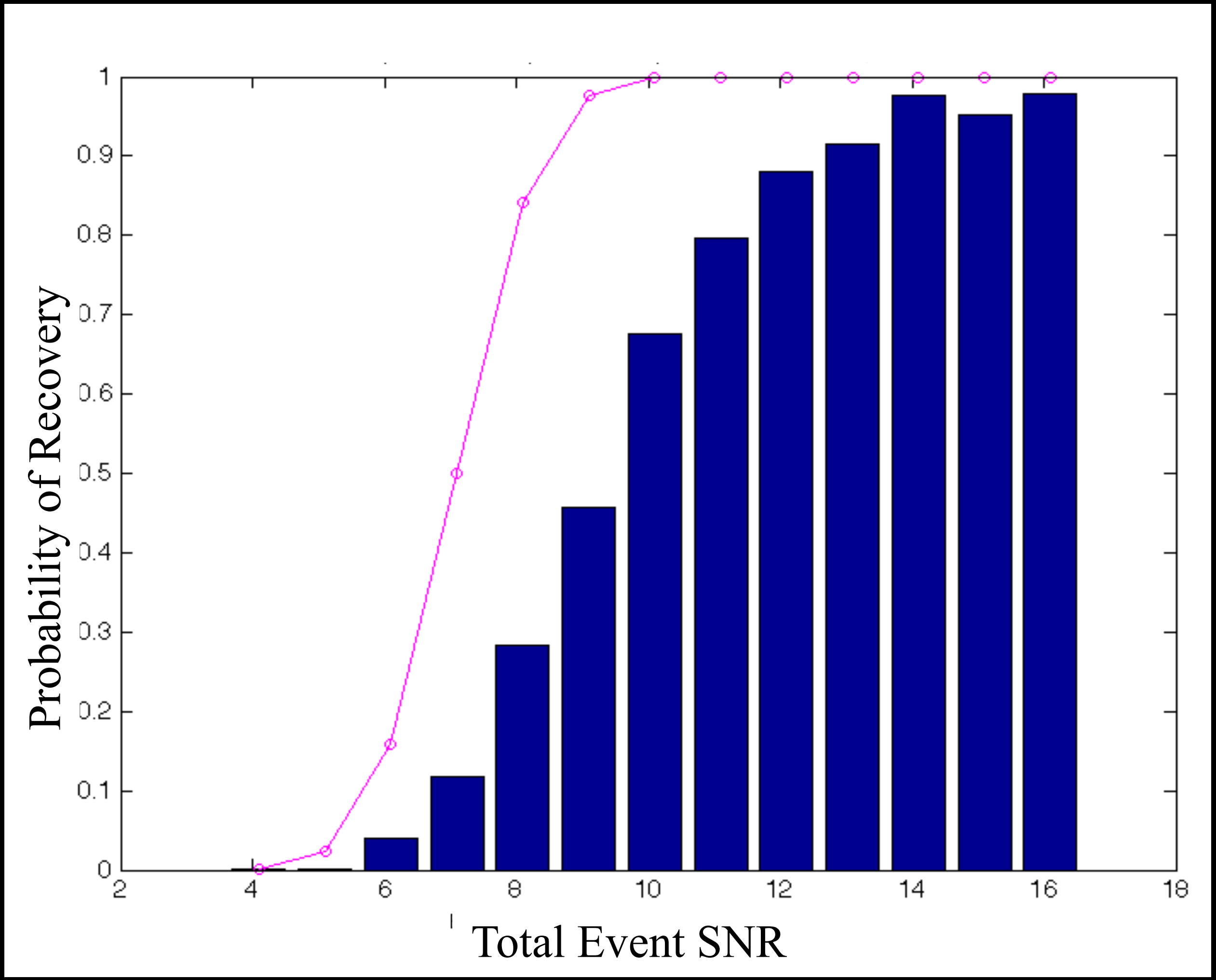

Monte Carlo transit injection in10,080 flux timeseries.

Planets with sizes ranging from 0.5 to 3.0 Re

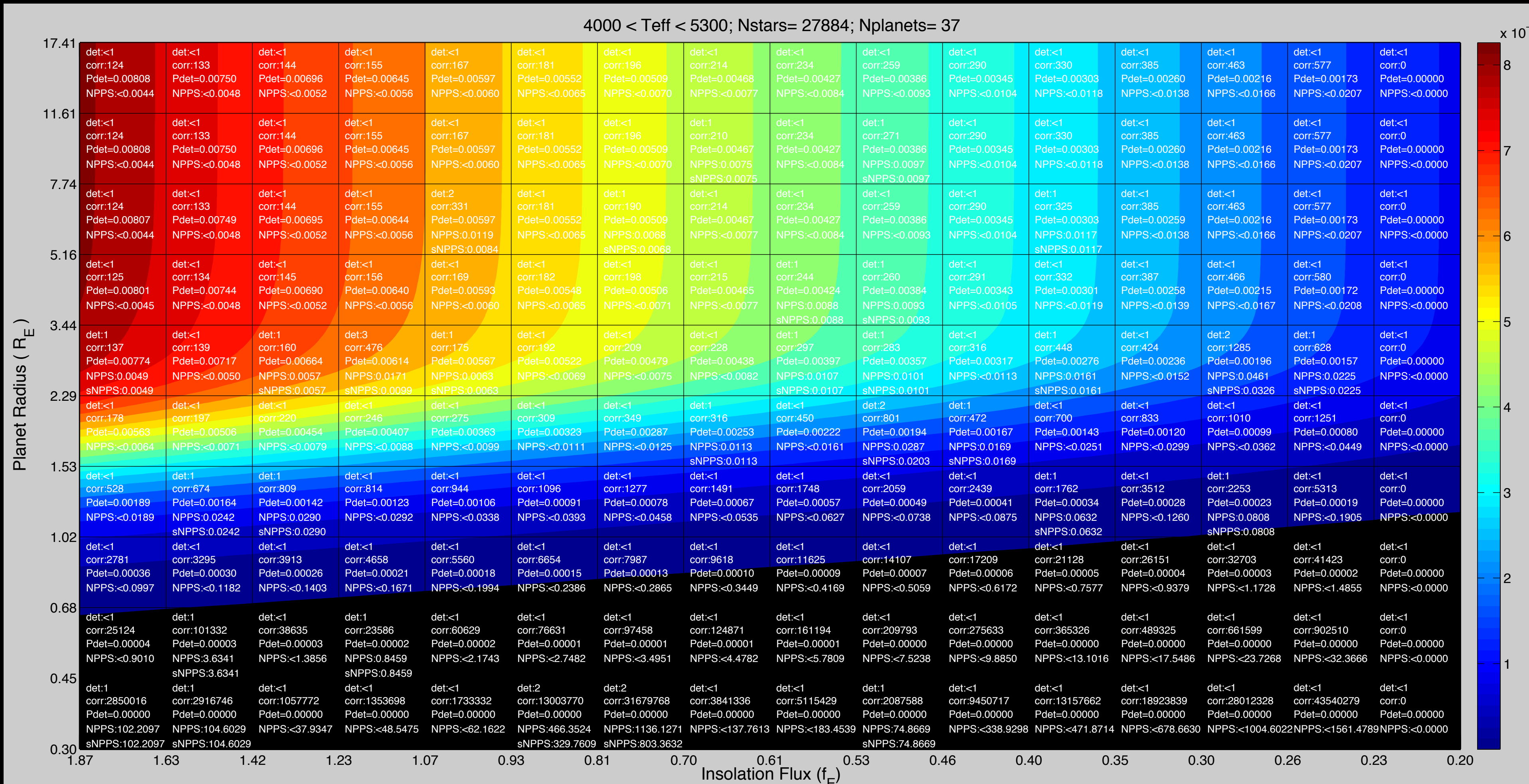
Periods ranging from 50 to 150 days

Pink: gaussian error function

Seader et al. 2013, ApJS, 206, 25

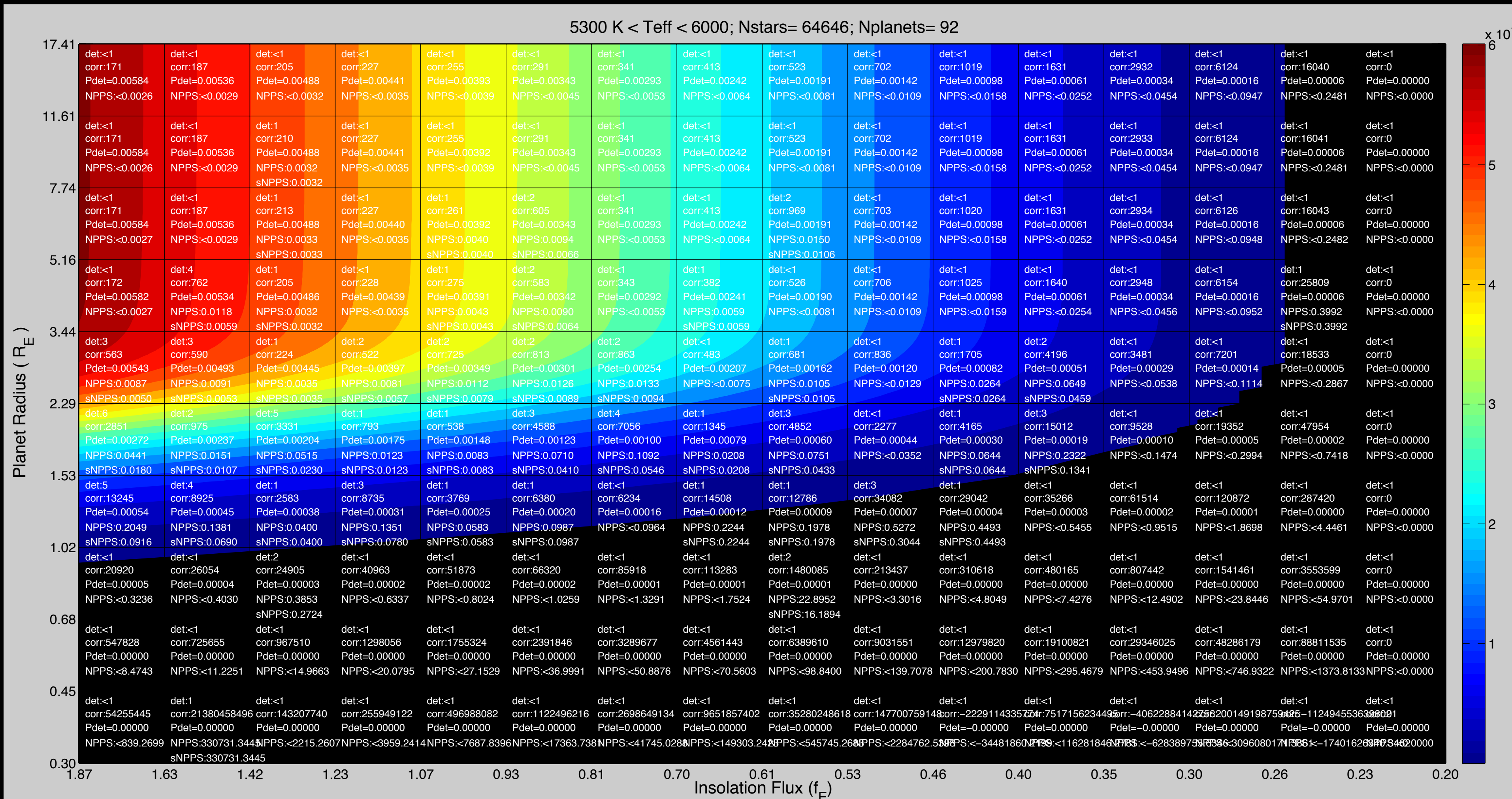


# Knowledge of detection completeness




Sum of all detection probability contours (radius versus insolation flux) of K-type main sequence stars.
Credit: Eduardo Seperuelo Duarte

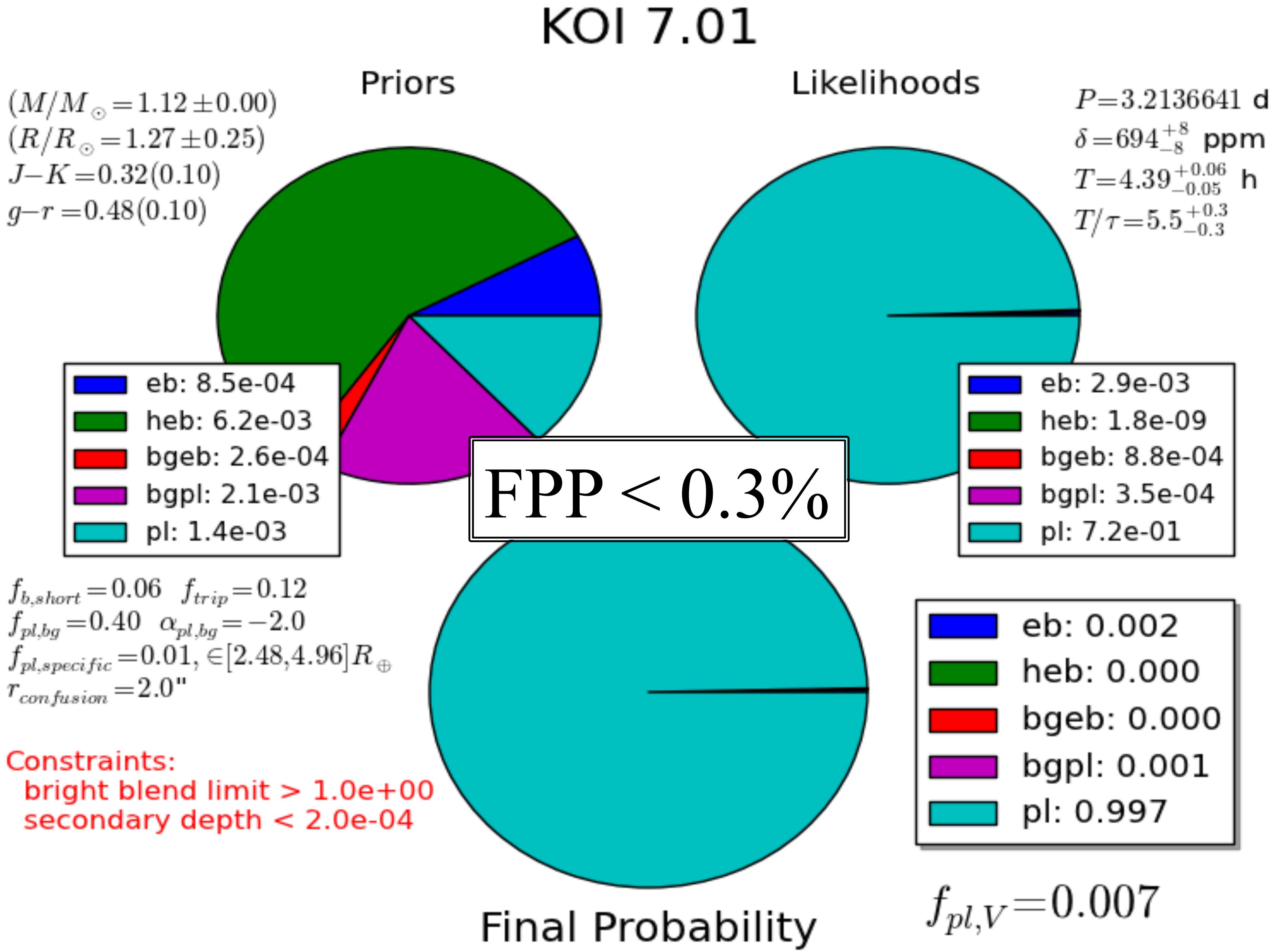
# Knowledge of detection completeness




Sum of all detection probability contours (radius versus insolation flux) of G-type main sequence stars.

Credit: Eduardo Seperuelo Duarte

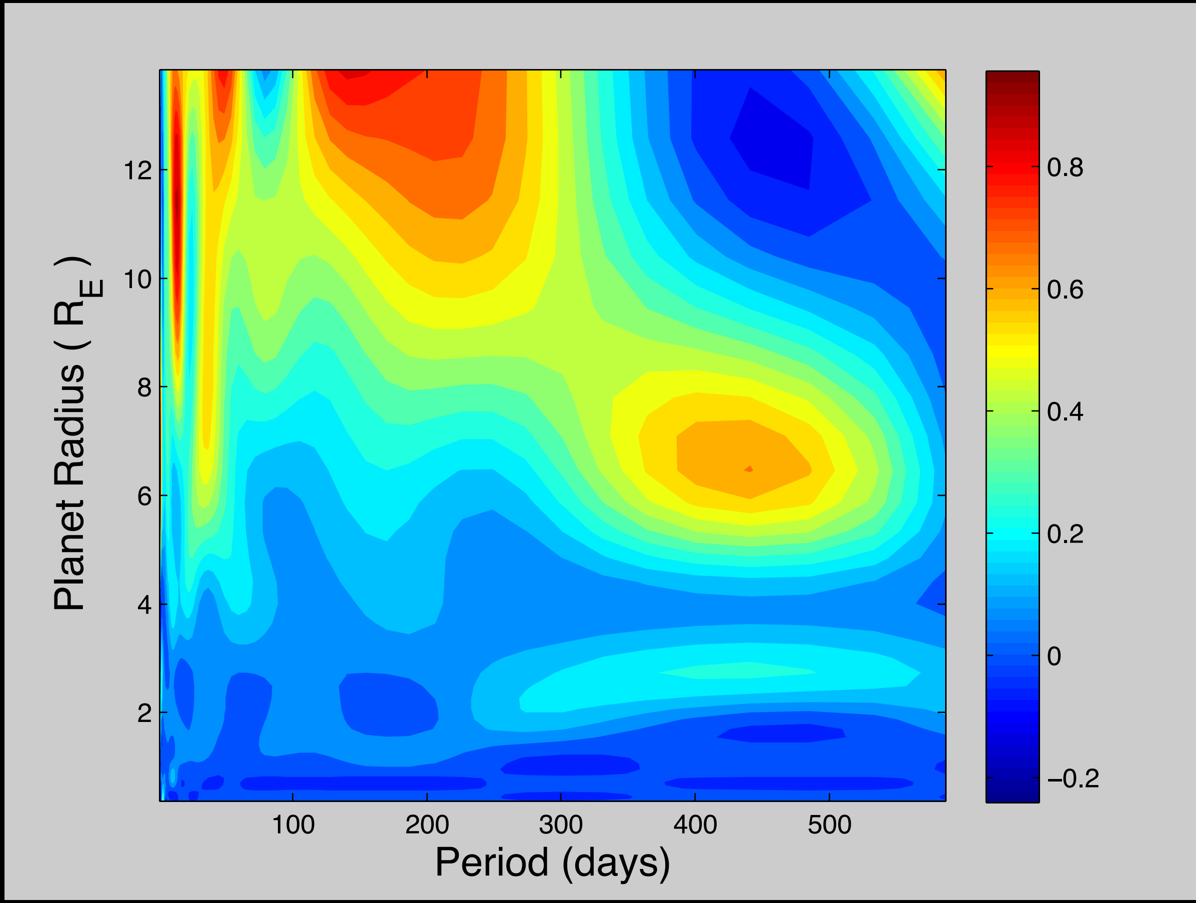

Quantitative
Knowledge of the
Catalog Reliability



KOI 7.01
Priors
Likelihoods
$(M/M_\odot = 1.12 \pm 0.00)$
$(R/R_\odot = 1.27 \pm 0.25)$
$J{-}K = 0.32(0.10)$
$g{-}r = 0.48(0.10)$
$P = 3.2136641$ d
$\delta = 694^{+8}_{-8}$ ppm
$T = 4.39^{+0.06}_{-0.05}$ h
$T/\tau = 5.5^{+0.3}_{-0.3}$
eb: 8.5e-04
heb: 6.2e-03
bgeb: 2.6e-04
bgpl: 2.1e-03
pl: 1.4e-03
eb: 2.9e-03
heb: 1.8e-09
bgeb: 8.8e-04
bgpl: 3.5e-04
pl: 7.2e-01
FPP < 0.3%
$f_{b,short} = 0.06$ $f_{trip} = 0.12$
$f_{pl,bg} = 0.40$ $\alpha_{pl,bg} = -2.0$
$f_{pl,specific} = 0.01, \in [2.48, 4.96] R_\oplus$
$r_{confusion} = 2.0"$
Constraints:
bright blend limit > 1.0e+00
secondary depth < 2.0e-04
eb: 0.002
heb: 0.000
bgeb: 0.000
bgpl: 0.001
pl: 0.997
$f_{pl,V} = 0.007$
Final Probability

# Catalog Reliability: False Positive Probabilities

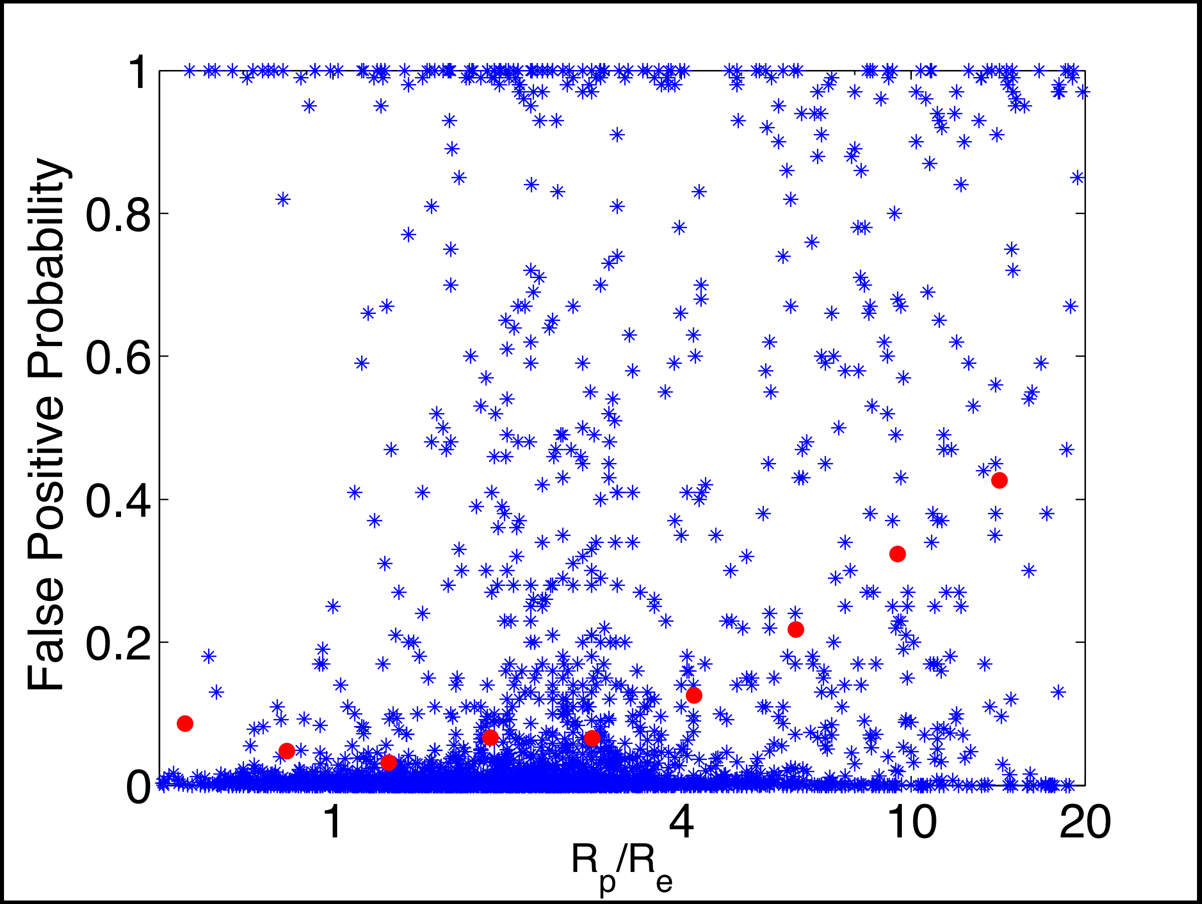
A false positive probability is computed for each planet candidate.
Contour is the average FPP as a function of radius and period.




Credit: T. Morton

# Catalog Reliability: False Positive Probabilities

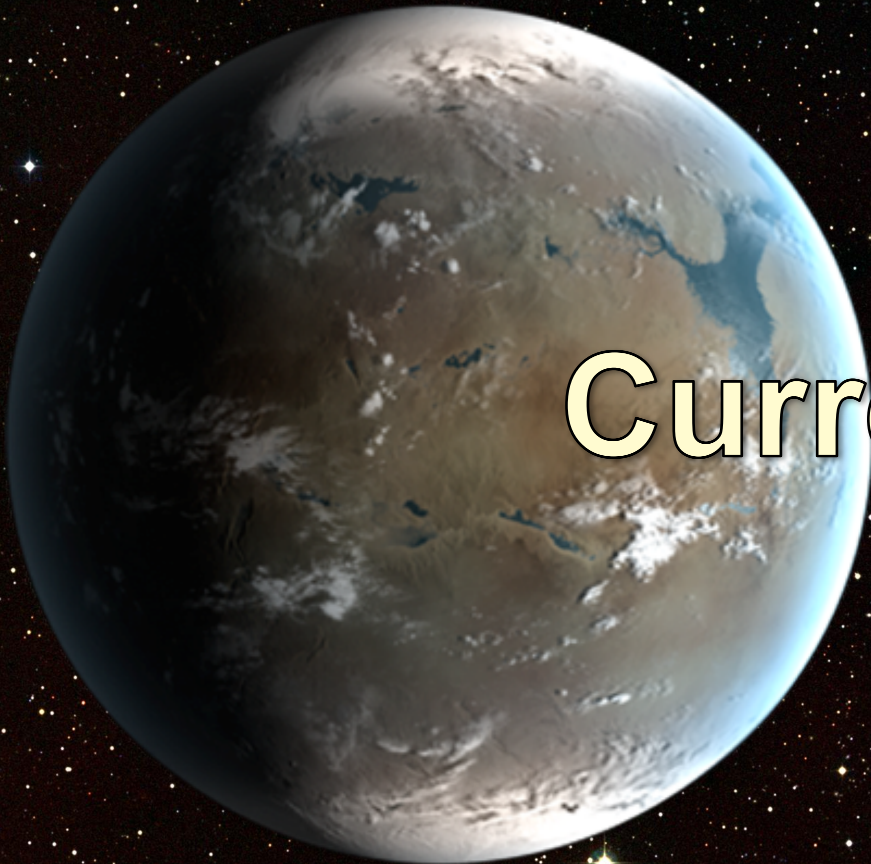
A false positive probability is computed for each planet candidate.
Each red point is the mean for that radius bin.




80% have
FPP < 5%

Credit: T. Morton

# Current Status

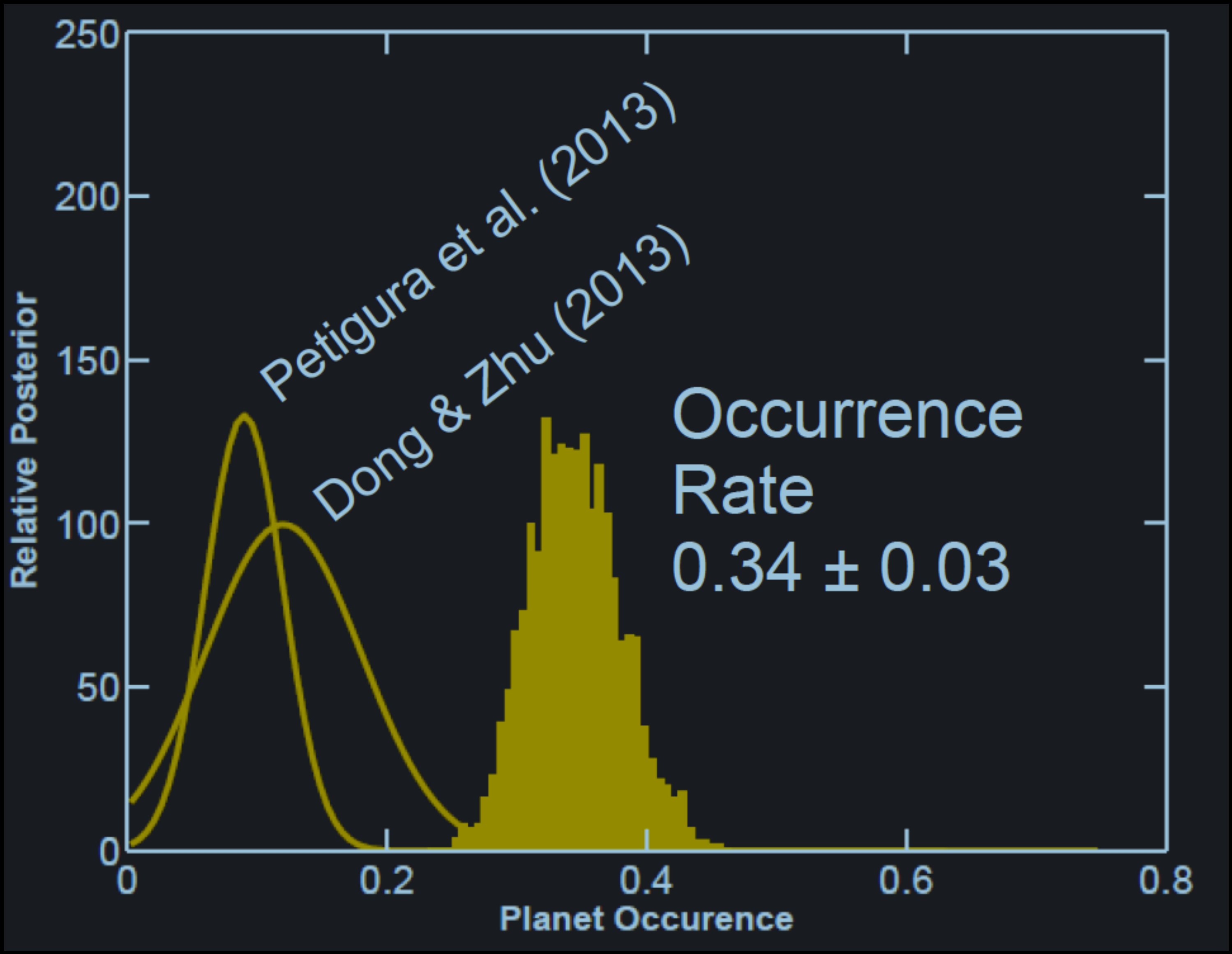


- Borucki et al. 2011 ApJ 736 19, Characteristics of Planetary Candidates Observed by Kepler. II.
- Catanzarite & Shao 2011 ApJ 738 151, The Occurrence Rate of Earth Analog Planets Orbiting Sun-like Stars
- Youdin 2011 ApJ 742 38, The Exoplanet Census: A General Method Applied to Kepler
- Gould & Eastman 2011 arXiv:1102.1009
- Traub 2012 ApJ 745 20, Terrestrial, Habitable-Zone Exoplanet Frequency from Kepler
- Howard et al. 2012 ApJS 201 15, Planet Occurrence within 0.25 AU of Solar-type Stars from Kepler
- Dong & Zhu 2012 arXiv:1212.4853, Statistics of Kepler Planet Candidates Up to 0.75 AU
- Mann et a. 2012 ApJ 753 90, The May be Giants: Luminosity Class, Occurrence, and Metallicity Relations
- Gaidos & Mann 2013 ApJ 762 41, Objects in Kepler's Mirror May be Larger than they Appear
- Beaugé & Nesvorny 2013 ApJ 763 12, Emerging Trends in a Period-Radius Distribution of Close-In Planets
- Fressin et al 2013 ApJ 766 81, The False Positive Rate of Kepler and the Occurrence of Planets
- Dressing & Charbonneau 2013 ApJ 767 95, The Occurrence Rate of Small Planets around Small Stars
- Swift et al. 2013 ApJ 764 105, Characterizing the Cool KOIs. IV
- Petigura et al 2013 ApJ 770 69, A Plateau in the Planet Population Below Twice the Size of Earth
- Gaidos 2013 ApJ 770 90, Candidate Planets in the Habitable Zones of Kepler stars
- Kopparapu 2013 ApJ 767 8, A Revised Estimate of the Occurrence Rate of Terrestrial Planets in the HZ around M-dwarfs
- Morton & Swift arXiv:1303.3013, The Radius Distribution of Small Planets Around Cool Stars
- Gaidos 2013 ApJ 771 18, An Understanding of the Shoulder of Giants: Jovian Planets…
- Petigura et al 2013 PNAS 110 19273, Prevalence of Earth-size planets orbiting Sun-like stars.
- Mulders, Pascucci, Apai 2014 arXiv:1406.7356 A stellar-mass-dependent drop in planet occurrence rates and the abundance of Earth analogs from noisy, incomplete catalogs
- Kane, Kopparapu, Domagal-Goldman 2014 arXiv:1409.2886 On the frequence of Potential Venus Analogs from Kepler Data

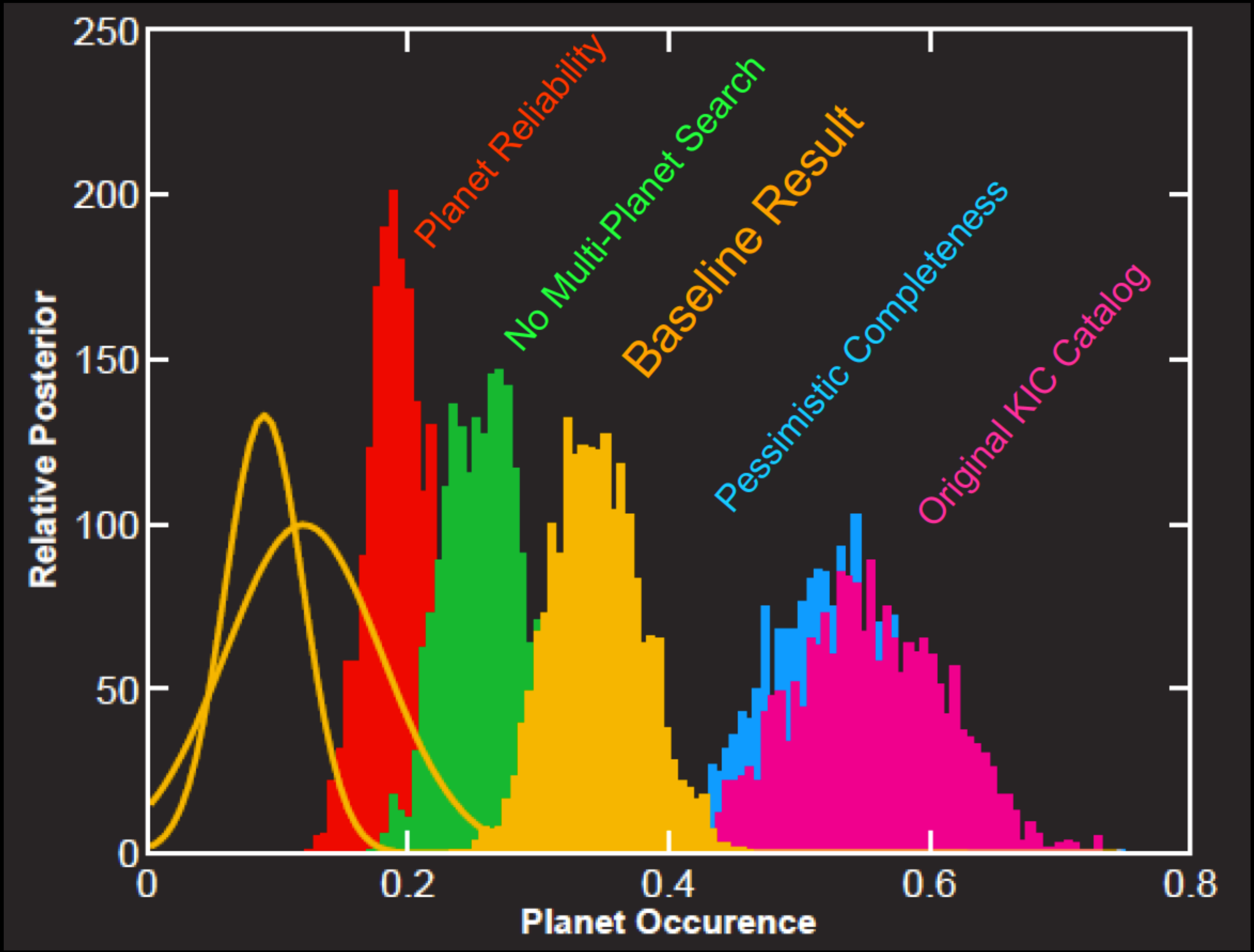
# Q1-Q16 Occurrence Rates: Ingredients

- Stellar Parameters
  - Q1-16 Star Catalog (Huber et al. 2014)
  - Teff < 7000 K ; Rstar < 1.15 Rsun
  - 119,600 Kepler targets
- Planet Parameters
  - Q1-16 Planet Candidate Sample (Mullally et al. in prep.)
  - 50<Porb<300 day ; Rp < 5 Rearth
  - 397 Planet Candidates
- Pipeline Completeness
  - Analytic Star-by-star Detection Model
  - Calibrated With Transit Injection & Recovery
- Methodology
  - Parametric (broken power-law) fit to marginalized distributions (Youdin 2011)

# Q1-Q16 Occurrence Rates

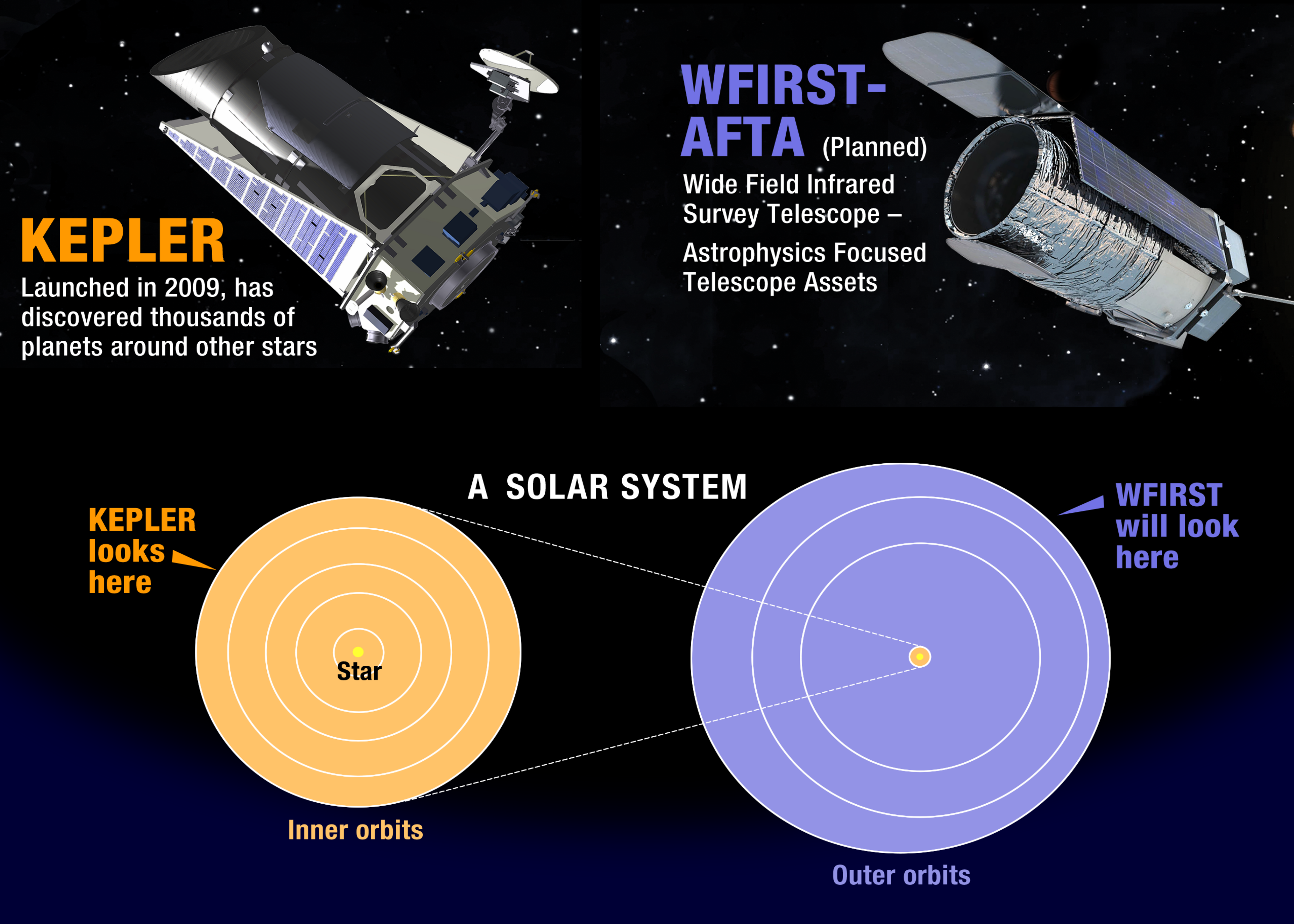


# Q1-Q16 Occurrence Rates: Sensitivity Analysis

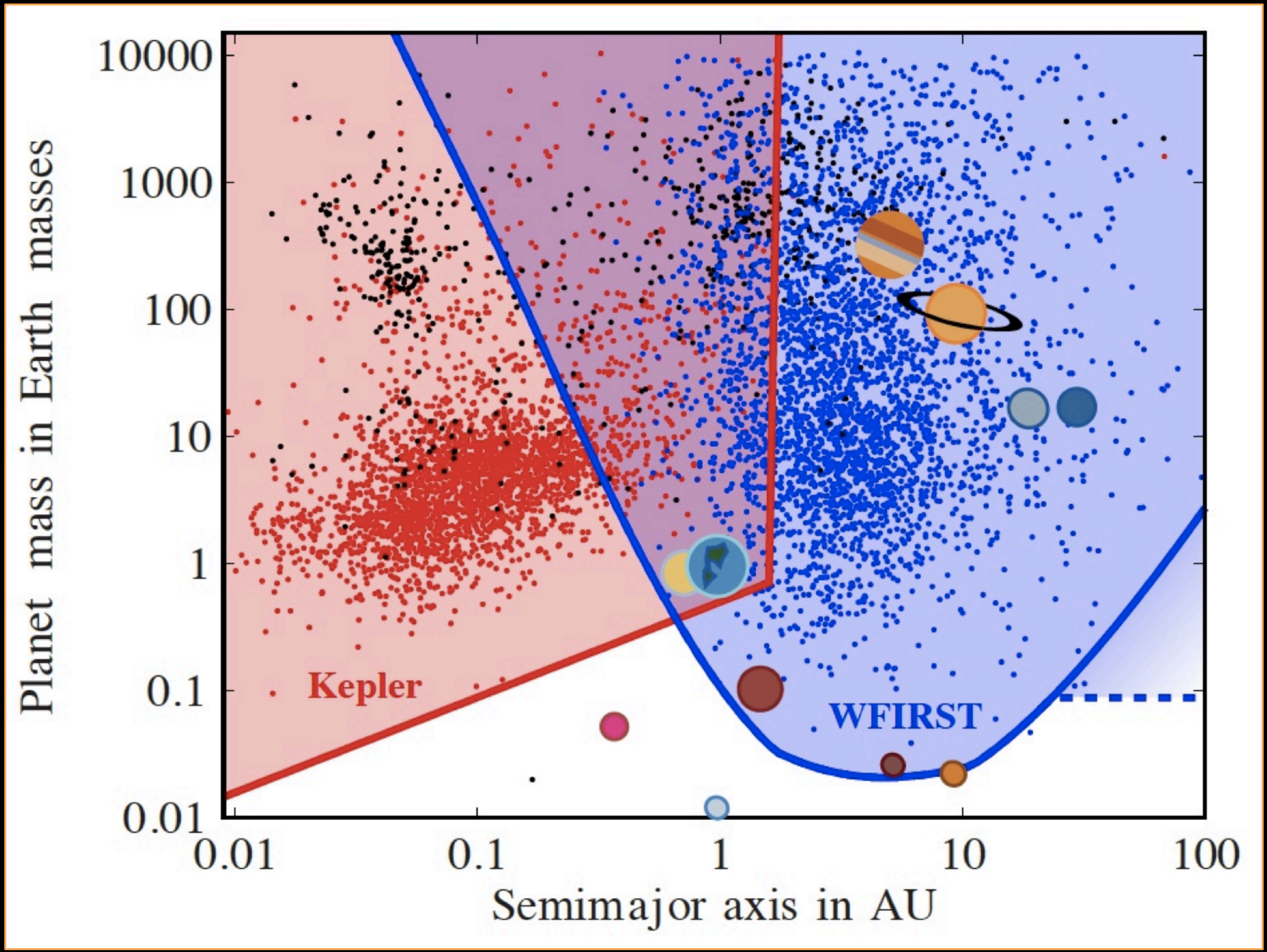


# Summary of Preliminary Results

- Non-negligible systematic errors in planet occurrence rates exist when observational biases are not considered.
- The impact of inadequate star properties on planet occurrence rates is comparable to ignoring completeness and reliability.
- Resulting Occurrence Rate
  - 1 < Rp < 5 Re; 50 < Porb < 300 days
  - 0.79 +/- 0.06 planets per star
  - Systematic range: 0.44 to 1.24
- Burke et al., in preparation

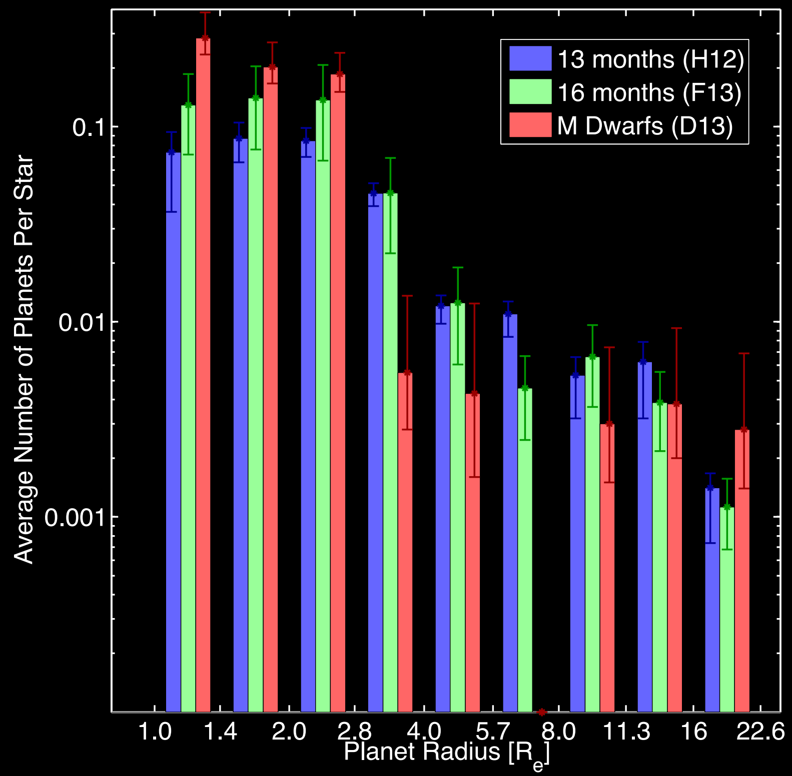

KEPLER
Launched in 2009, has discovered thousands of planets around other stars
WFIRST-AFTA (Planned)
Wide Field Infrared Survey Telescope –
Astrophysics Focused Telescope Assets
A SOLAR SYSTEM
KEPLER looks here
Star
Inner orbits
WFIRST will look here
Outer orbits

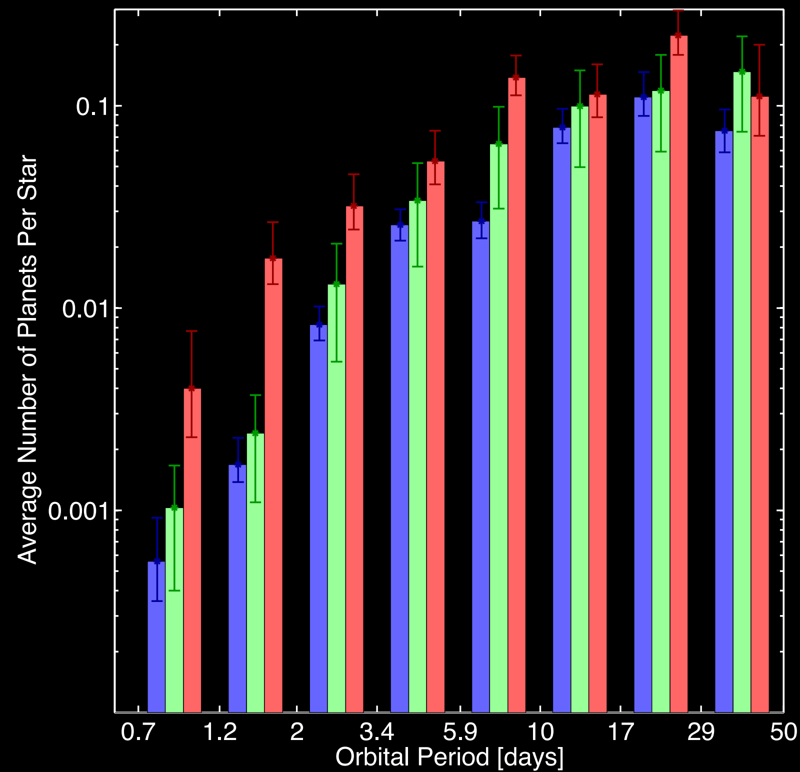
# Kepler and WFIRST: completing the statistical census




Both likely to yield small sample statistics for the earth-analogs

# Planet Occurrence Rates: Selected Results




Log Radius Distribution (periods less than 50 days) :

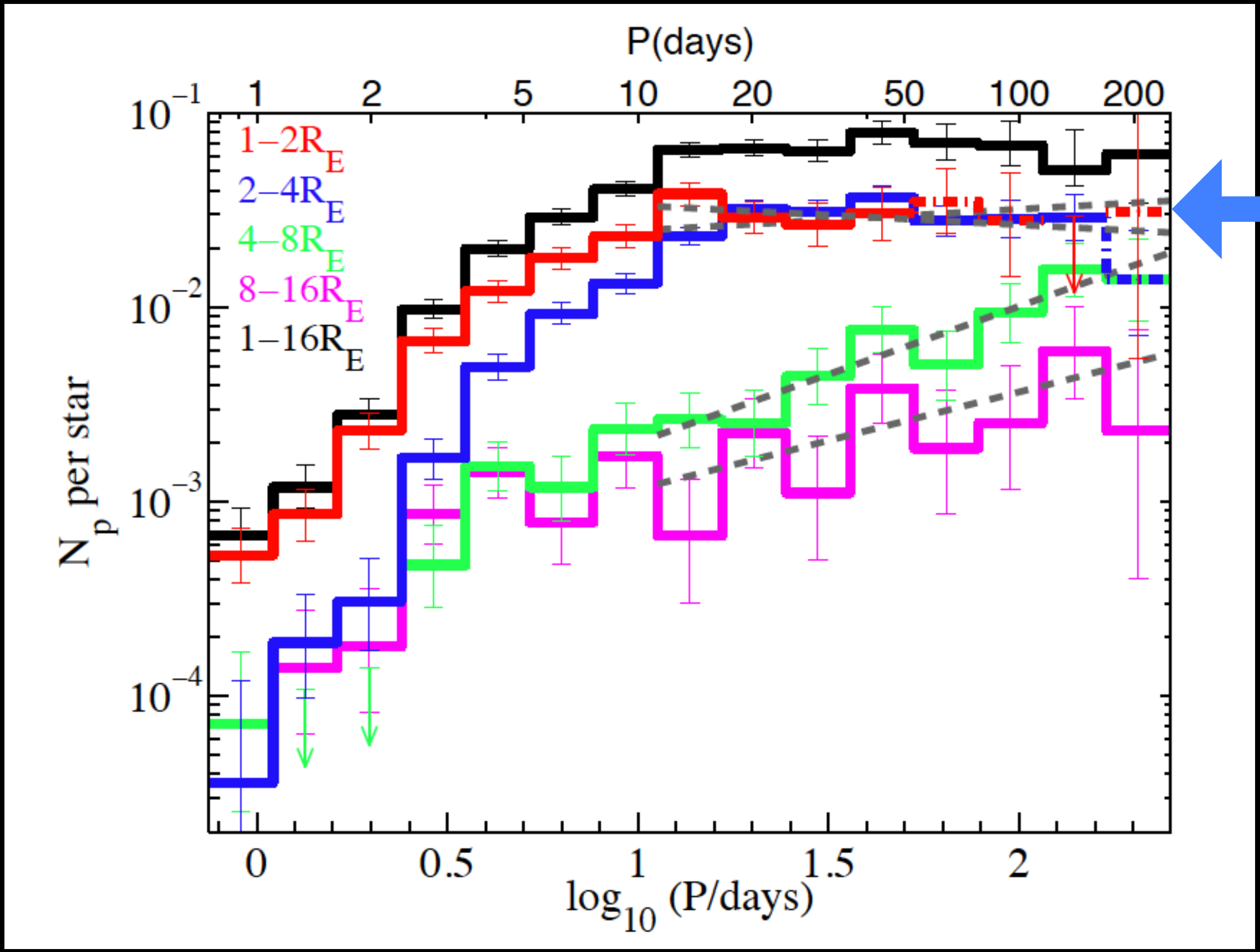
- Sharp rise at ~ 3 $R_e$
- Flat from 1-3 $R_e$

Log Period Distribution ($0.5$ to $22$ $R_e$):

- Rises sharply but flattens out beyond 10 day orbits.

Average number of ($R_p$=0.8 to 1.25) planets per star: $16.5 \pm 3.6\%$

# Planet Occurrence Rates: Selected Results

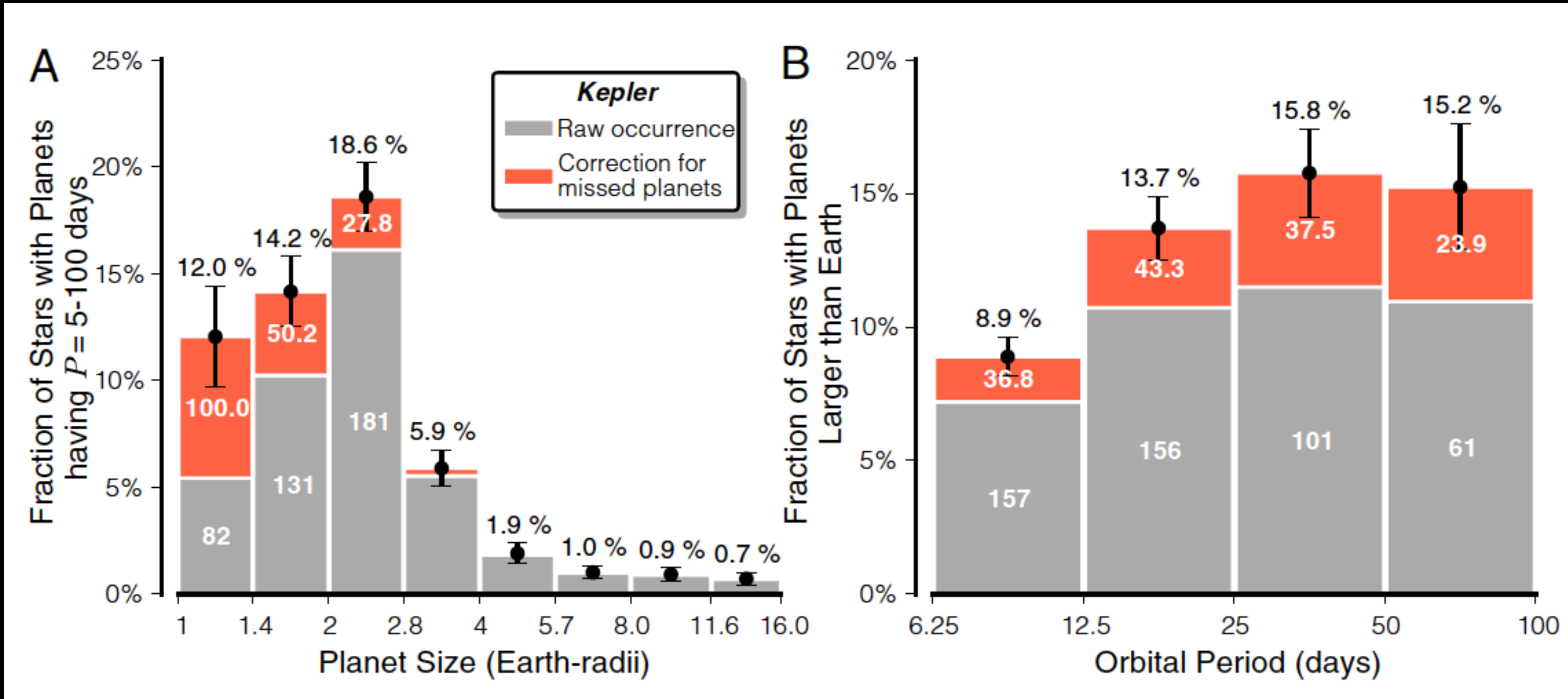



1 - 4 $R_e$

Period distribution (logarithmic) is flat from 10-200 days for small planets while giant planets are increasing in frequency.

Dong & Zhu 2012

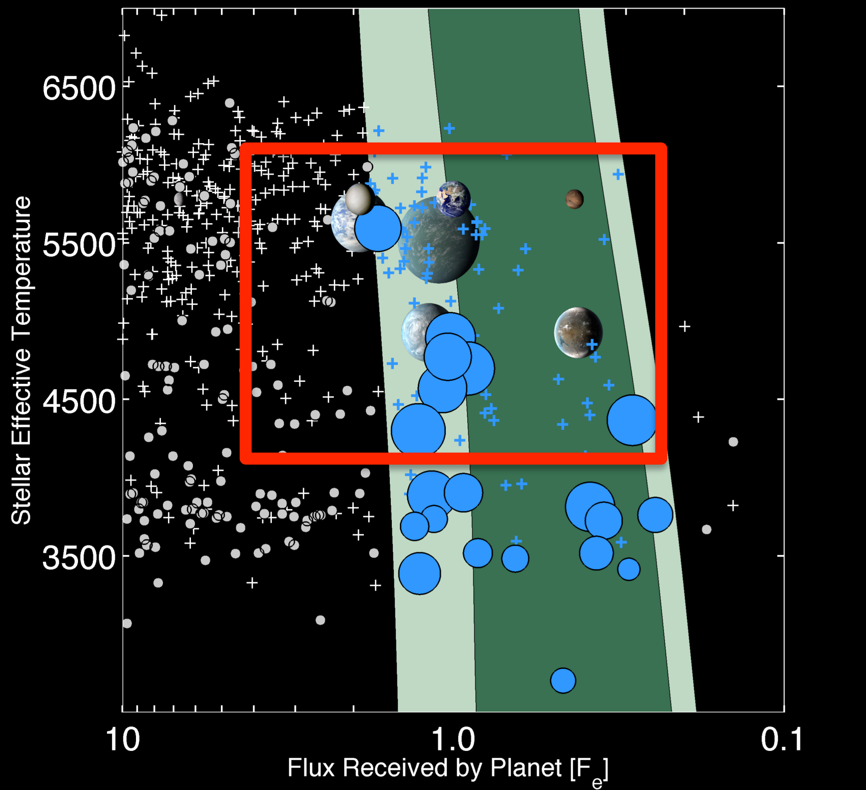
# Planet Occurrence Rates: Selected Results

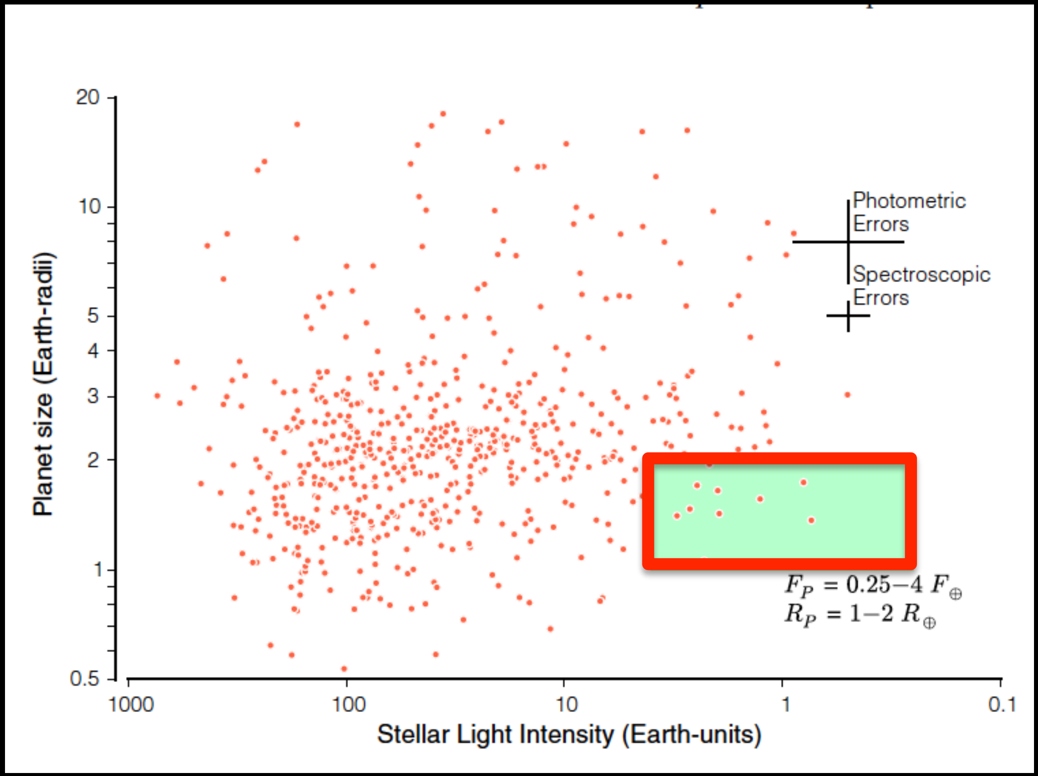



- Independent transit detection pipeline applied to Q1-Q16 data
- ~42,000 (bright) G & K stars (4100-6100 K)
- Identified ~600 planet candidates
- Measured detection efficiency via transit injection

Petigura, Howard, & Marcy 2013

# Planet Occurrence Rates: Selected Results

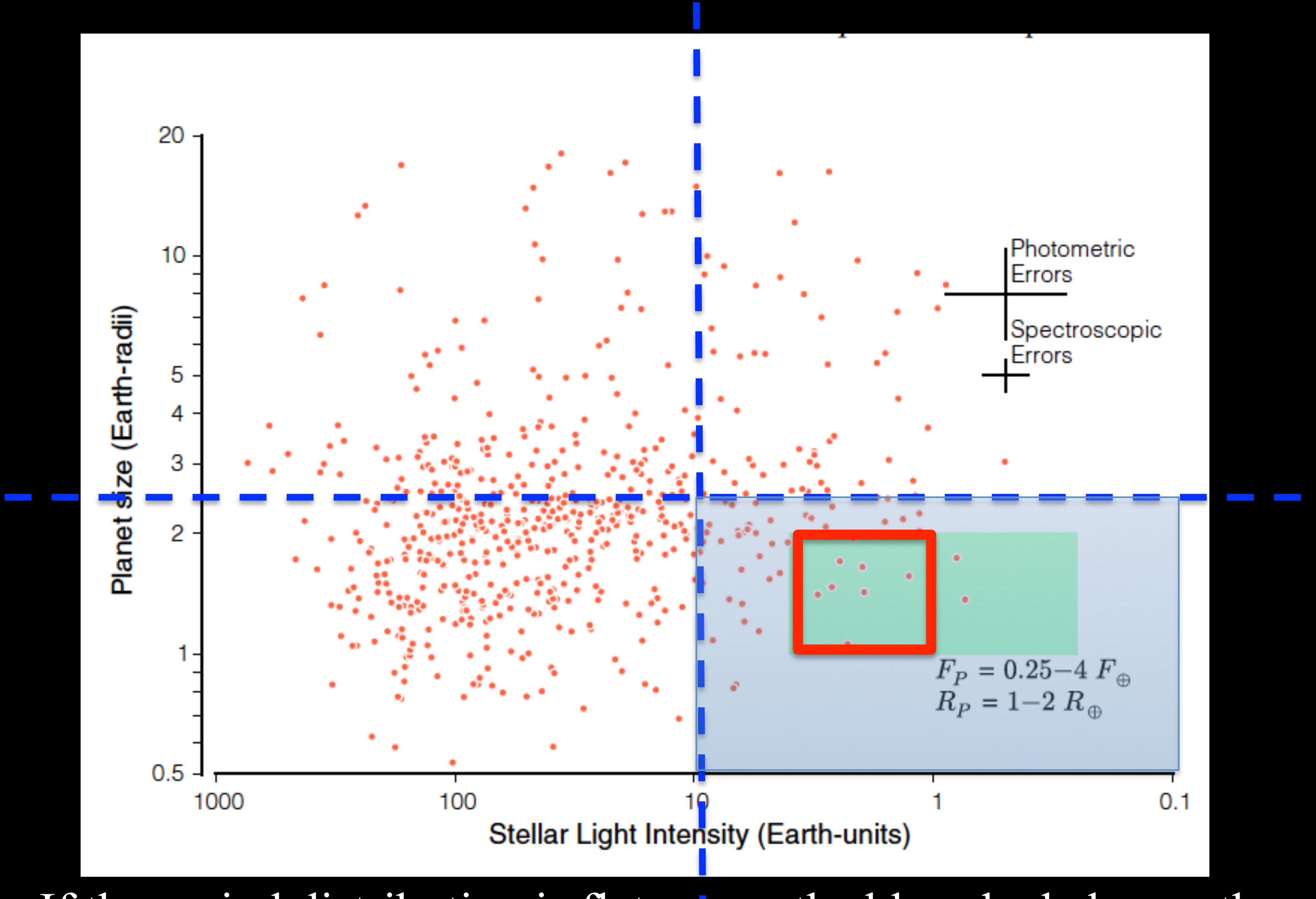



Average number of planets smaller than 1.4 $R_e$ in the HZ of M dwarfs is 0.53 ± 0.17 (Dressing & Charbonneau 2013; Kopparapu 2013; Gaidos 2013)

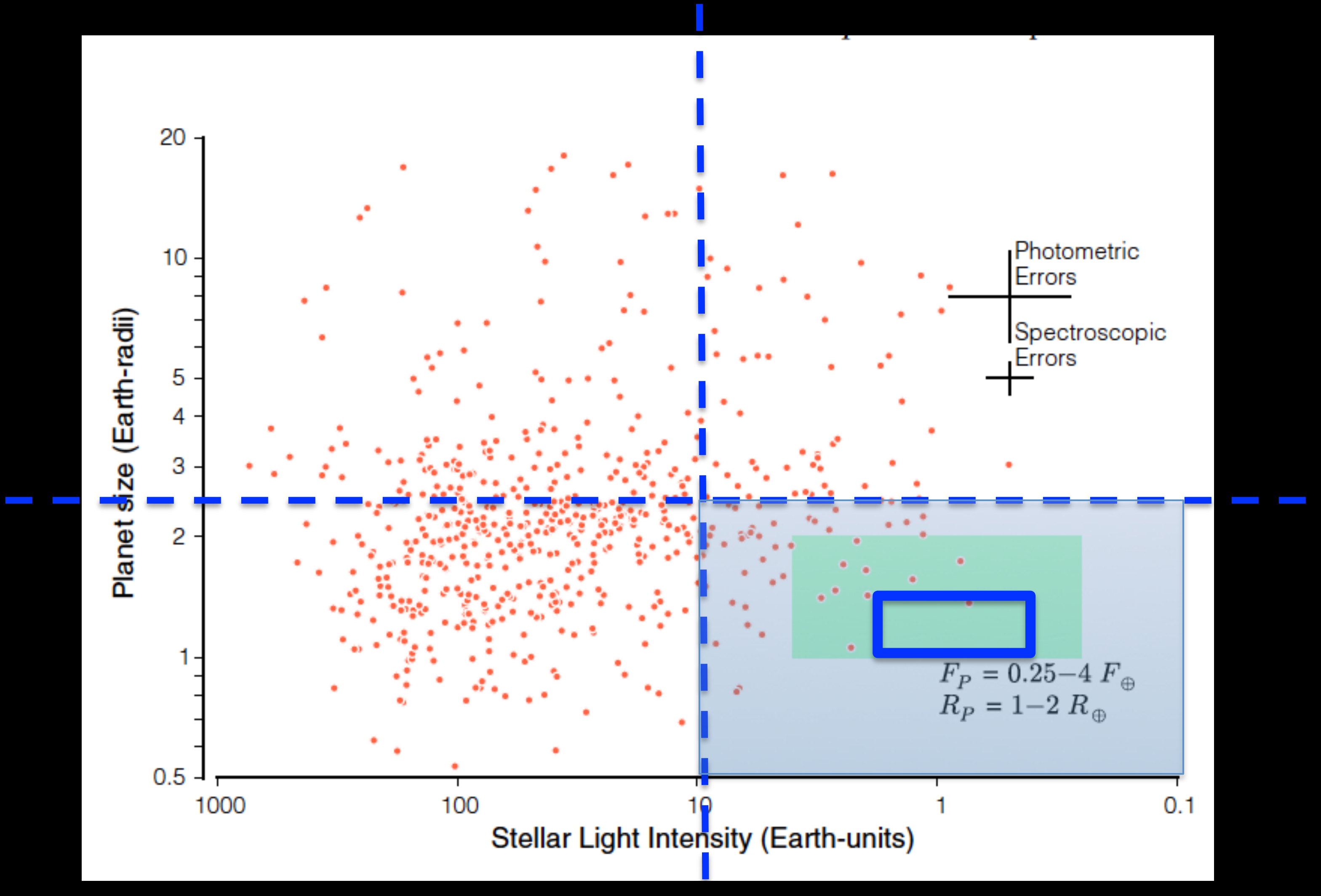
22% of G & K dwarfs harbor a planet smaller than 2 $R_e$ for HZ defined by 0.25 – 4 $F_e$ (Petigura et al. 2013).

# Planet Occurrence Rates: Selected Results



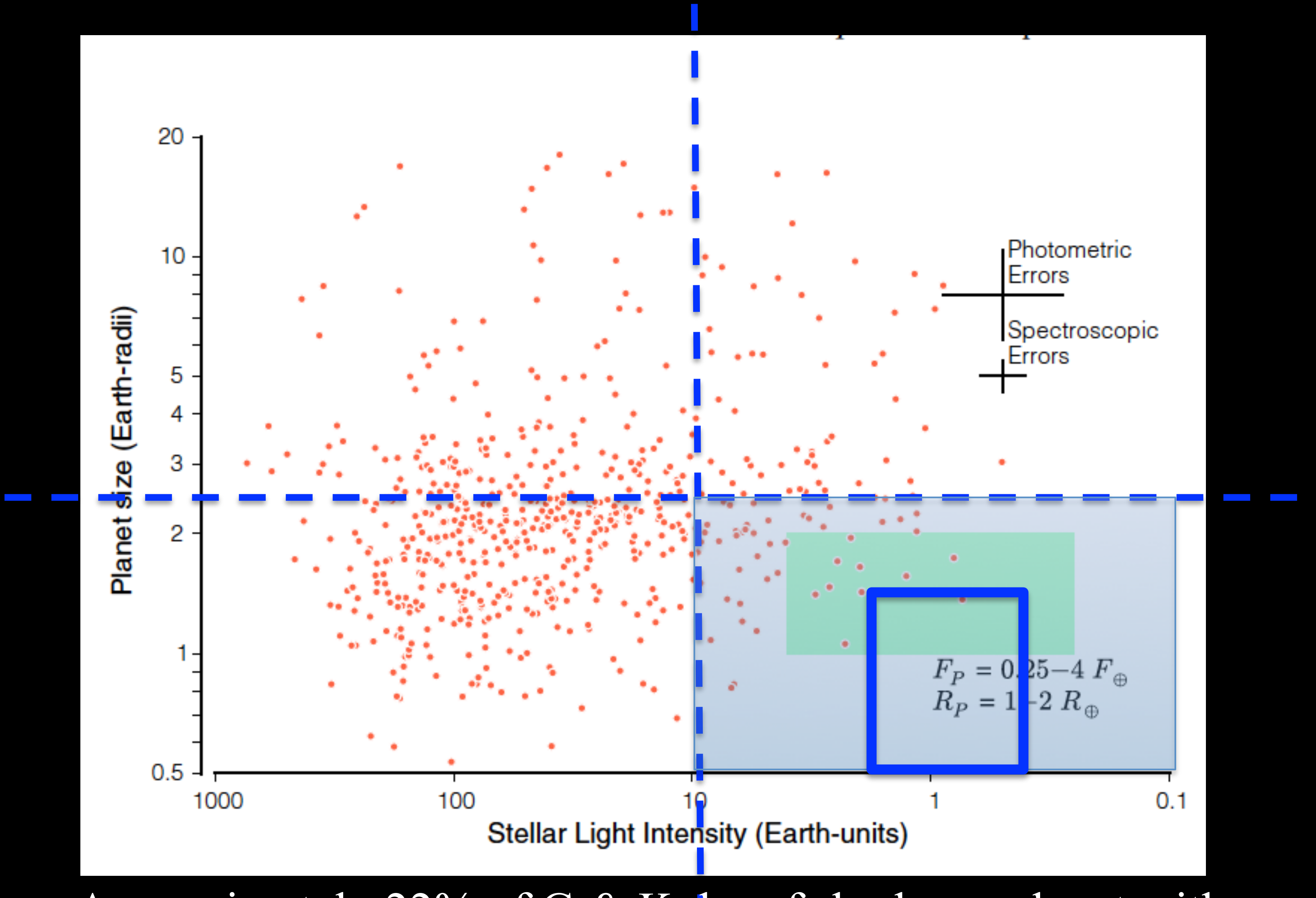

If the period distribution is flat across the blue shaded area, the occurrence rate will be the same in any equal area box contained within that region. Red box: 11%; Green box: 22%

# Planet Occurrence Rates: Selected Results




Approximately 7% of G & K dwarfs harbor a planet with $R_p$ = 1 to 1.4 $R_e$ in the optimistic (empirical) HZ.

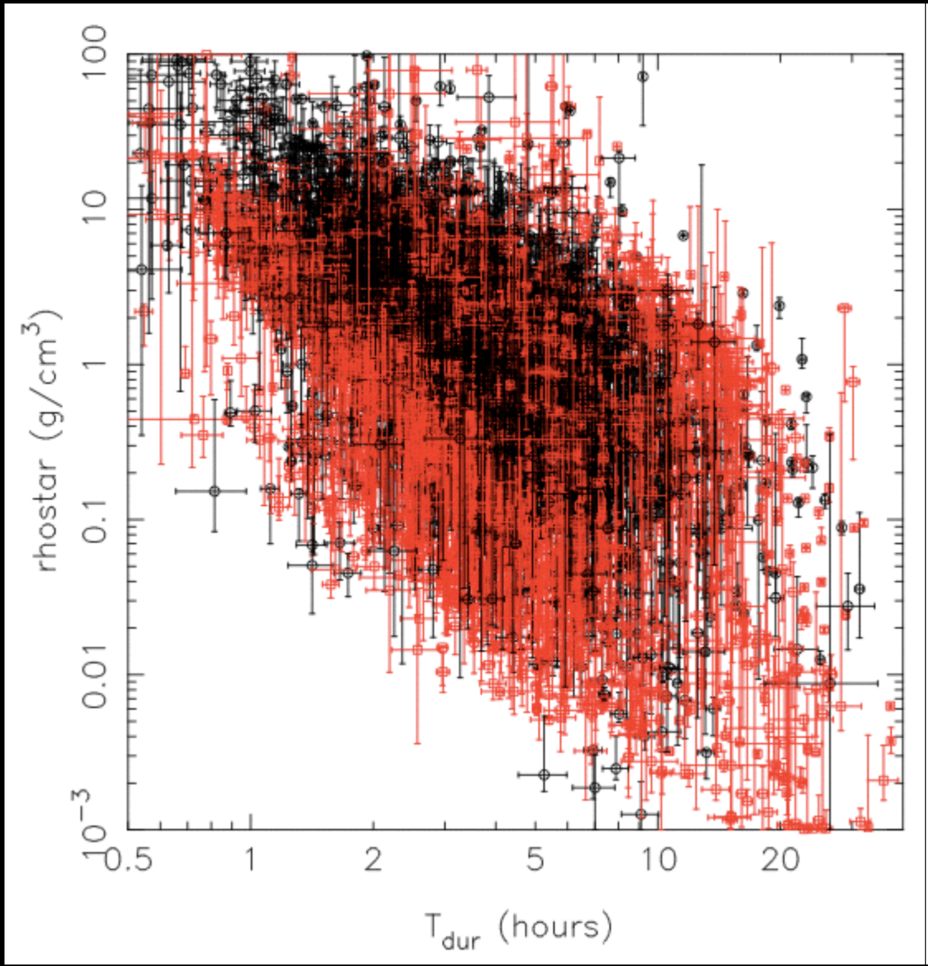
# Planet Occurrence Rates: Selected Results

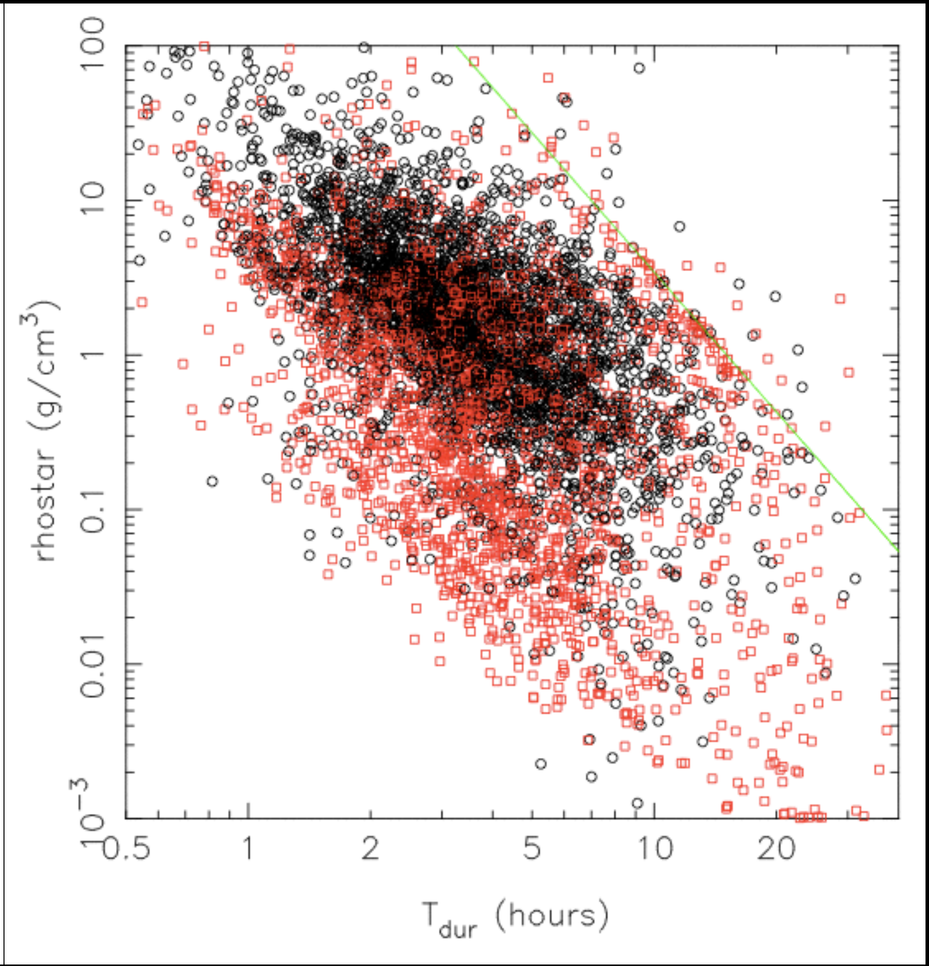



Approximately 22% of G & K dwarfs harbor a planet with $R_p$ = 0.5 to 1.4 $R_e$ in the optimistic (empirical) HZ.

# Stellar Density vs. Transit Duration

$$\frac{\rho_*}{\rho_\odot} \equiv \frac{M_*/M_\odot}{(R_*/R_\odot)^3} = \left[\frac{4\pi^2}{P^2G}\right]\left[\frac{(1+\sqrt{\Delta F})^2 - b^2(1-\sin^2\frac{t_T\pi}{P})}{\sin^2\frac{t_T\pi}{P}}\right]^{3/2}.$$



# Transit Injection Tests

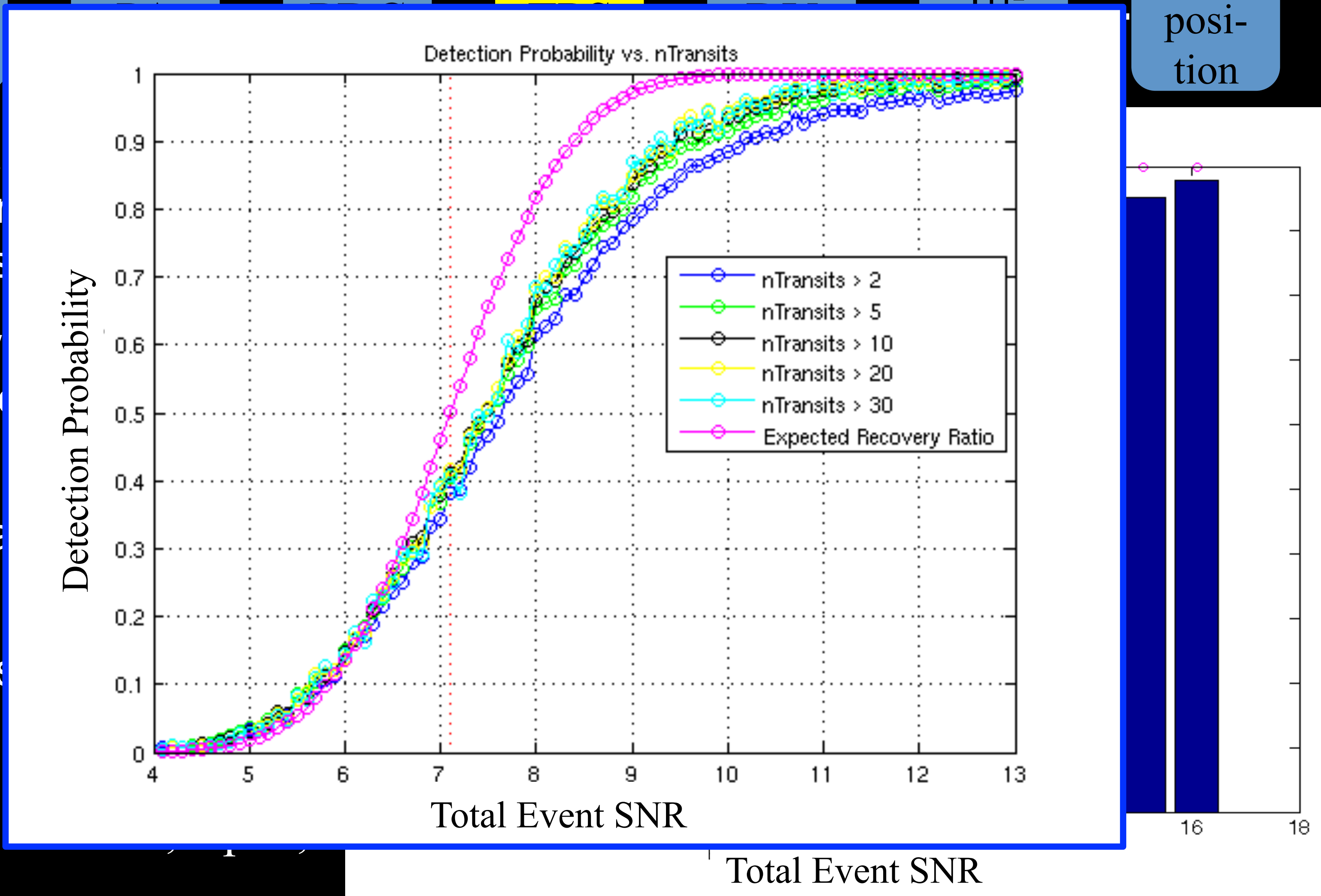
CAL | Tri- | Dis-posi-tion

Monte Ca
injection i
timeseries

Planets w
ranging fr
Re

Periods ra
50 to 150

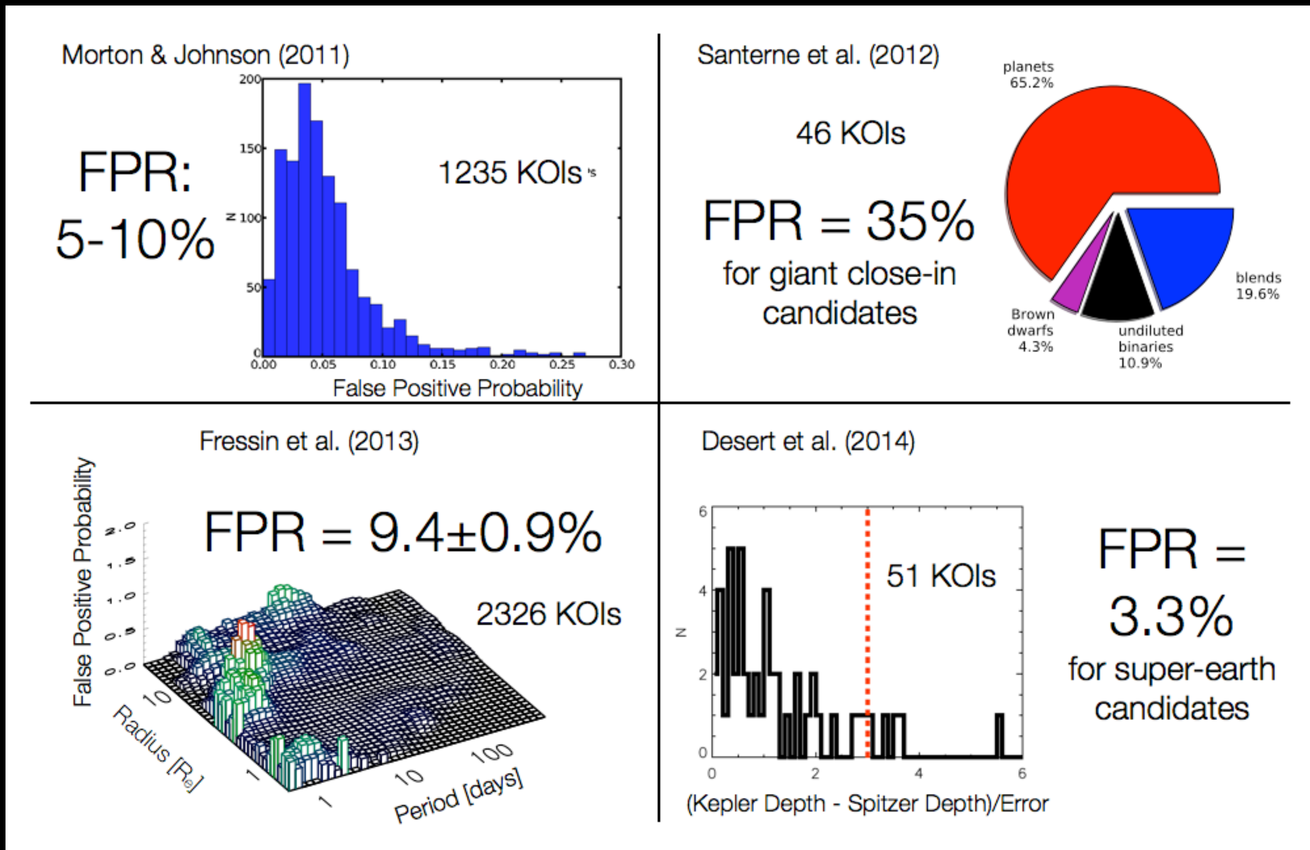
Pink: gaus
function




Seader et
206, 25

# Knowledge of catalog reliability

Globally averaged, the false positive rate (FPR) is less than 15% but is highly dependent on planet properties.




Empirical measurements yield FPR=35% for the close-in giant exoplanets and < 5% for super-earths.