Reviewer #1: The manuscript by Berginski et al presents results obtained by automated analysis of TIRF images to obtain information on focal adhesion dynamics. The method appears interesting and useful and the data generally seem to be of good quality (although see concerns below on cell and experiment numbers). Certainly this type of analysis is of interest for many screening applications.

Aside from the points outlined below, I have a general thought that authors might do a better job of pitching their data by emphasising the methodological aspects and de-emphasising the biological story. It strikes me as a methods paper at heart, but the main text is mostly concerned with the output of the analysis rather than how the analysis was performed. Although well within the remit of PLoS one, the biological results are rather weak on their own; whats more interesting is how they were obtained. Authors might bring some aspects out of the methods section into the main results. Likewise, some of the (13!) supplementary figures might be presented as figures, even if at the expense of some main figures.

We appreciate these comments and have tried to place greater emphasis on the methodological aspects within the paper. This has mainly been done through editing of the text and changing the emphasis in areas of the introduction as well as results. We feel that these changes address these comments and place emphasis on the methodology while still using the biological studies (e.g. the S178A experiments) as a proof-of-concept example of how the methodology can be used. We also have made it clear within the text that the software for this approach is provided under an open-source license freely available on the Gomez lab website.

MAJOR POINTS

1. Total number of adhesions analysed is rather abstract without knowing the number of cells analysed. Why is a value for number of adhesions per cell not presented? see related points below

There are 21 EGFP-Paxillin, 9 EGFP-PaxillinS178A and 10 EGFP-FAK time-lapse image sets included in the data sets analyzed. Each image set typically contains a single cell (but see below). The number of movies is lower in the FAK and S178A data sets because they were gathered for a longer time and thus produced more analyzable adhesion assembly and disassembly events. These cell movie counts have been made more clear in the captions of figures 2, 4 and 5 and the end of the "Microscopy" section of the methods.

We do not present the number of adhesions per cell as it is a value that changes continually throughout the experiment's duration. We have added the average number of adhesions in each cell in "FA Tracking" section of the methods.

2. P7, methods, Fig2D: is there a lower size limit for static identification of FAs?

We did not set a limit for lower size of FA identification.

The dynamic analysis is good because it uses behaviour to classify FAs. But the text mentions that fewer dynamic FAs are identified than static FAs. Which static FAs don't stand up to dynamic analysis, is it mostly the smaller static contacts ?

This may have been confusing. In this particular study, we quantified the dynamics of long-lived adhesions only. This was primarily due to the desire to look at adhesions where we could see their entire life cycle (birth, maturity, death). Thus we required that adhesions must be present in the data set for at least 20 minutes for them to be considered for dynamics analysis (although it should be noted that this duration is a user adjustable parameter). Adhesions that are present for shorter time periods are often smaller adhesions. For these, we do characterize their static properties. We have tried to make this clearer in the text.

What percentage of statically identified FAs are dynamic?

For this work, we have been quite restrictive in looking only at a class of long-lived adhesions (> 20 minutes), where assembly and disassembly rates could be accurately calculated. This condition does exclude the majority of adhesions from our dynamics analysis. In the EGFP-Paxillin data, 1.5% of adhesions have assembly rates calculated, while 1.8% have disassembly rates calculated. The assembly percentages for the EGFP-FAK and the S178A mutation are 2.7% and 2.2%, while the disassembly percentages are 2.4% and 2.0% for EGFP-FAK and S178A, respectively. We attribute the percentage increases in the number of adhesions that could be analyzed to the longer FAK and S178A time-lapse movies.

How do properties of size, position, and intensity compare for FAs identified by static and dynamic analysis? The underlying concern is how to be sure things identified by static analysis are really contacts. P7L7 claims "The segmentation methods successfully identify the adhesions in each image regardless of the background Paxillin fluorescence intensity." But there's no gold standard for comparison. With the dynamic analysis you can be pretty sure that any blob which behaves a certain way is probably an adhesion, but with static analysis, how do you know a blob is an FA and not noise, an aggregate, an endosome, or a bit of autofluorescence? Authors seem to imply that everything labeled with paxillin which shows up in TIRF is an FA.

We are operating under the assumption that after background correction, high-frequency noise filtering and intensity thresholding, all of the discrete, Paxillin- or FAK-labeled structures contained in a TIRF image are focal adhesions. The use of these threshold and filtering steps help minimize background fluorescence and noise and should limit false detection from any of the above sources of noise. We also note that our criteria for deciding whether something is an adhesion or not is at least as rigorous as current standardized and published manual approaches. The detection algorithms are similar to those from a prior publication (Zamir 1999 reference in paper) and the parameters of the segmentation method can be modified. Whether a small number of these structures are something other than focal adhesions is another matter, but we feel that our initial assumption is correct for the purposes of this manuscript. The automated methods of adhesion detection at least guarantees that human bias is not playing a role in determining what is and is not an adhesion. A description of data preprocessing is in Methods. We also note that, as part of our testing, multiple different simulation studies were performed where we have a "known-truth" for comparison. Also, it is also important to note that just because an adhesion does not meet our strict criteria for automated dynamics analysis does not mean that it is not dynamic. It only means that there is not a sufficient number of time points sampled from its life-cycle to reliably calculate rates and other dynamic properties.

3. How many cells have been analysed of each type, and how many experiments have been performed? The robustness of analysis is an important issue. Although figures list number of adhesions analysed, its not clear how many cells or experimental replicates this relates to. Likewise, P9L16 states 10 movies were analysed, but how many cells and experimental replicates does this represent?

For the P9L16 reference, this represents 10 cells. We have updated the Methods and Results sections to include the number of cells analyzed and experimental replicates performed.

P8L15 refers to "all adhesions identified by our software (n = 21 cells)" -is that for the whole manuscript or just Paxillin-GFP in 3T3s?

This part of the paper is only referring to the EGFP-Paxillin cells. This has been clarified in the text of the paper.

If this is meant to be a robust analysis method for high content screening, I would like to see that the same cells analysed in separate experiments give similar results.

We feel the results are quite robust. In fact, the results presented encompass experiments that were conducted over the course of a year and were even performed on different microscopes, while still giving consistent results. We view the analyses presented in this paper a proof of concept for applying this software tool to larger scale experiments and downstream high-content screening.

4. Continuing with robustness, how sensitive is the analysis to the TIRF illumination angle? People often set the illumination angle to a value which looks good by eye, but what is the effect on the analysis if illumination angle is a little bit different from one experiment to the next?

As long as the TIRF angle is set so that the signal to noise ratio in the movie is high enough to reliably detect the adhesion at the time of formation, the methods are robust to the specific angle selected. However, as with any quantitative study using TIRF (or other microscopy system), it is essential that the illumination angle is kept constant when making direct comparisons as, for example, we did with cells expressing WT and S178A Paxillin.

5. Finally, with regard to robustness, how constant is the "empirically determined threshold value" referred to on P17L4 across different experiments, probes, and cell-types. Does it have to be tweaked for every set of experimental conditions? Are there any other parameters which need to be manually determined?

We used the same threshold for the adhesion segmentation for the analysis of all the data presented in the paper. In addition, we investigated the effect of varying this parameter on the differences in assembly and disassembly rate detected in the S178A mutant (Figures S6-7) and found that the differences in rates were robust to the specific level of this parameter.

6. P8L20: ?"estimated from far fewer measurements?" I think this should be expanded. This is essentially a methods paper, so the authors should play to their strength here. What are the possible sources of human bias? The implicit hypothesis is that larger/brighter FAs are more likely to be manually selected for analysis and this will skew the results. Authors could test this idea by comparing analysis of larger and/or brighter FAs with dimmer and/or smaller FAs. Do the results cluster with FA size and/or intensity? Authors could really prove the point that human bias is likely to skew results. -Or maybe big bright and small dim contacts behave similarly, in which case no bias.

You are correct that we implicitly hypothesize that large, bright adhesions will be those that are readily identified with manual methods and thus any relationship between those properties and the determined assembly or disassembly rates will be similarly biased. We empirically determined for instance, that the Paxillin-labeled adhesions of Webb et al, 2004 appear to be sampled from the large/bright/high-rate subpopulation of adhesions found within a cell, which correspond to the upper quantile of adhesion assembly/disassembly rates that we characterized globally. We found, but did not include in the paper, that average adhesion size and intensity explains between 14%-42% of the variance in assembly or disassembly rate. However, describing and quantifying the potential biases in the process of adhesion image analysis is in and of itself a significant project worthy of a dedicated publication. Instead of emphasizing these analyses, in this work we chose to focus on describing the methodology along with proof-of-concept examples of its capabilities.

7. There are two features of interest which authors appear to have access to in their data but have not presented: sliding velocity of FAs and the sliding path length. This information is different than, but complimentary to the zones of assembly and disassembly (P9L5-9). Also, it would be very interesting to know how sliding relates to FA size, intensity, and assembly/disassembly kinetics. Do bigger adhesions slide faster or slower? If you want to present an analysis method for high-content screening, it should be able to extract as much content as possible from the data. Sliding velocity and path length seem like necessary basic analysis outputs for focal adhesions.

We agree that analyzing the sliding behavior of adhesions is of interest, especially in cells directionally migrating or being subjected to shear stress. While our software system does provide a nice base from which to work towards adding sliding parameters, there are other considerations that make this type of analysis non-trivial. For instance, deciding whether or not change in adhesion position and/or morphology is due to sliding vs assembly vs disassembly, or some combination of the three processes is difficult. We are considering techniques such as optical flow for the study of such aspects of adhesion behavior, but again, the quantification and analysis of such behaviors is more complicated than it may first appear. Our focus for this paper was on the global characterization of static and dynamic adhesion properties of wide interest to the FA community. We have modified the future developments paragraph in the "Discussion" to reflect future plans to add adhesion sliding measurements.

MINOR POINTS

P6L5 says 1000 - 10000 adhesions per cell were analysed, whereas P7L15 says over 44,000 adhesions were analysed, and neither agrees well with numbers of adhesions presented in histograms. see related major point above.

See comments to point 1 above.

P6L11: "to" missing

Corrected in revised manuscript.

Fig 2D: why has this histogram bin size been selected? virtually all the data points are in the first bin. not very informative.

The average adhesion size clusters at the bottom of the size scale, various bin sizes were tested and all produce similar appearing histograms.

Fig3B: why has this been normalised such that maximum intensities occur around 0.55?

The normalization is based on the maximum intensity detected in the set of images analyzed, not for individual adhesions.

P8L5: "? over 50% of all adhesions analysed (n = 1183 for assembly and n = 1487 for disassembly)?" is the point of this to say that half of the curves of each type had R-squared values of 0.79 or better? "over 50%" seems ambiguous, why not give the actual value?

We agree that this sentence from the "Kinetics of FA Assembly and Disassembly" section was confusing in the original paper. We have simplified it in the revised manuscript.

P8L8: "?relative to the statics properties of figure 2?" -i think i know what the authors mean here, but not sure why they say "statics properties".

We agree, the sentence has been revised in the manuscript to make it clear that we were referring to the set of adhesions that produced the data in figure 2.

P9L19: not sure if it is reassuring or alarming that paxillin and FAK have "identical" kinetics in different cell types. would you expect them to?

Webb et. al, in Nature Cell Biology 2004, determined assembly and disassembly rates for fluorescently labeled FAK, Paxillin and Zyxin. FAK and Paxilin had statistically indistinguishable

rates in their experimental setup, while Zyxin differed in the assembly rates. Thus, at this point in time we do find it reassuring that they are similar. The similarity between FAK and Paxillin kinetics may be due to interactions at the focal adhesion that dictates the ratios of FAK and Paxillin recruitment, but we feel that determining the exact nature of this relationship is beyond the scope of this paper.

P10L7: Not clear which cells and probes have been used in Paxillin mutant experiments. Authors start with Paxillin analysis in 3T3s, then switch to FAK analysis in MEFs, then Figure 7 and related text talk about adhesion assembly without stating which type of cell and probe were used to asses it. Also not clear which cells used in Figure 8.

Thank you for catching this ambiguity. The same type of cell was used for all experiments, and it was not clearly given that 3T3s and MEFs are the same cell type. We have standardized the cell line name throughout the paper as NIH 3T3s.

P10L13: "?known effects (of S178A mutation) on cell motility?" -there is loose association made here and elsewhere between the effects of this paxillin mutant on FA dynamics and on cell migration. see also P10L19 and P3L17. P12L1 even talks about a mechanism. authors should present a coherent argument somewhere about how their data on mutants tie in with mutant cell migration. ...or maybe emphasise the methodological aspects of their work and drop the argument altogether?

We have tried to deemphasize the proposed connection between the S178A mutant and FA assembly and disassembly rates.

P16L13: missing "and"

Fixed.

do authors think this approach could be adapted for flexible substrates? could it be done in epi?

Yes, we think these analysis methods could be adapted to flexible substrates, as long as images of focal adhesions have a high enough signal to noise ratio. The methods presented are potentially applicable to examining focal adhesions imaged using epifluorescence. TIRF illumination raises confidence that the objects detected are near the surface of the substrate and thus represent adhesive structures. Under epi illumination, there are likely to be several classes of organelles and intra-cellular vesicles that would contaminate the analysis. Adapting the analysis system would require algorithmic developments to classify and remove those contaminating objects.

Reviewer #3: Comments on PONE-D-11-05518

1. I think authors should apply their analysis techniques to imaging a tagged integrin. Although Paxillin and FAK are both good FA markers, they are not integral membrane proteins.

Paxillin is present in focal contacts at all stages of their life cycle and is one of the first molecules to enter the adhesion complex following integrin activation (Zaidel-Bar et al 2003, also reviewed in Zaidel-Bar 2004). FAK is also present in nascent focal contacts (Choi 2008) with Paxillin, as well as other adhesion components including Vinculin. While GFP-Beta-3 integrin has been used as a marker of adhesions, Paxillin along with Vinculin, FAK and Talin are generally regarded as reliable markers for adhesions (Adams MC et al, 2004; Webb DJ et al, 2004; Kanchanawong P et al, 2010).

2. Page 3: Not sure what is meant by "system perturbations"

We have tried to clarify this in the "Author Summary" section where the idea of system perturbations is first mentioned.

3. Page 4: Please do not say that FAs are "born" (and throughout)

We have discussed this work with several colleagues who study adhesions and/or cell motility and they have been in favor of using birth and death to describe the beginning and end of a single adhesion's presence in a time-lapse image set. We are open to other ways of describing this process if specific suggestions could be provided.

4. TIRF should probably be better introduced and referenced

We have provided provided a relevant reference in the Introduction.

5. Page 5: Please don't say "birth, lifetime and death" (and throughout)

See comments to point 3.

6. Page 6: "to allow the community TO use our methods"

Thank you for catching this error, it has been fixed in the revised paper.

7. mRFP-tagged Paxillin has been used by several groups, could two colour imaging of FAK and Paxillin be employed and analysed by this system to allow for differential time-course analysis?

Yes, it can be used. However, when we attempted to express FAK and Paxillin constructs simultaneously within our particular system, we encountered cell viability issues. We are looking into alternate methods to enable two color imaging as this is decidedly an approach that we wish to include in future research.

8. It needs to be clearly stated what exposure times were used and the time between subsequent images.

The "Microscopy" sub-section of the methods has been changed reflect that the images were gathered once every minute for all experiments.

How do the authors know that the automated tracking is not combining two FAs into one, both in terms of effective lateral resolution, and temporally, i.e. that one FA is not disappearing just as another FA nearby is forming and the computer is not putting them together into a single trace?

Our methods of segmenting adhesions are fundamentally limited by the resolution of the microscopes used to gather the images. Spatially, if adhesive structures are so close (and generally small) that no difference in pixel intensity is sufficient to separate them via segmentation, then they will be identified as a single adhesion. If such adjacent adhesions eventually grow to become larger independent objects, the software system will automatically split them into separate adhesions. Note that the segmentation of large adjacent adhesive structures uses a modified watershed segmentation based on the work of Zamir, et. al 1999. As detailed in the "Image Processing" section of the methods. Also note that this potential problem would happen even with manual analysis as, at some point, a decision must be made as to whether or not a particular object is actually composed of 2 smaller ones. In general, manual inspection indicates that such events are relatively rare.

Temporally, if there are adhesions that form and then disassemble entirely in less then 1 minute (the sampling rate used in this work), we are unlikely to see them (or have only a single image of

their appearance). This can be compensated for by decreasing the time between image samples. In theory, it is possible for an adhesion to disappear while a new adhesion is forming in the near vicinity. In general, this does not happen, as adhesions tend to form in protrusions and thus groups of adhesions tend to born spatially and temporally close to each other in this protrusions. Retractions generally show correlated disassembly.

While extremely unlikely, if 1) an adhesion did begin forming next to a disassembling one, 2) it perfectly coincided in time such that they were never together within an image at the same time, 3) they were extremely close to each other and 4) we still mis-identified them as being the same object, the output data for these adhesions would look extremely different from the norm as we would see two sets of assembly, mature and disassembly phases connected to each other. Our algorithms identify specific "normal" behaviors and thus we are readily able to detect when unusual events occur so that we can investigate these in greater detail manually. We have yet to see such a process occur.

9. What are the kinetics for the ~50% that are not "log linear"?

We apologize for the confusion as the section of the text that describes this was confusing. The vast majority of adhesions do actually follow log-linear models. The adhesions that don't follow these models well appear to either assemble or disassemble so quickly that we are unable to accurately assess their properties (and thus they cannot be placed into any category, including log-linear). As the methods described are agnostic to the rate of image sampling, future studies taken at higher sampling rates will be able quantify the kinetics of those that we cannot currently characterize. In addition, since the software is being released under an open source license, other researchers will be able to build on our work by adding new adhesion assembly or disassembly models. We have tried to make this more clear within the text.

10. Basically the latter part of this manuscript (especially many of the supplemental figures) is full of data that is statistically significant, but of very small magnitude differences. This begs the questions which of the observations are really biologically relevant?

Our primary goals with this paper were to describe a set of methods for automating the analysis of focal adhesion dynamics and showing a case study for how the software might be used. With large numbers of measurements and appropriate statistical methods, it is possible to discern significant, but potentially very subtle differences in variables under different conditions. However, in many cases, the underlying biological significance of such differences is not obvious. Understanding whether or not such differences are biologically meaningful is obviously important and future work is being directed at such questions.

As stated earlier it might be good to directly compare FAK and Paxillin in the same cells. How about the WT and Mutant Paxillins?

We agree that comparing the markers in the same cells would be an excellent experiment, see comments in point 7.

11. The is an inappropriate line break at the top of Page 10

We can't see any inappropriate page breaks, on Page 10, but we did double check that the section breaks are in the appropriate locations.

12. Page 13: Should the new paper from Lippincott-Schwartz in Nat Cell Biol be referenced?

This paper came out after the time of our submission (April, 2011), but we agree that it is relevant. We have included a reference in the revised paper.

13. Page 15: "penicillin-streptomycin", not "penicillinstreptomycin"

This has been corrected in the submitted revision.

14. Page 16: "(epifluorescence excitation)"

This has been corrected in the submitted revision.

15. It would be nice to see this applied to a perturbation in real time within a single set of cells, for example addition of the Src kinase inhibitor Dasatinib should cause very interesting effects that could potentially be detailed easily

Real-time perturbations are definitely an area of interest as are many other areas of research that we would like to pursue. Photactivatable proteins being developed in the Hahn lab are of great interest as they can also be used as real-time, localized perturbations (e.g. Karginov et. al., J Am Chem Soc, 2010). In this work we are focusing on establishing an automated computational approach that allows focal adhesion dynamics to be quantified. Future work will pursue more detailed study of changes in adhesion dynamics in response to various chemical and/or genetic perturbations.

16. Should the automated analysis be compared to manual to make sure it is working properly?

The output has been manually inspected to ensure that it is producing consistent and reliable results. In addition to other issues, manual analysis has it's own biases and will vary between users. At a minimum, a computational approaches makes the fundamental assumptions/ parameters explicit, and while certain setting may introduce their own biases, the user knows specifically what those are. As detailed primarily in the supplemental materials, we were also able to validate our system using simulation studies where the ground truth is known.

17. Should the GFP-tagged markers be tested for colocalisation with endogenous FA proteins?

This has been done for all of the constructs used in this study. Please see Huang et al. Nature 2003, Huang et al JCB 2004 and Rajfur et al 2002 NCB for colocalization of these constructs with either endogenous adhesion components or adhesions identified with Interference Reflection Microscopy.

18. Figure 3: Need to put in explanation of inset image in D

A description of the inset of Figure 3D has been added.

19. Figure 6: Graphs don't really look difference, even though statistically significant...

This is true. As discussed in point 6 above, the large number of adhesions measured allows us to identify subtle differences between adhesion populations/properties. Biological significance is, however, a more complicated aspect of the story and requires additional experiments to decide whether these statistical differences translate into meaningful biological ramifications.