Transcription results:

| | |
|---|---|
| Interviewer: 00:00 | record again. Okay. So first, just to kind of remind ourselves what happened in the workshop, so cast our mind back to March because it's been several months since then, were you able to attend both days? |
| Researcher: 00:15 | Yes, yes. |
| Interviewer: 00:15 | Okay. Great. What did you think of the workshop? |
| Researcher: 00:19 | I think, overall, it's an excellent workshop, and indeed I have attended another workshop before this one. Compared to that I think this is a little bit shorter, but much more practical, yeah, because for freshman in this field, I think it's very hard to get started. In that workshop, they helped you have all the software installed and told you how to open it, link this one, and download some packages, how to open and remove it. I think this is a very important step. Other, then, we can explore other things ourselves. Yeah. It's always very good, I think. |
| Interviewer: 01:04 | Nice. And what do you think your biggest takeaway was? |
| Researcher: 01:08 | The biggest takeaway. I think there might be several. Just discuss it in detail. Is that right? Yeah, because I think there may be two. One is I can get started using the Jupyter Notebook. Yeah. It's much easier than I thought, yeah. Another thing is the first time I learned about the GitHub, the version control system. Yeah. There were two things first to me. |
| Interviewer: 01:45 | Oh, good. Good. So we can talk about those more in just a second. Was there anything that would have made it maybe a better learning experience for you? |
| Researcher: 01:57 | I think it's more ideally they maybe have sort of-- maybe a very small project, but it may need more time [laughter] to do that. Yeah. A small project that you can start from the very beginning, complete it, and have some results, I think, maybe. Yeah. But I know it's hard. |
| Interviewer: 02:18 | Yeah. Would this be something from your own research or a project that everyone could work on together? |
| Researcher: 02:23 | Yeah. I think everyone can work on together should be fine. |
| Interviewer: 02:26 | Okay. Yeah. That's a great idea. |
| Researcher: 02:28 | Yeah. Then you have some sense how to initiate a project, how to solve a problem step by step by this kind of solving. |
| Interviewer: 02:38 | Yeah. No, that's a great idea. All right. So let's take a look at your workflow. So we drew back in probably January or February. It's been a while. |
| Researcher: 02:51 | Oh, I can remember |
| Interviewer: 02:52 | Do you remember everything you said? |
| Researcher: 02:54 | Yeah, I think so. |
| Interviewer: 02:55 | I think you had the most drawing of anybody. So just to kind of summarize. So we were talking about kind of [body part] development-- |

| | |
|---|---|
| Researcher: 03:03 | Yes, that's true. |
| Interviewer: 03:04 | --using some single-cell sequencing. |
| Researcher: 03:06 | Yeah, and some [methods]. |
| Interviewer: 03:10 | Yeah. And so it looked like you were using kind of R already, and you were a little bit kind of unhappy about how it was working, or-- |
| Researcher: 03:19 | Yeah, yeah. And I've slightly shifted to Python now. |
| Interviewer: 03:25 | Oh, you have? |
| Researcher: 03:25 | Yeah, yeah. |
| Interviewer: 03:26 | Great. So let's talk about that. And then what was all this down here? This was kind of-- |
| Researcher: 03:32 | This is preparing figures, I think, yeah. Using other software, combine with R and Python to generate figures. |
| Interviewer: 03:42 | Okay. And then these were all the figures that we were creating? |
| Researcher: 03:44 | Yeah. I think the experiments are the same, but now I spent some time to analyze some images, but mainly manually that is Image J, yeah. Yeah. I think the the biggest change here is now I can use python to plot figures now. |
| Interviewer: 04:05 | Great. Okay. |
| Researcher: 04:06 | Yeah. It's amazing. |
| Interviewer: 04:09 | Actually, with our pink, anything that's different now, yeah, go ahead and write it and-- |
| Researcher: 04:16 | I don't use R, I just use python to visualize figures I use some boxplots, rasta plot and others. Another plot I often use is the pie one. A beautiful pie chart. |
| Interviewer: 04:41 | It even looks like a donut chart, actually. |
| Researcher: 04:43 | Yeah. I just made donut-shaped pie. Yeah. It's amazing. Indeed, I have two layers. |
| Interviewer: 04:52 | Oh, wow. |
| Researcher: 04:52 | Yeah. You can further divide it into different categories, really powerful. |
| Interviewer: 04:59 | And why did you decide to switch from R to python for your visualizations. |
| Researcher: 05:03 | Oh. This is hard to answer. why I stick to python. Indeed, I have looked for a lot of articles comparing R to Python. There are endless discussions of it. |
| Interviewer: 05:27 | Indeed. |
| Researcher: 05:28 | Yeah. And so I say it is related to my plan B. My plan A is to maybe stay in academic and then doing some research. My plan B is to go to industry and doing some things and Python's more universal to do that. So now shift to python. |
| Interviewer: 05:50 | Oh, cool. |
| Researcher: 05:50 | To do anything I want. |
| Interviewer: 05:52 | Yeah. |

| | |
|---|---|
| Researcher: 05:53 | Yeah. Yeah. That's the reason. And they are similar powerful. They can learn from each other since they're quite similar. |
| Interviewer: 06:06 | Yeah. And what has that kind of done for your research, to use python instead of R? |
| Researcher: 06:12 | Oh. Now I mainly use for generating figures, publishing figures. In the future, I may also want to use it to analyze some images. Yeah. Because it says there are a lot of good packages to analyze images. But now, an image is not so good quality. not so easy to recognize so I analyze manually now. |
| Interviewer: 06:40 | Oh, okay. |
| Researcher: 06:40 | Yeah. Yeah. And I may also used it to-- as I told you last time I also still have a very small, single cell data set because I may also use it to analyze this one. it seems that python also have good packages compared to Seurat and the-- yeah. |
| Interviewer: 07:03 | Yeah. So is there any other changes that you've made to this workflow? |
| Researcher: 07:09 | Not so many, I think. Maybe the data analyze and data visualization. But at this stage, no. |
| Interviewer: 07:19 | So then the other color was going to be green. So this was for future changes, things you're planning on doing differently. |
| Researcher: 07:24 | I see. Maybe this part for image analyze I just, ideally it's automatic. Yeah that would be everything for me. |
| Interviewer: 07:39 | Yeah. |
| Researcher: 07:40 | Yeah. |
| Interviewer: 07:40 | And that would be on python? |
| Researcher: 07:41 | Yeah. Yeah. because there are good machine learning algorithms that may help me. Yeah. I hope it works. |
| Interviewer: 07:49 | And were you familiar with those before-- |
| Researcher: 07:51 | No. No. |
| Interviewer: 07:52 | --or is that something you learned about in a workshop? |
| Researcher: 07:54 | Yeah. I just learned from the workshop. And I think there may be still a lot good training online, so I can also learn some of them. Yeah, I think that another great resource is-- the library also have some one-to-one consulting. I can also make an appointment. |
| Interviewer: 08:14 | Have you done that yet? |
| Researcher: 08:15 | Not yet, because they're at Mission Bay. But at start of this, I definitely will go there because it's more complex than this kind of simple data visualization. |
| Interviewer: 08:27 | So maybe a future route would be to kind of keep going with Python and then use it for the project. |
| Researcher: 08:32 | Yeah, yeah, yeah. And this one, I must say. |
| Interviewer: 08:36 | And also the single-cell. |

| | |
|---|---|
| Researcher: 08:37 | Single-cell, yeah. Single-cell data sets. And I also have some new planned experiment. Maybe it's similar to use RNA seq. Yeah, although different methods, but as a similar concept to analyze it. |
| Interviewer: 08:52 | Also in Python with that? |
| Researcher: 08:54 | Yeah, yeah. |
| Interviewer: 08:55 | Interesting, yeah. And you mentioned Git and GitHub. Is that something that you're using at all now? |
| Researcher: 09:03 | I don't use it quite often, but I look for a lot of resource there, yeah, because there was some open codes there and there is also good resource for some training, yeah. And yeah, that's it. I just-- |
| Interviewer: 09:24 | So not so much for your own code, but to find other people's code. |
| Researcher: 09:28 | Other people's code, yeah, and some famous people, I can also follow them. |
| Interviewer: 09:31 | Oh, cool. And had you used GitHub before the workshop? |
| Researcher: 09:36 | No. |
| Interviewer: 09:36 | Okay, so it was new |
| Researcher: 09:37 | Yeah, yeah, new. And I found it really powerful, and I never found it, and yeah, that's why so many people use it [laughter]. |
| Interviewer: 09:45 | Yeah. That's a really great resource. And what about the UNIX shell or command line? |
| Researcher: 09:55 | Yeah, the command line-- because I use Windows system, but after the workshop, I also stick to the-- Git Bash is also a command line. I just want to yeah, so yeah every time I just use that to start a Jupyter Notebook [laughter], then yes. So help me be familiar with the single commands, the present folder CD enters, yeah. |
| Interviewer: 10:24 | And is that the only time you use it, just to start up your Jupyter Notebooks? |
| Researcher: 10:27 | Yeah, I think so. For copying a removed file, I just stick to the Windows system. It's easier [laughter]. Yeah. |
| Interviewer: 10:35 | Yeah, and so with this Python, what you're doing for visualizations, that's all through Jupyter? |
| Researcher: 10:41 | Yeah, all through the Jupyter Notebook. |
| Interviewer: 10:43 | And had you used Jupyter before the workshop? |
| Researcher: 10:46 | I just try to use it, but I don't know how to use it at that time, but after the workshop, I learned how to import it out, output data. That's convenient. |
| Interviewer: 10:55 | Nice. Kind of similar-- because I remember you liked R-- what's it called? |
| Researcher: 11:03 | Yeah, at the very beginning, RStudio. |
| Interviewer: 11:07 | RStudio. |
| Researcher: 11:07 | RStudio is more user-friendly than Jupyter Notebook [laughter]. Yeah, but now, it seems Jupyter also have its own advantage. They are very strict structure, then you have different modules there. |

Interviewer: 11:21 Yeah, yeah. That's great. So actually, is there anything else that you changed about your workflow or future changes that you've been thinking about?

Researcher: 11:36 Yeah, I think just mainly data analyze all the imaging of the single-cell data, but yeah. And now I have a focus on data visualization because, yeah, having a good figure is very important for publishers.

Interviewer: 11:51 Oh, yeah. Very true. And so one thing we're trying to figure out is-- you take a two-day workshop and then you go off and some people they get right into it and some people don't. Is there anything that kind of was extra motivating or kind of enabled you to make this change to Python? Anything that promoted success?

Researcher: 12:16 Oh you mean how can I stick to the two day workshop?

Interviewer: 12:20 Yeah. How did you kind of take what you learned and actually start using that?

Researcher: 12:25 I say it's just because my plan B. I want to have more choice in the future. And another thing is very important to learn at least something and it uses I think the library provided the most high quality and low cost workshop. Because for the same lesson, I understand is very expensive I think, several hundreds I think? But I think the library one was no more than $30?

Interviewer: 13:03 $25.

Researcher: 13:04 Okay. Wow. If you have intermediate or advanced lessons, I will also still attend.

Interviewer: 13:14 So more classes would be helpful?

Researcher: 13:15 Yeah. I think so. For me, I now focus on data visualization. I also attended several lessons since they only just simple introduction of several softwares. But that was helpful for me to choose which one we were using.

Interviewer: 13:37 And what about barriers? Was there anything between the workshop and now that made it harder for you to kind of use the things you've learned?

Researcher: 13:48 I think the main barriers is sometimes you have bugs there. But from the workshop, they also teach you how to solve it. There's another thing I forget to mention. Just google it then. In most of the case you can solve it. But you may need to spend maybe 20 minutes to look for the solutions.

Interviewer: 14:14 So a barrier is when you have a mistake?

Researcher: 14:17 Yes. Sometimes. But I say for data visualization is sometimes very simple it's just a summation. I can google it and after looking for less than maybe five website and I can solve the problem. But another barrier is sometimes I don't know why I just copy their code put in there, it works. Sometimes just these kind of things. But I know because Python has different versions. They made one very fast, maybe some structure belongs to the old versions and the new one. But at long as it works I won't spend too much time on thinking why this happen, why don't that happen. Just do it, it works.

Interviewer: 15:16 And do you have other people that you work with that use Python or R?

Researcher: 15:22 Yeah. Most of our lab members I think they use R because they use it to analyze the single-cell data.

Interviewer: 15:30 So you're kind of a rebel going off for Python more?

| | |
|---|---|
| Researcher: 15:33 | Yeah. But I think they may learn some. I don't know yeah. But I think most of the point I just saw in the the google search because that's great. |
| Interviewer: 15:46 | So google's what you mostly use for resources. |
| Researcher: 15:48 | Yes. |
| Interviewer: 15:51 | That makes you a real software engineer [laughter]. |
| Researcher: 15:53 | Yeah. Something like that. |
| Interviewer: 15:56 | That's what everybody does. |
| Researcher: 15:57 | But it's still a lot of fun it's a problem you finally solve it, it's very exciting, yeah [laughter]. |
| Interviewer: 16:04 | It's very satisfying. |
| Researcher: 16:05 | Yeah. And I say another advantage is my mentor is a little-- is very open for us to learn something new. He just no force, you should do this today, you should do that day, then I've tried to learn these new things. |
| Interviewer: 16:25 | Oh nice, so your mentor, PI, makes time for you? |
| Researcher: 16:28 | Yeah. I think that's-- I think, for some people there may be some barriers for them if they want to enter this two day workshop because they say no, you can do lot of things during these two days. You should not waste your time, yeah. |
| Interviewer: 16:43 | And does your mentor, do they also know how to program? |
| Researcher: 16:47 | No. |
| Interviewer: 16:48 | No. Okay. |
| Researcher: 16:49 | Yeah. He didn't know but he's very supportive. I'm very happy. |
| Interviewer: 16:52 | Yeah, that's great. |
| Researcher: 16:53 | Yeah, very thankful for that. |
| Interviewer: 16:55 | Good, good. |
| Researcher: 16:56 | Yeah. That's why I can enter a lot of trainings from the library. |
| Interviewer: 17:01 | What other trainings have you done? |
| Researcher: 17:03 | Oh, a lot. I think the first lesson is learn how to use Pubmed to-- you can make structured search, I never learn that before. |
| Interviewer: 17:13 | Oh yeah. |
| Researcher: 17:14 | Several combinations makes more efficiency. Another is learn to write how to write a systematic review. Yeah, that's also similar. I didn't learn that before-- I didn't before that I just thought a systematic review was a kind of review nothing then after that I learn much more than. I prefer to look for systemic review now, yeah. And there also the literature manager software. Before that I used EndNote but after here I use Zotero writing. |
| Interviewer: 17:54 | Oh, Zotero? |

| | |
|---|---|
| Researcher: 17:55 | Yes, Zotero. |
| Interviewer: 17:57 | Do you use that in your workflow now? |
| Researcher: 18:00 | I say in my writing paper. I have a-- I forget-- I also prepare my manuscript, yeah. I indeed, at the beginning I wanted to use the LibreOffice, the open office combined with Zotero but I think because my boss use the Microsoft software so to help them to easily revise my script so I use the Microsoft. |
| Interviewer: 18:29 | So you use Word but with Zotero? |
| Researcher: 18:30 | Yes. |
| Interviewer: 18:31 | Oh, nice. |
| Researcher: 18:31 | Yeah, yeah, it's support really well, I think. |
| Interviewer: 18:33 | Yeah, that's great. |
| Researcher: 18:34 | Yeah. That's very cool. |
| Interviewer: 18:36 | Very cool. You've been taking so many classes. |
| Researcher: 18:38 | Yeah. Library's amazing, I know. I also entered, there's a 3D printer. It's lot of fun. |
| Interviewer: 18:44 | Oh yeah, the maker's lab? Yeah, that's great. So just to kind of wrap up then, we've got just a little checklist here. So this is just about the way that it is currently. So not looking at future stuff but with the pink. And so with this current workflow do you use any programming languages like R Python or the Command line? |
| Researcher: 19:07 | Yeah. |
| Interviewer: 19:08 | Yeah? |
| Researcher: 19:08 | Yeah. |
| Interviewer: 19:10 | You use Python and R. Do you have any places where you've transformed a step-by-step workflow into scripts or functions? |
| Researcher: 19:17 | You mean I have my own module? I can write a function myself? |
| Interviewer: 19:23 | Yeah. |
| Researcher: 19:23 | Not yet I think it's just, yeah. Download |
| Interviewer: 19:27 | Yeah, yeah. Do you use any version control to manage your code? |
| Researcher: 19:32 | No. No, no. Yeah. |
| Interviewer: 19:35 | Do you use any open-source software? |
| Researcher: 19:36 | Yeah. |
| Interviewer: 19:38 | Do you share any of your code publicly? |
| Researcher: 19:42 | Because I don't have too much my own code so not yet. But I will definitely share when I publish paper, I think for the Github and because this is also relevent to make my research reproducible by others is very important, I think. |
| Interviewer: 19:57 | Yes, definitely. And do you share any of your computational workflow or protocols publicly? |

| | |
|---|---|
| Researcher: 20:08 | My workflow? Not yet, I must say. Yeah. Because yeah, this is-- most of biological is we have old protocol and we don't have so much workflow, I think. Yeah. |
| Interviewer: 20:23 | Yeah. Doesn't always apply. Great. So then the last question was kind of what you just started talking about, actually, which is the-- one of the goals with this workshop is we want you to learn Python so you can use Python at your work but overall something we're hoping to do is help people make their work more computationally reproducible. |
| Researcher: 20:43 | Yeah, that's true. |
| Interviewer: 20:44 | So I'm kind of wondering, do you think that participating in this workshop has helped you make your work more reproducible? |
| Researcher: 20:50 | Sure, I think. Especially the version control concept. I don't use it quite often because I now use another kind just same as in my home computer. Yeah. |
| Interviewer: 21:02 | For your code as well? |
| Researcher: 21:03 | Yeah because another reason is for Python, for Jupyter notebook I haven't found an easy way to parse the file there. And I just put the files in the Google Drive so, yeah. |
| Interviewer: 21:25 | But you think that the workshop helped you? |
| Researcher: 21:29 | Yeah. Because for data visualization, I think you don't need version control but for other things that you need to adjust some parameters is very important, yeah. |
| Interviewer: 21:43 | Great. And anything besides the version control aspect of it? |
| Researcher: 21:48 | I think-- you mean anything besides version control to help me? |
| Interviewer: 21:54 | Be more reproducible. |
| Researcher: 21:55 | Be more reproducible. I used to use another thing. Maybe if you can besides the version control you should share this code. Then that may help you to know if someone else use your code immediately they may have some feedback. |
| Interviewer: 22:16 | Yeah. |
| Researcher: 22:17 | Yeah. |
| Interviewer: 22:17 | And you said you were thinking of doing that for future code as you write it? |
| Interviewer: 22:21 | Yeah, I think so. Even for data visualization I can do some, yeah. |
| Researcher: 22:26 | Yeah. I always like to see how'd you do it. |
| Interviewer: 22:29 | Yeah, you mentioned it, yeah. Okay. Yeah, that's a good suggestion. |
| Researcher: 22:34 | Now you're motivated. I like that. |
| Interviewer: 22:35 | Yeah. |
| Researcher: 22:36 | That's great. Anything else to share about any of your workflow or thoughts about the workshop or anything that you've learned? |
| Interviewer: 22:44 | Yeah. I think the workshop, all the workshop is great. You usually have it every three months because these things are so popular, I love them. |
| Researcher: 22:53 | It's very popular. |

| | |
|---|---|
| Interviewer: 22:54 | Yeah. And I just want, if there's a less focus on maybe it's harder and smaller audience like data visualization with R or Python, just stick to more focused topic, yeah. |
| Researcher: 23:11 | Yeah. I think we do have a plan to do more R visualization but there's more classes coming, so. |
| Interviewer: 23:19 | I see, yeah. |
| Researcher: 23:20 | Yeah, good to know. |
| Interviewer: 23:21 | Yeah. |
| Researcher: 23:22 | Great. |
| Interviewer: 23:23 | Yeah, lots of lessons then. Great, I must say. Oh, I forgot, I also entered a statistic class for R. |
| Interviewer: 23:30 | Cool. You really are, you're our best customer. |
| Researcher: 23:33 | Yeah. |
| Interviewer: 23:35 | Taking everything. |
| Researcher: 23:35 | Very old customer [laughter]. |
| Interviewer: 23:37 | Well, that's great. Great, well thanks so much for chatting. I'm going to just turn this off. |