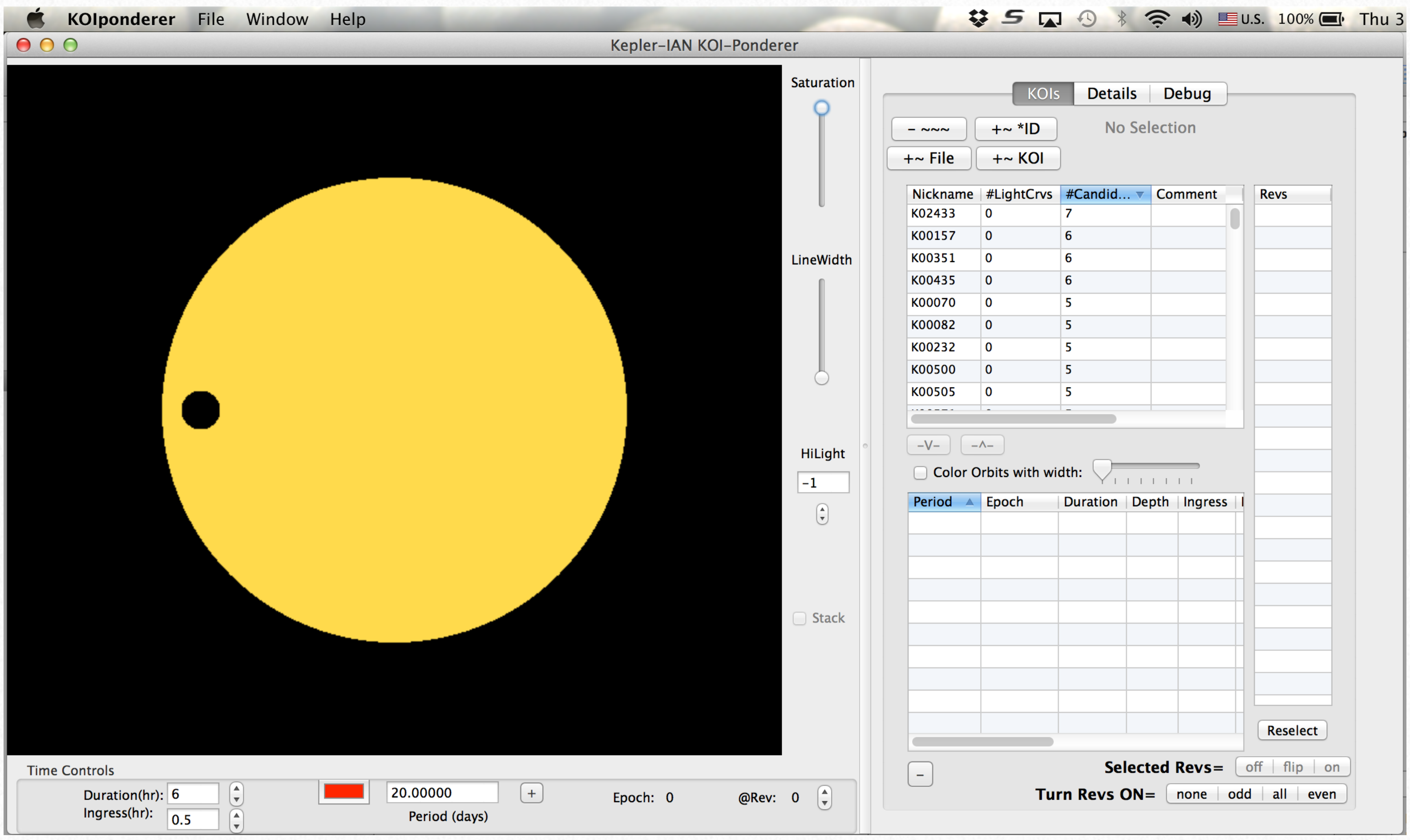
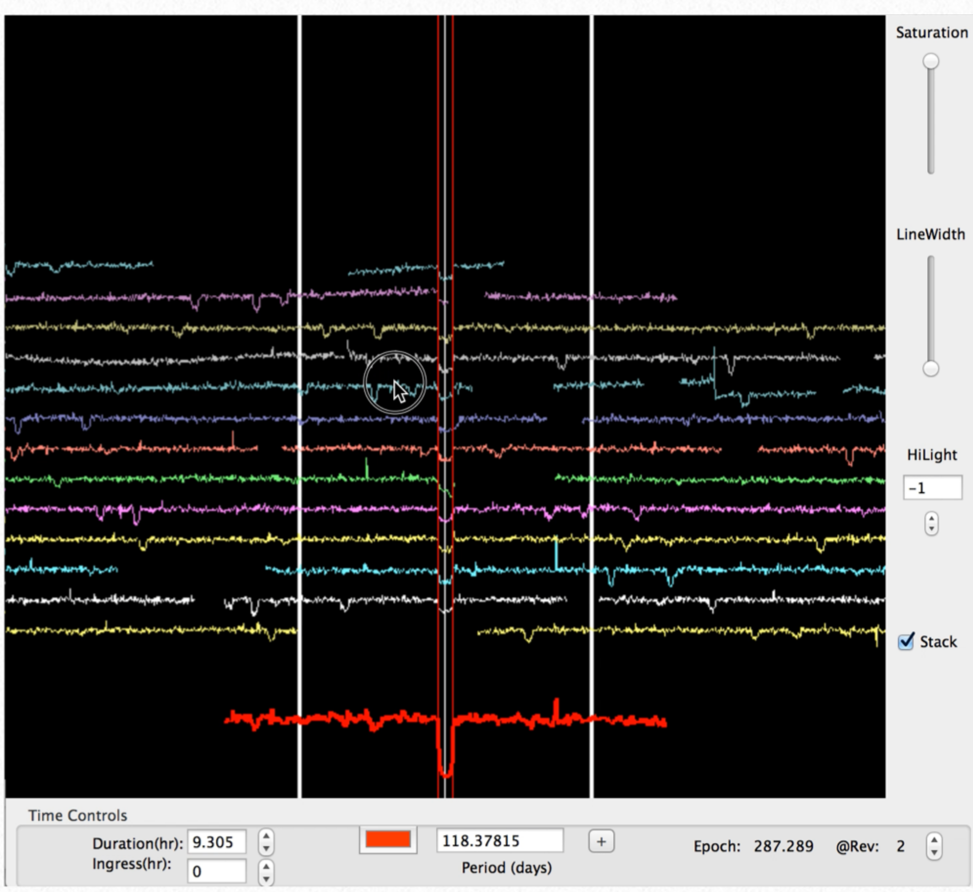
R. B. Nerf, Jr.

KOI Ponderer

Interactive exploration of Kepler Objects of Interest





Overview

> "I used to measure the skies, now I measure the shadows of Earth. Although my mind was sky-bound, the shadow of my body lies here."
>
> *--Epitaph for and by Johannes Kepler*

The Kepler spacecraft carried a telescope system designed to stare intently at over 100,000 stars, precisely recording their light, so that any eclipses caused by orbiting planets could be detected and analyzed. The processing of this data was an enormous task, requiring attention to detail, precise correction for known instrumental effects, and the search for signals obscured by a variety of nasty noises.

The data displays and diagnostic materials were useful for professionals with a detailed understanding of the system, but were overwhelming for a less-experienced person trying to get "a feel for the data".

KOI Ponderer was developed as an attempt to use hand/eye coordination to "impedance match" the Kepler light-curve data to an understanding of the exoplanet orbits derived from it. What follows here will only be as precise in its nomenclature as is needed to understand the basic items that appear in the app's interface.

The first entity to be considered is an individual star, which NASA assigns a unique 8-digit identifier (e.g. 11968463). Associated with each star are one or more "long cadence" light-curves. These curves integrate the star's light flux over half-hour intervals for up to 90 days; the quarter-year reflects spacecraft operational limitations. If analysis of a star's light-curves reveals a possible sequence of eclipses, then the star becomes a "Kepler Object of Interest" and is assigned a KOI identifier (e.g. K02433). Each potential sequence of eclipses is identified by appending a hundreths decimal to the KOI identifier, e.g. K02433.01 to K02433.07 identify seven potential exoplanets orbiting this star.

Each potential exoplanet is characterized by a suite of orbital (and stellar) parameters derived from the details of the light curve. We will not explicitly consider the stellar parameters, nor any inter-planet resonance interactions.

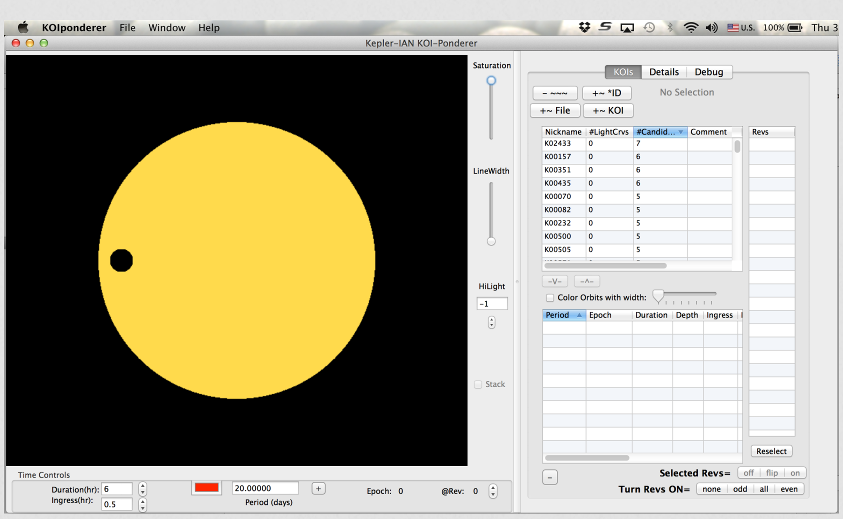
Within *KOI Ponderer*, each potential (candidate) orbit is defined by:

> *period* -- The time interval (in days) between eclipses.
>
> *epoch* -- A time (in days) when (the center) of an eclipse occurred. This parameter is not unique, since eclipses occurred both before and after the recorded epoch.
>
> *duration* -- The time in hours between the beginning and end of each eclipse.
>
> *ingress* -- The time in hours between when the eclipse begins and when the planet is fully overlapping the star's surface.
>
> *depth* -- The maximum fraction of the star's light (in parts-per-million) that the exoplanet obscures.

The *KOI Ponderer* application initially contains a database of some seven thousand candidate eclipse parameter records. Each star that shows eclipses is also represented in the database, but, to save disk space, the light-curves associated with all the stars are not immediately available. The expected work flow starts with selection of one of the stars, downloading the associated light-curves, and displaying those light-curves aligned according to the orbital parameters of one of its candidate eclipse sequences. At this point, various orbital and display parameters can be interactively modified to provide insight into the analysis, particularly with respect to the effect of phenomena that are not incorporated into the design of the analysis software.

In the following chapters, the workflows for using *KOI Ponderer* will be discussed in increasing detail. Appendices will describe under-the-hood minutia of the implementation.

Figure Overview.1 *KOI Ponderer* launch window.

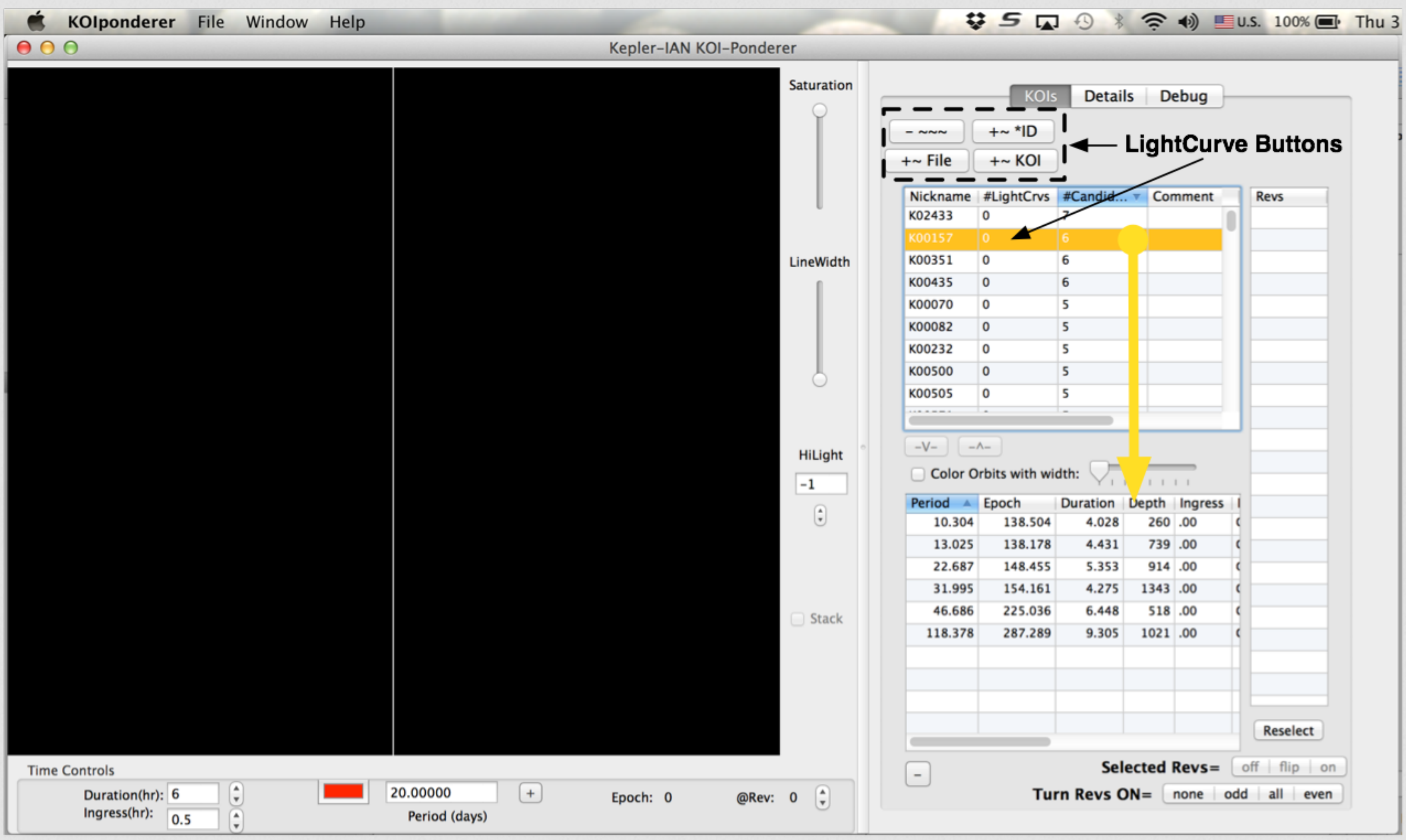


1

Basic Workflow

"Nature uses as little as possible of anything."
— *Johannes Kepler*

Figure 1.1 Effect of selecting a star.

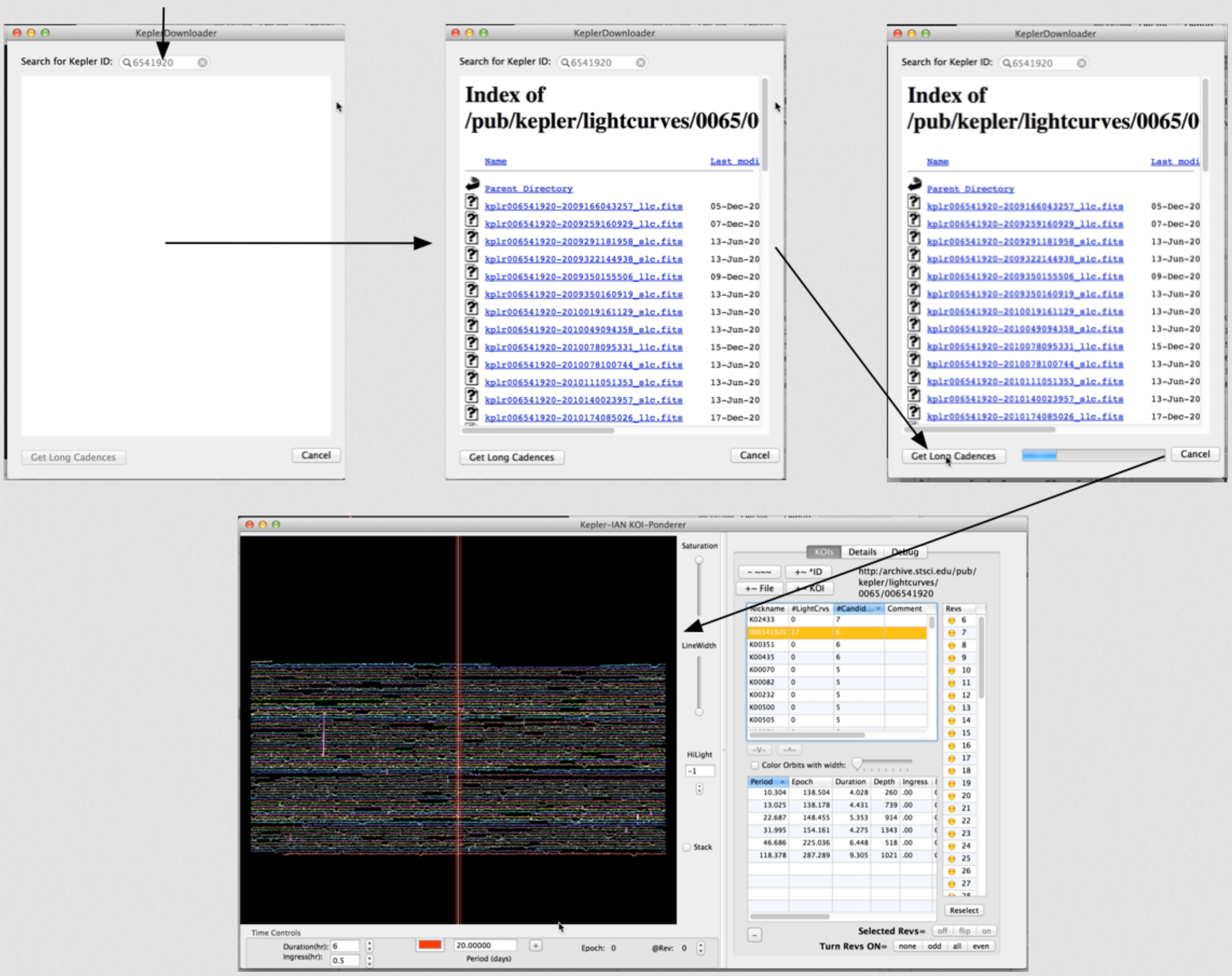


When *KOI Ponderer* first launches and no interaction has occurred, the app's window resembles the screen-shot at the bottom of the previous page. At left, an icon represents a yellow star with a transiting planet, and, at right, several tables are arranged.

The wide table at upper right displays a row for each star. If a star's row is selected (indicated by the yellow highlight), then two things can occur. First, the wide table at lower right will be populated to display the star's candidate exoplanet orbits (indicated by the yellow arrow). Second, any downloaded light-curves for that star will be displayed in the black area to the left. In this image, the "*#LightCrvs*" for the star is shown as zero, so nothing appears except a central vertical line.

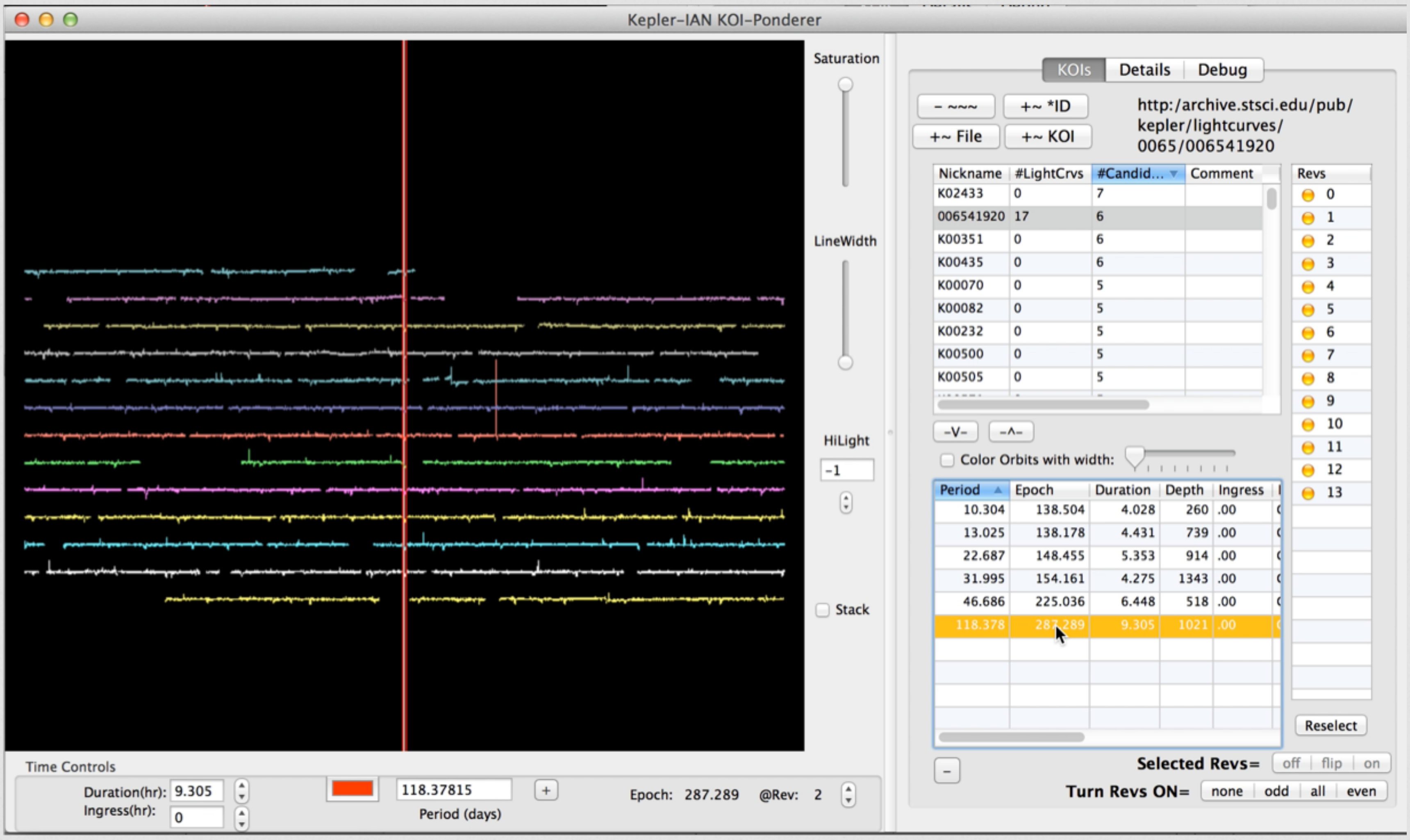
We will use one of the *LightCurve Buttons* enclosed in the dashed box to download light-curves from NASA's server. If an Internet connection is available, pressing the button [**+~ KOI**] will initiate the following sequence:

Figure 1.2 Sequence of windows while downloading light-curves.



A *WebView* window will appear, with the star's 8-digit Kepler ID in the field at the top of the window. Although no "in progress" indicator is shown, a request has been made to NASA's server. When the server responds, a list of the star's data files will appear (second

Figure 1.3 Light-curves aligned by candidate orbit.



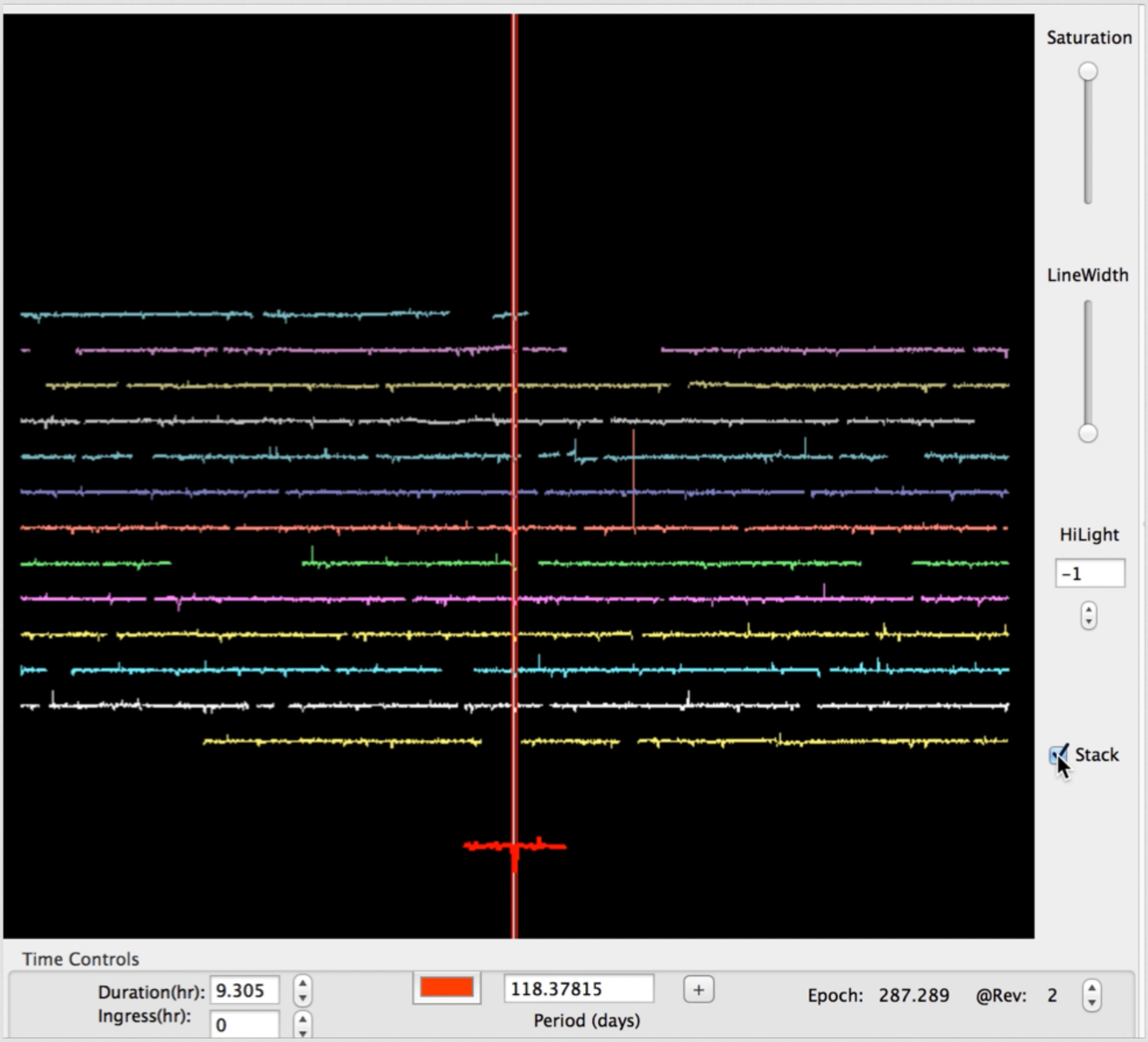
window) and the [**Get Long Cadences**] button will be un-grayed. Pressing this button brings up a thermometer (third window) and the light-curves will appear on completion.

Now that there are some light-curves in the database, selecting a candidate in the wide table at lower right (yellow highlight) will cause the light-curves to be aligned according to the orbital parameters. Each eclipse in the sequence will underly the central vertical line. The parameters are also copied into the fields below the light-curve display, where they can be interactively modified in several ways.

The screen shot below focuses on the light-curve display area; it differs from the image above only in the selection of the *"Stack"* checkbox, which causes the display of the small red curve below the light-curves, as explained below.

Each of the longer color-coded curves displays an orbital period's worth of data. They are aligned so that a vertical line will intersect data that differs in time by a multiple of the orbital period. The central vertical line marks where the *epoch* time, plus or minus multiples of the orbital period, would appear. (The central line has structure that can barely be seen at this scale. It will be explored below.)

Figure 1.4 **Light-curves and Stack**

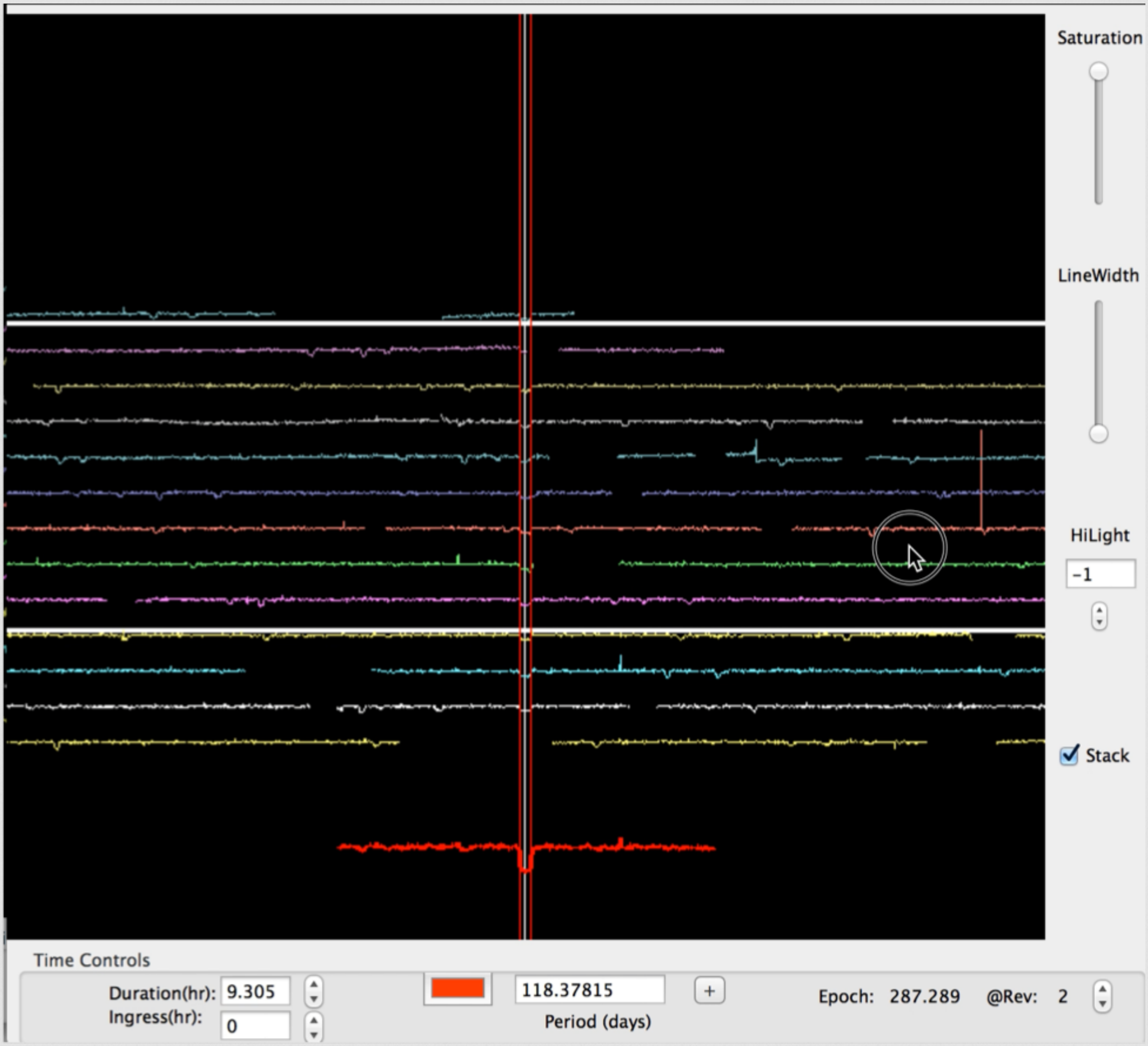


Note that the *Epoch* parameter below the light-curves display is associated with an integer following the @*Rev* label, which is followed by a stepper control to change the integer's value. This accommodates the ambiguity inherent in $T_{mid} = T_0 + n_{rev}T_{period}$. The data for revolution "@rev" is displayed in white (the light-curve second from the bottom). A second reason to call-out a particular revolution is to indicate the position about which the display will "pivot" when the eclipse period is changed.

Some interaction with the light-curves and orbital parameters can be accomplished by entering values into the fields beneath the curve display, but the preferred method is to click-and-drag within the black area of the display. There are nine parameters, orbital and display, that can be modified using a mouse-drag in concert with a modifier keypress. The details are presented in the next section.

Horizontally-oriented parameters

Figure 1.5 **Mouse drag adjusting horizontal (time) scale.**

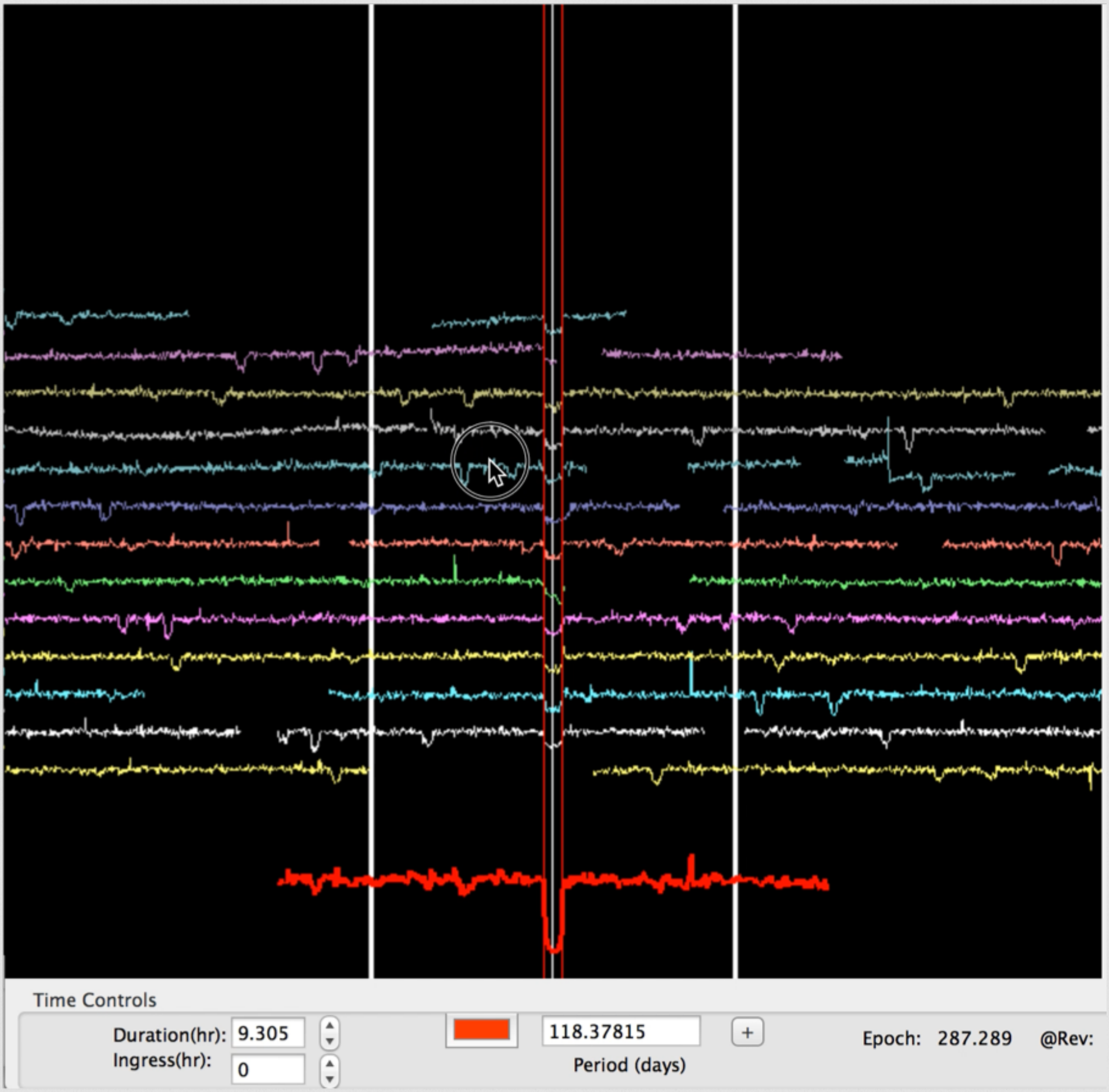


If a mouse-drag is initiated with no keypress, it is assumed to adjust a horizontal parameter; each third of the display controls a different parameter. In the display above, the middle third (as delimited by the white lines) controls the horizontal (time) scale, so one can zoom in on an eclipse sequence.

A mouse-drag in the upper third would adjust the *Period*, and in the lower third would adjust the *Epoch.*

Vertically-oriented parameters

Figure 1.6 **Shift-key + Mouse-drag adjusts vertical (flux) scale.**

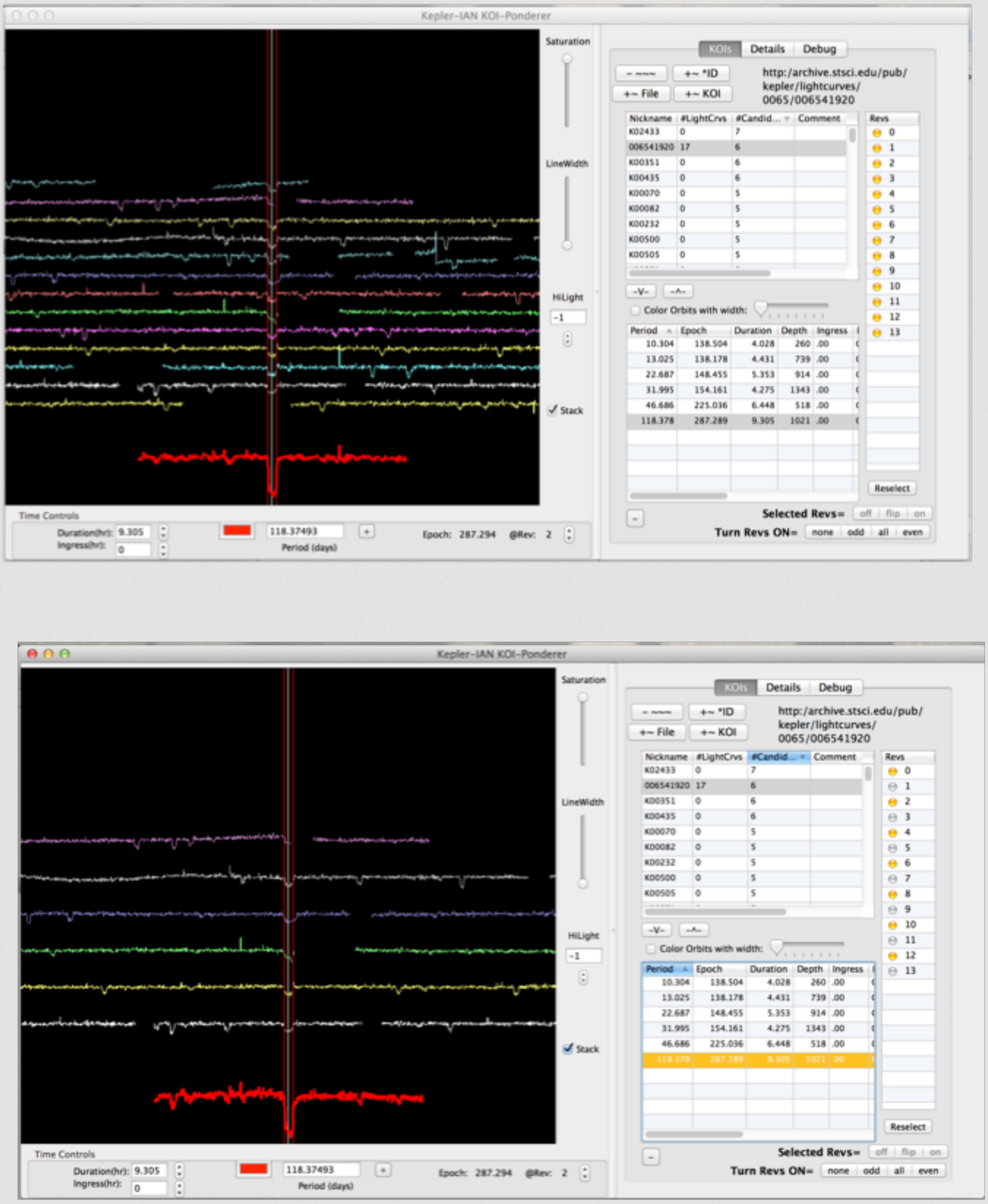


In the figure above, A vertically-oriented *shift-mouse-drag* in the middle third controls the flux amplitude scale. In the left third the drag will control the vertical location of the curves, and in the right third will control the spaces between the curves of each revolution.

Function-mouse-drag motions in the left and center thirds change the vertical scale and position of the *Stack* curve, analogous to the “all curve” behavior of *Shift-mouse-drag*. However a vertical *function-mouse-drag* in the right third of the screen will change the horizontal extent of the *Stack* curve.

Stack Curves

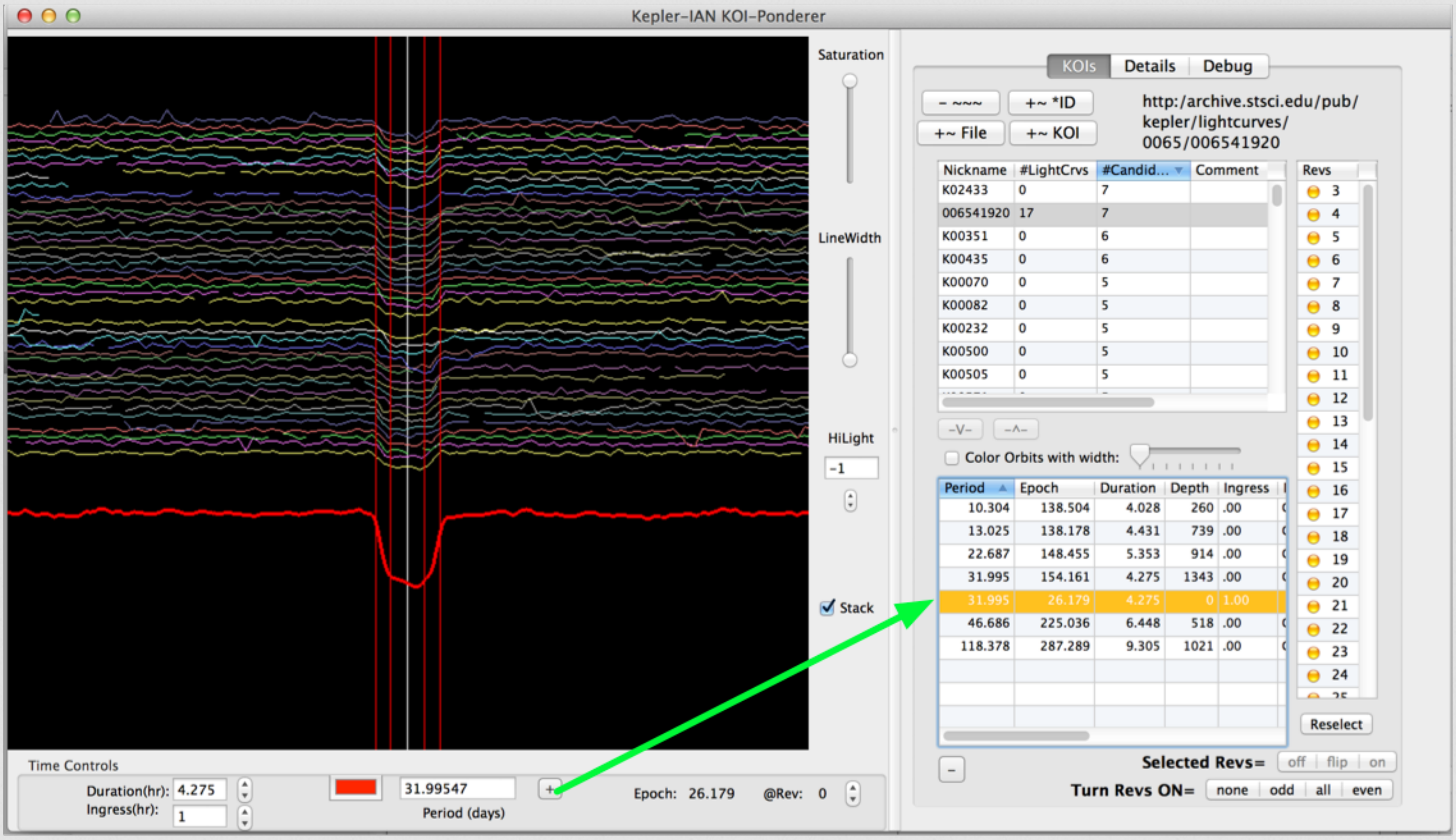
Figure 1.7 Stacks for all revolutions and even-numbered revolutions.



The *Stack Curve* (usually) represents a vertical summation of the displayed light-curves. (The exceptions involving the *Details* tab are discussed in the Appendices.) The narrow table at far-right shows which revs are included; three interactions are possible: double-clicking in the table, using the bottom row of buttons, and selecting within the table, and using the buttons in the middle row.

Picking & Revising Candidate Orbits

Figure 1.8 Adding new candidates.



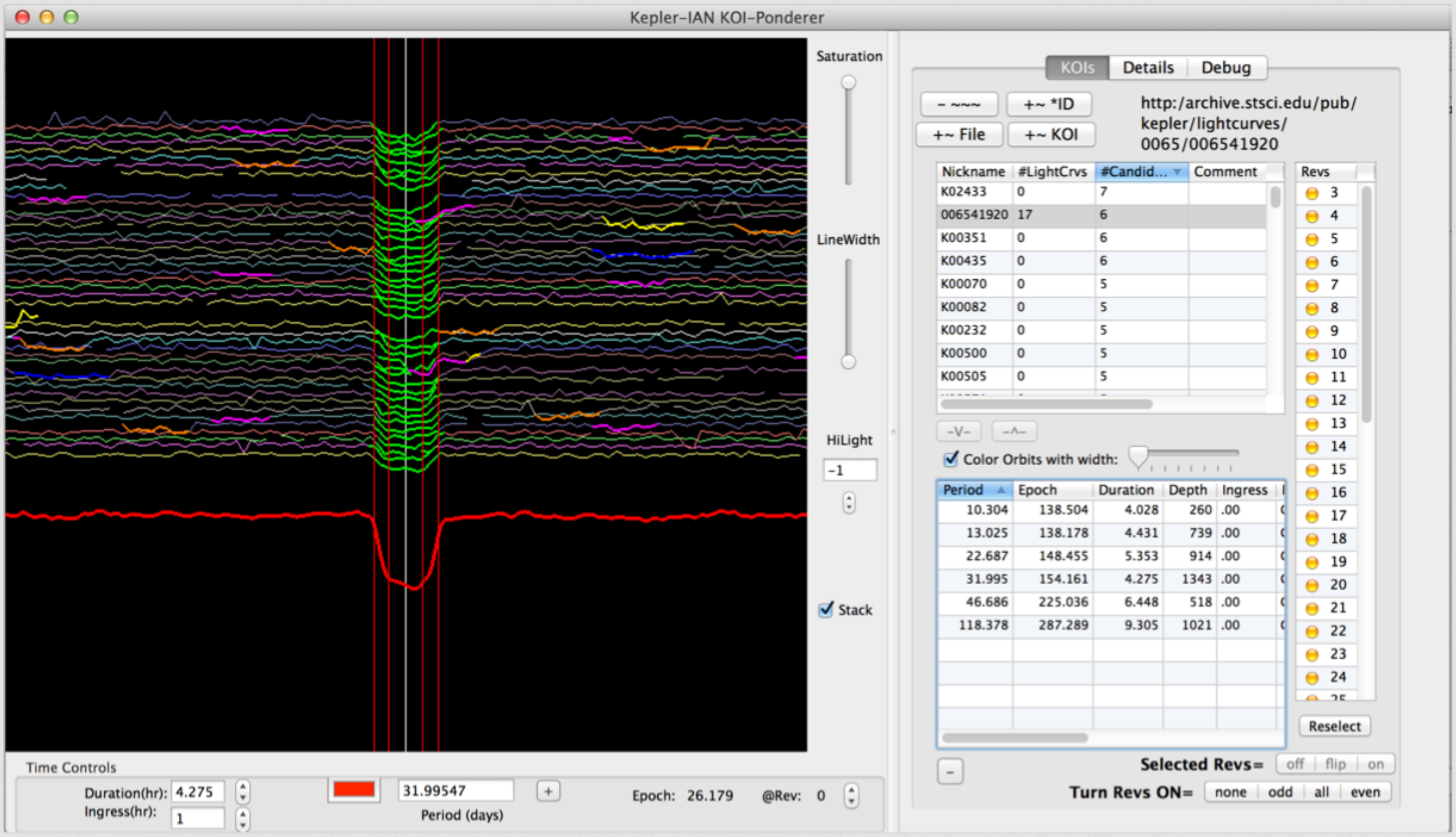
This display has been zoomed-in to show the details of the eclipse sequence. There are five vertical lines: the white central line marks the center of the eclipse, the spacing between the outer red lines marks the duration of the eclipse, and the inner red lines mark the ingress and egress of the eclipse.

The [+] button below the light-curve display causes the current orbital parameters to be entered into the candidate orbits table.

The [-] button below the candidate orbit table will remove the selected row from the database.

Bells & Whistles

Figure 1.9 Eclipse locations highlighted in color.



The relative areas allocated for the light-curve display at left and the tables at right can be modified by dragging the separator.

Selecting the "*Color* Orbits" checkbox above the candidate orbit table causes the location of each eclipse to be outlined with colors that are settable via the *fRed, fGreen, and fBlue* entries in the candidate orbit table. This display allows one to see potential interferences between various eclipse sequences.

The *Saturation* and *Linewidth* sliders in the middle affect the lines which draw the light-curves.

The *Hilight* controls are obsolete.

Changes to the star and candidate databases are not stored until either the *Save* or *Quit* commands are issued.

2

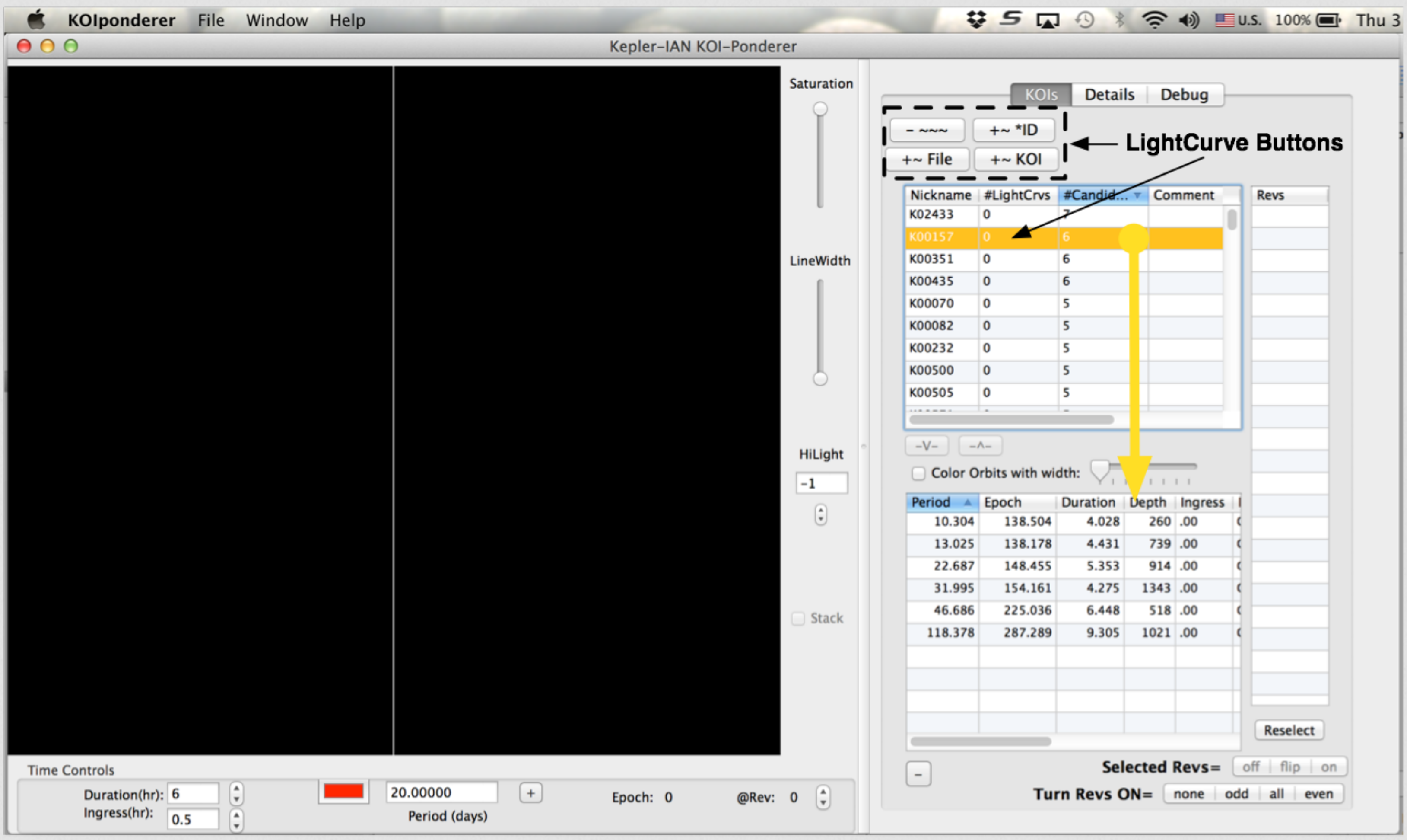
Advanced Workflows

> “I too play with symbols... but I play in such a way that I do not forget that I am playing.
> — *Johannes Kepler*

This chapter contains implementation details that may need to be considered when doing more than eyeballing the data.

Advanced Light-Curve Management

Figure 2.1 **Other ways to load and unload light-curves.**



In the Basic Workflows we loaded light-curves from NASA's server by selecting a star in the upper table and pressing the [+~ **KOI**] button (mnemonic: add curves using KOI). This automagically inserted the 8-digit Kepler ID in the retrieval dialog. If there is no KOI entry, but the star's 8-digit identifier is known, pressing the [+~ ***ID**] button (mnemonic: add curves using star ID") will bring up the dialog and allow the manual entry of the ID, followed by the "enter" key. If the star's curves are in the NASA server, the same sequence of dialog windows and button-presses will happen.

If FITS-formatted light-curve files are available within a folder on the local file system, then pressing the [+~ **File**] button (mnemonic: add curves from file system) will bring up a file-system dialog. Select the folder that contains the long-cadence light-curve files.

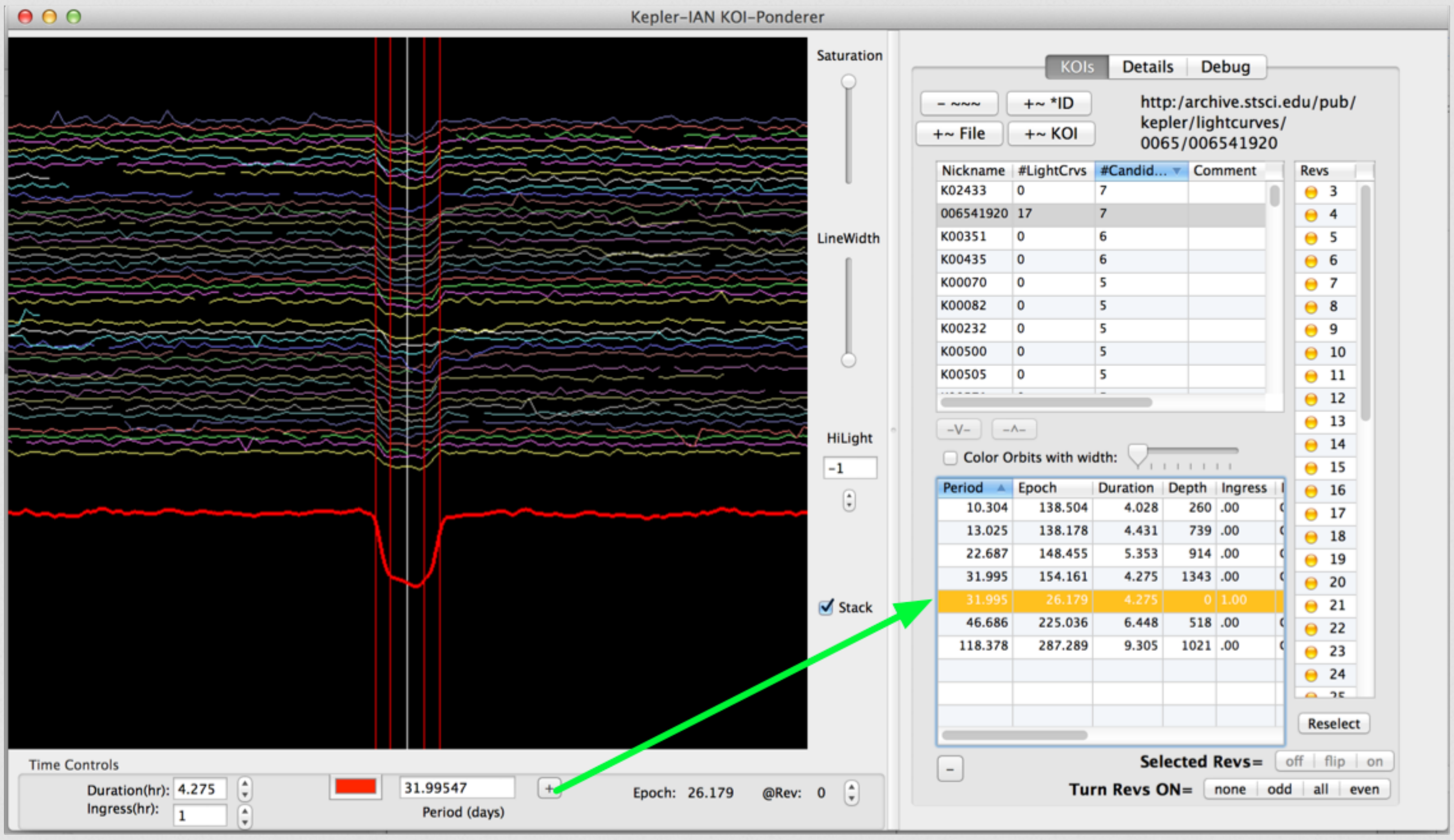
The following workflow gives a concrete example of how light-curve files might be transferred to the local file system:

> In general, how the curves arrive on the Mac is an exercise left to the student, but a brute force example can be demonstrated using the NASA server. Here by accessing a URL for a star, e.g. http://archive.stsci.edu/pub/kepler/lightcurves/0117/011709124/ the default web browser will display a list of files associated with the star. Each "long cadence" file can be downloaded by right-clicking or control-clicking on the corresponding entry and choosing "Download linked file" from the contextual menu.
>
> Collect all the long cadence files into a folder -- the name of that folder will be used as the "Nickname" entry in the stars table. This is the folder that will be selected when the [**+~ File**] button is pressed.
>
> However there's a Murphy's Law annoyance to be overcome first. OSX adds a gratuitous *.txt* extension to the downloaded files, so use *Terminal* to "cd" to the nickname folder and run the following script to strip the *.txt*:
>
> **for f in `ls *`; do mv $f `basename -s .txt $f` ; done;**

The [**- ~~~**] button (mnemonic: remove curves) will delete light-curves associated with the selected star. However, the freed disk space will only become available when *sqlite* memory management frees the corresponding memory pages.

Fake Data (useful & dangerous)

Figure 2.2 Using Fake Data.



During quality-assurance testing, the ability to inject synthetic data into *KOI Ponderer* was implemented to make sure that the various sub-systems were in sync. That functionality is available from the [~**v**~] and [~**^**~] buttons which are located between the Star and Candidate Orbit tables in the right half of the window. (The **v** & **^** characters show which way the synthetic eclipse will be oriented in the curve.)

Pressing either of these buttons, respectively adds or subtracts a synthetic eclipse sequence (corresponding to the selected candidate orbit) to the light-curves. This can give an indication of how well the orbital parameters represent the actual data. The first danger is that this process modifies the light-curve data, so that if the operation is not reversed, the light-curves in the database will be corrupted.

A second danger results from the synthetic data being a point calculation, not including the effects of the half-hour integration that original Kepler data represents.

A third danger lies in possible oversimplification. The synthetic data assumes a very large star radius, and no limb-darkening.

Stack Calculation

The *Stack* curve is calculated, not from the original light-curves, but from a more-finely resampled representation of that data. For ease-of-comparison with the light-curves displayed in the interface, the resampled curves use a linear interpolation between data points. Mathematically, this may not be the most accurate resampling method and it ignores any difficulties associated with gaps in the data sequence.

3

Appendices

“I much prefer the sharpest criticism of a single intelligent man to the thoughtless approval of the masses.”
— *Johannes Kepler*

System & Hardware Compatibility

The app has been developed and tested using:

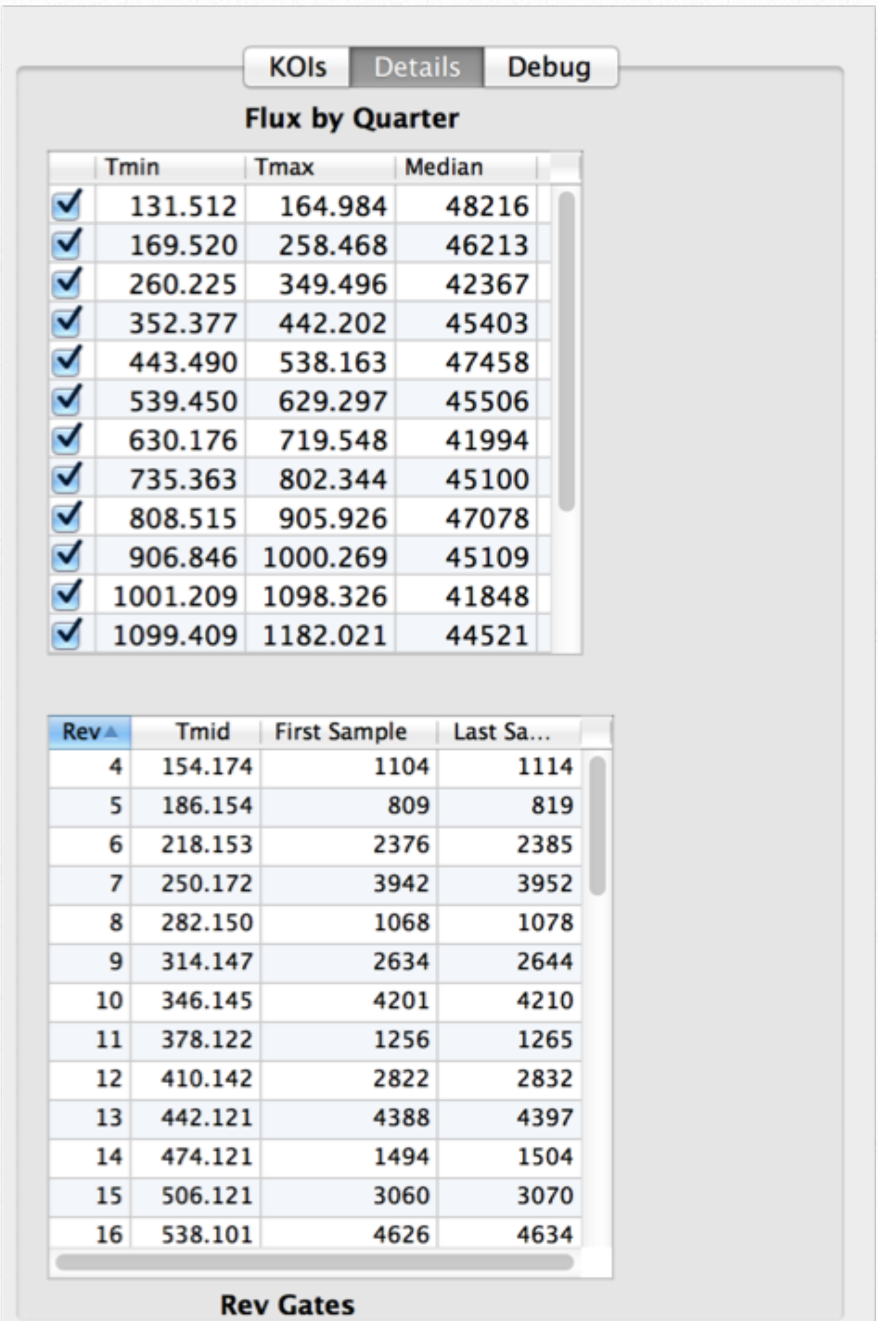
- OS X 10.9.4
- Mac Book Pro - Retina, 13-inch, Late 2013.
- iMac - 21.5-inch, Mid 2010

Bugs, Missing Features & Questions:

- Empty folders appear in ~/Downloads for each star downloaded. They can be deleted.
- Glitch when first adjusting stack-width using *function-mouse-drag* in right third of curve-display window.
- Stack-curve display sometimes disappears when new flux curves are downloaded. Triggering a redisplay, e.g. by changing window size will make it reappear.
- Need some ability to do arithmetic on the *Period* to make it easier to look for aliased picks.
- No in-app documentation of mouse-gestures for changing display and orbital parameters

- Epoch and flux can be oddly linked to period if a large change in period is typed into the period field. One method of recovery when this problem is detected is to use [**+**] button to pick a set of orbital parameters, and to edit the *Epoch* field to 0.0.

Details



Tmin	Tmax	Median
131.512	164.984	48216
169.520	258.468	46213
260.225	349.496	42367
352.377	442.202	45403
443.490	538.163	47458
539.450	629.297	45506
630.176	719.548	41994
735.363	802.344	45100
808.515	905.926	47078
906.846	1000.269	45109
1001.209	1098.326	41848
1099.409	1182.021	44521

Rev	Tmid	First Sample	Last Sa...
4	154.174	1104	1114
5	186.154	809	819
6	218.153	2376	2385
7	250.172	3942	3952
8	282.150	1068	1078
9	314.147	2634	2644
10	346.145	4201	4210
11	378.122	1256	1265
12	410.142	2822	2832
13	442.121	4388	4397
14	474.121	1494	1504
15	506.121	3060	3070
16	538.101	4626	4634

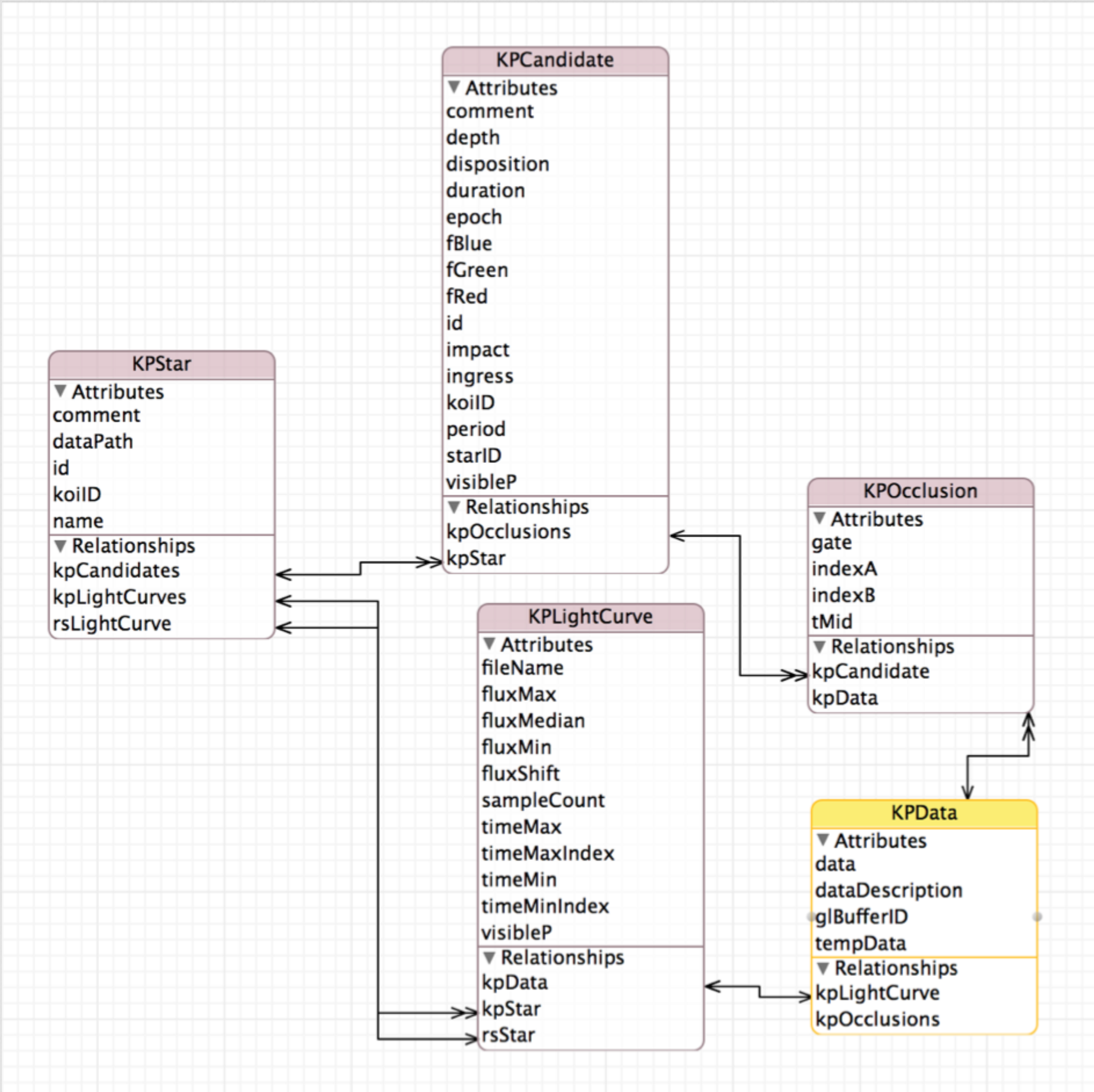
The tables shown in the *Details* tab contain potentially-useful information for app debugging, but, in one case, might be misleading within the *Stack* curve calculations.

The upper table displays minimum and maximum times, together with median flux values for each quarter's data. Toggling the associated check box will cause that quarter's data to disappear from the display; however, unlike the *Rev* table, there is <u>no</u> effect on the *Stack* curve.

The lower table displays information about the *Rev* gates, relative to each quarter's data.

Data Model

Figure 3.1 **Core-Data Model**



The diagram shows the Core Data model, which is implemented as an sqlite database at ~/Application Support/com.gnerph.KOIponderer/. Initially only KPStar and KPCandidate entities are stored, but as light-curves are downloaded, the KPLightCurve and associated KPData entities are populated, along with KPOcclusions, which correspond to the color highlights in the plots for each eclipse sequence.