

## GET-Evidence - Bug #5872

### 23andMe report pipeline is broken

04/30/2015 10:52 PM - Abram Connelly

<b>Status:</b>	In Progress	<b>Start date:</b>	04/30/2015
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assigned To:</b>	Abram Connelly	<b>% Done:</b>	0%
<b>Category:</b>		<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	Upgrade work	<b>Totalhours:</b>	
<b>Billable:</b>		<b>Resolution:</b>	
<b>Estimatedhours:</b>			
<b>Hours:</b>			

#### Description

When participants upload their 23andMe data, this kicks off a pipeline run through Arvados (now) through the same code path that the dataset submission goes through from Tapestry. The pipeline is failing to correctly generate reports for 23andMe data.

Porting the pipeline over to Arvados, the format for the input was changed and that's why the pipelines are failing on 23andMe data. I think the best thing to do going forward is to be more liberal about what input collection is specified so that 23andMe reports along with older reports can be run (or re-run) without having to worry about updating to the new format.

The old format just specifies a portable data hash. The new format expects <PDH>/<filename>.

#### History

#1 - 05/14/2015 09:57 PM - Abram Connelly

- Target version set to Upgrade work