From: **Lau, Matthew K.** matthewklau@fas.harvard.edu
Subject: Re: Follow up on GATK on insects
Date: February 14, 2017 at 5:18 PM
To: Sarah Young stowey@broadinstitute.org

Fantastic, thanks Sarah!

Matt

On Feb 14, 2017, at 4:12 PM, Sarah Young <stowey@broadinstitute.org> wrote:

Hi Matt,

I know you were recently emailing about GATK usage.  I did a bit of a follow up with some of our computational biologists here who are also working with insects and got a bit more information for you. From Seth Redmond :

"Not sure about wider insects but I don't think we've found anything significantly better for mosquitoes than using GATK. For the AG1000g study we used Unified Genotyper with some manual filtering based on a validated set of F1 crosses. Methods are all up on Biorxiv (in the supplemental): http://biorxiv.org/content/early/2016/12/22/096289

UnifiedGenotyper was largely used because HaplotypeCaller can take a *very* long time with highly heterogeneous samples (the reassembly graph gets very large), but personally I'd lean on haplo caller if you have the time, particularly for INDELs and particularly if you don't have a validation set of markers to compare against."

If you have any more direct questions, I'm happy to put you directly in touch with him.

~Sarah

--
Sarah Young
Director, Operations, Finance and Computation
Broad Technology Labs
The Broad Institute
75 Ames Street
Cambridge, MA  02141
T:  (617) 714-8508
E:  stowey@broadinstitute.org

From: **Sarah Young** stowey@broadinstitute.org
Subject: Follow up on GATK on insects
Date: February 14, 2017 at 4:13 PM
To: Lau, Matthew matthewklau@fas.harvard.edu
Cc: Caroline Cusick ccusick@broadinstitute.org, James Bochicchio jboch@broadinstitute.org

SY

Hi Matt,

I know you were recently emailing about GATK usage.  I did a bit of a follow up with some of our computational biologists here who are also working with insects and got a bit more information for you.  From Seth Redmond :

"Not sure about wider insects but I don't think we've found anything significantly better for mosquitoes than using GATK. For the AG1000g study we used Unified Genotyper with some manual filtering based on a validated set of F1 crosses. Methods are all up on Biorxiv (in the supplemental): http://biorxiv.org/content/early/2016/12/22/096289

UnifiedGenotyper was largely used because HaplotypeCaller can take a *very* long time with highly heterogeneous samples (the reassembly graph gets very large), but personally I'd lean on haplo caller if you have the time, particularly for INDELs and particularly if you don't have a validation set of markers to compare against."

If you have any more direct questions, I'm happy to put you directly in touch with him.

~Sarah

--
**Sarah Young**
Director, Operations, Finance and Computation
Broad Technology Labs
The Broad Institute
75 Ames Street
Cambridge, MA  02141
T:  (617) 714-8508
E:  stowey@broadinstitute.org