CIViC U01 - Update on APTRC supplement and collaborative set-aside activities

Obi Griffith
April 4, 2019
Activities to Promote Technology Research Collaborations (APTRC) for Cancer Research (Admin Supp)

- Leverages two unique NCI programs focused on supporting technology development research with transformative potential for advancing cancer research
  - The Innovative Molecular Analysis Technologies (IMAT) program supports highly innovative, data-generating platforms and methodologies while
  - Informatics Technologies for Cancer Research (ITCR) program supports powerful data processing and visualization technologies.

- The goal of the APTRC:
  - Accelerate the development of new enabling cancer technologies
  - Support and encourage new multidisciplinary scientific collaborations between IMAT and ITCR
  - Bring together complementary technology platforms and methodologies to enhance the capabilities of either in a way that advances cancer research or clinical care
  - Activities must be within the overall scope of active parent grants

See “Invitation to propose ITCR-IMAT collaborations” email from Juli!
APTRC Aims build directly from parent aims

Aim 1. Perform comprehensive variant curation and develop evidence-based prioritization for selection of high-value clinically actionable targets for panel design

Aim 2. Develop an inexpensive, ultrasensitive experimental platform to enable clinical detection of actionable variants

Ultimate goal: Demonstrate that an open-access and open-source approach can be used to support evidence-based community-driven cancer sequencing panels
WashU Aim 1.1 - additional gaps will be identified and filled by targeted curation of key sources.

Curation for APTRC difficult to separate from overall curation activities. Targeted curation continues to prioritize the most clinically “actionable” variants.
**WashU Aim 1.2 - Development of an evidence-based prioritization scheme for automated target selection and panel design**

Create CIViC score to determine extensively curated variants

**Figure 1. Example of CIViC entry for KRAS (G12/G13).**

- The first panel details the evidence summary at the gene level.
- The second panel details the evidence summary at the variant level.
- The third panel details evidence items that support the variant and gene level summaries.

**CIViC Score = \( \sum (\text{Evidence Level}) \times (\text{Trust Rating}) \)**

<table>
<thead>
<tr>
<th>EID</th>
<th>DESC</th>
<th>DIS</th>
<th>DRUGS</th>
<th>EL</th>
<th>ET</th>
<th>ED</th>
<th>CS</th>
<th>VO</th>
<th>TR</th>
</tr>
</thead>
<tbody>
<tr>
<td>622</td>
<td>Adding tramet...</td>
<td>Pancreatic Ad...</td>
<td>Trametinib, G...</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4★</td>
</tr>
<tr>
<td>990</td>
<td>KRAS G12/G1</td>
<td>Pancreatic Ad...</td>
<td>Trametinib</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3★</td>
</tr>
<tr>
<td>134</td>
<td>Chemotherap...</td>
<td>Colorectal Ca...</td>
<td>Cetuximab</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3★</td>
</tr>
<tr>
<td>135</td>
<td>KRAS mutation...</td>
<td>Colorectal Ca...</td>
<td>Cetuximab</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3★</td>
</tr>
</tbody>
</table>

- EID662 = B(4) * 4 = 16
- EID990 = B(4) * 3 = 12
- EID134 = B(4) * 3 = 12
- EID135 = C(3) * 3 = 9

**CIViC Score = 49**
WashU Aim 1.3 - Modification of CIViC interface to display variant prioritization score and probe design

- In order to facilitate widespread usage and integration of the prioritization score and probe design modifications to the CIViC interface will be needed.
- Scores will be displayed at gene and variant level indicating cumulative score of “clinical relevance”
- Scores will be integrated into existing search and filtering features of the site
- Facilitate rapid identification of the complete set of variants currently eligible for inclusion in a CIViC-based panel.
- Designed probe sequences for each variant will be made available through the public interface, providing complete transparency and facilitating widespread adoption of the technology.
- As CIViC variants are added and achieve threshold score, new probes will be added allowing the continual evolution of a comprehensive and current precision medicine resource.
CIViC Variant Evidence Score is dynamically calculated and displayed

**VARIANT V600E**

**Aliases:** RS113488022 and VAL600GLU

**Allele Registry ID:** CA123643

BRAF V600E has been shown to be recurrent in many cancer types. It is one of the most widely studied variants in cancer. This variant is correlated with poor prognosis in certain cancer types, including colorectal cancer and papillary thyroid cancer. The targeted therapeutic dabrafenib has been shown to be effective in clinical trials with an array of BRAF mutations and cancer types. Dabrafenib has also shown to be effective when combined with the MEK inhibitor trametinib in colorectal cancer and melanoma. However, in patients with TP53, CDKN2A and KRAS mutations, dabrafenib resistance has been reported. Ipiilimumab, regorafenib, vemurafenib, and a number of combination therapies have been successful in treating V600E mutations. However, cetuximab and panitumumab have been largely shown to be ineffective without supplementary treatment.

**Variant Type:** Missense Variant

**HGVS Expressions:**
NM_004333.4:c.1799T>A, NP_004324.2:p.Val600Glu, NC_000007.13:g.140453136A>T, and ENST00000288602.6:c.1799T>A

**ClinVar ID:**
13961

**CIViC Variant Evidence Score:**
1019

**Evidence for V600E** 154 total items (showing 147)

<table>
<thead>
<tr>
<th>EID</th>
<th>DESC</th>
<th>DIS</th>
<th>DRUGS</th>
<th>EL</th>
<th>ET</th>
<th>ED</th>
<th>CS</th>
<th>VO</th>
<th>TR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1409</td>
<td>Phase 3 randomized clinical tr...</td>
<td>Skin Melanoma</td>
<td>Vemurafenib</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3017</td>
<td>Patients with BRAF V600E-mut...</td>
<td>Non-small Cell Lung Carcinoma</td>
<td>Trametinib, Dabrafenib (Comb...</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Panel-development incorporated into CIViC search and download features

### Search Variants

**Example Searches:**
- CHR1 Chromosome is 2
- CHR1 Start between 16 and 60K
- Variant type contains frameshift

**Match all** of the following conditions:

- **CIViC Variant Evidence Score**: is above 20
- **Pipeline Type**: is DNA-Based

### Search Results

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variant Name</th>
<th>Variant Group(s)</th>
<th>Variant Types(s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALK</td>
<td>F1174L</td>
<td>Crizotinib Resistance</td>
<td>Missense Variant</td>
<td>ALK F1174L has been observed as recurrent in patients with non-small cell lung cancer</td>
</tr>
<tr>
<td>ALK</td>
<td>EML4-ALK L1196M</td>
<td>ALK Fusions, Crizotinib Resistant</td>
<td>Transcript Fusion, Missense Variant</td>
<td>In patients with non-small cell lung cancer</td>
</tr>
<tr>
<td>ALK</td>
<td>R1275Q</td>
<td>N/A</td>
<td>Missense Variant</td>
<td>ALK R1275Q has been observed as recurrent in patients with non-small cell lung cancer</td>
</tr>
<tr>
<td>ALK</td>
<td>ALK FUSION I1171</td>
<td>N/A</td>
<td>Missense Variant, Transcript Fusion</td>
<td>Multiple case reports indicate that the I1171 mutation is recurrent in patients with non-small cell lung cancer</td>
</tr>
<tr>
<td>ALK</td>
<td>ALK FUSION G1202R</td>
<td>N/A</td>
<td>Transcript Fusion, Missense Variant</td>
<td>ALK fusion with G1202R mutation in the patients with non-small cell lung cancer</td>
</tr>
<tr>
<td>AKT1</td>
<td>E17K</td>
<td>N/A</td>
<td>Missense Variant</td>
<td>AKT1 E17K is a recurrent mutation that has been observed in patients with non-small cell lung cancer</td>
</tr>
</tbody>
</table>
Developed OpenCAP resource to walk users through panel development

OpenCAP describes methods for capture panel development using variants within the CIViC database.

http://opencap.org
smMIPs provide an ultra-sensitive, cost-effective method for assaying variants

- The single-molecule Molecular Inversion Probes (smMIPs) method modifies established assays for multiplex sequence capture using MIPs such that each copy of a target is “tagged” with a Unique Molecular Identifier (UMID), a unique, random DNA sequence, during the capture reaction.
- Couples scalable target enrichment with sequence read error correction:
  - High multiplex capabilities: >50,000 reactions, targeting up to 125kb without intra-probe interference.
  - UMID retained on all copies of a smMIP molecule during PCR and sequencing used to determine which sequence reads share a common origin.
  - Reads derived from the same capture event error-correct one another to yield an independent haploid consensus for each progenitor molecule.
  - Error-rate of sequencing and library preparation is effectively eliminated.
- Advantages: generally applicable to any target, cost-effective, ultrasensitive, quantitative, and scalable
UW (Salipante) Aims - Develop an inexpensive, ultrasensitive experimental platform to enable clinical detection of actionable variants

• 2.1 smMIP panel design
  • Design a smMIP capture panel targeting ~70kb of genomic sequence.
    • CIVIC-prioritized targets
    • full-gene tiling and mutational hotspot targeting
    • double-stranded capture design

• 2.2 Analysis pipeline
  • Data-cleaning, merging overlapping paired-end reads, grouping of reads based on shared UMID sequences, UMID-mediated consensus calling of reads originating from the same progenitor molecules, optimized alignment of consensus sequences to targets, and variant calling
  • Expand functionality to double-stranded capture design
  • Intersect variant calls with the CIViC database API to create clinical summaries for individual cancer patients

• 2.3 Validation Approach
  • Validate the CIViC smMIP panel by re-characterizing clinical specimens which have previously been submitted for genomic characterization
  • Evaluate at least 50 de-identified cancer specimens (predominantly lung adenocarcinoma, colorectal adenocarcinoma, and endometrial carcinoma specimens
    • evaluate sensitivity and specificity

• 2.4 Prospective application of panel to clinical specimens
  • Prospectively sequence at least 20 residual, de-identified clinical specimens submitted for genomic characterization by existing high-throughput sequencing clinical oncology assays offered by our institutions (UW-OncoPlex and WUCaMP)
  • Compare variants identified, de-identified clinical reports issued, assess whether the enhanced performance capabilities offered by CIViC-smMIP (ultrasensitivity and/or improved clinical interpretation) would have impacted patient care

Stephen Salipante
A demonstrative panel was created using CIViC/OpenCAP.
The panel contained 111 clinically relevant variants/regions/genes.
Samples with orthogonal sequencing were identified

<table>
<thead>
<tr>
<th>Malignancy</th>
<th>Symbol</th>
<th>Cases</th>
<th>dpGAP #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head and neck squamous cell carcinoma</td>
<td>HNSCC</td>
<td>5</td>
<td>phs001623</td>
</tr>
<tr>
<td>Small cell lung cancer</td>
<td>SCLC</td>
<td>9</td>
<td>phs0001049</td>
</tr>
<tr>
<td>Hodgkin’s lymphoma</td>
<td>HL</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Acute myeloid leukemia</td>
<td>AML</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>CRC</td>
<td>5</td>
<td>phs000159</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>22</td>
<td>3</td>
</tr>
</tbody>
</table>

- 22 Samples were identified to test CIViC smMIPs panel design
- Samples were derived from 5 different tumor subtypes
- Each sample had previous orthogonal sequencing (whole exome- or whole genome sequencing)
- Each sample has known true somatic variants that overlap with variants in the CIViC smMIPs capture panel
Variants observed on WEX/WGS were also observed using the CIViC smMIPs capture panel.
Variant Allele Frequencies were concordant between both sequencing methods

- VAF showed Pearson’s R correlation of 0.885
- Variants that were not observed on smMIPs but were identified on original sequencing was not associated with tumor type, total coverage, or DNA mass input
173 clinical variants were identified by smMIPs but not observed on original sequencing

- 162 variants showed variant support using smMIPs sequencing but no support on original sequencing
- VAF for these variants was below the level of detection based on original sequencing coverage

- 11 variants showed variant support on both smMIPs sequencing and original sequencing but were not labeled as a true somatic variant on original sequencing
- VAF for these variants was was directly correlated on both sequencing platforms
Identification of community-consensus clinically relevant variants and development of single molecule molecular inversion probes using the CIViC database


doi: https://doi.org/10.1101/479394

This article is a preprint and has not been peer-reviewed [what does this mean?].
CIViC U01 Collaborative set-aside proposal

- Goal: Develop collaboration between CIViC and ClinGen Somatic Working Group

- Task 1. Develop CIViC Assertion framework
  - Subtask 1.1 Develop assertions data model and UI
  - Subtask 1.2 CIViC-ClinGen face-to-face meeting

- Task 2. Develop the pipeline for ClinGen expert curation in CIViC and ClinVar submission
  - Subtask 2.1 Develop ClinVar submission pipeline
  - Subtask 2.2 CIViC-ClinGen Hackathon

<table>
<thead>
<tr>
<th>Timeline: Tasks/Milestones</th>
<th>U01 Year 2</th>
<th>U01 Year 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1: Develop assertion framework</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1 Develop assertions data model and UI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2 CIViC-ClinGen F2F meeting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 2: Develop the ClinGen -&gt; CIViC -&gt; ClinVar submission pipeline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1 Develop ClinVar submission pipeline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2 Hackathon</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ClinGen - NIH-funded resource dedicated to defining the clinical relevance of genes and variants for use in precision medicine and research

- Primarily focused on germline/constitutive disorders
  - Somatic Cancer Working Group

ClinGen Somatic Cancer Working Group

Subha Madhavan
Shashikant Kulkarni

>60 members

Pediatric Cancer Task force
- Gordana Raca
- Angshumoy Roy
- 17 members

Pancreatic Cancer Taskforce
- Gagandeep Brar
- Matthew McCoy
- 6 members

Genitourinary Cancer Taskforce
- Dr. Jue Wang
- Newly Forming!
Members of CIViC team joined last two ClinGen SWG in-person meetings to gather requirements

- Harmonization with MVLD
- CIViC “Organizations” to track curation activity of ClinGen groups and sub-groups
- Support for AMP/ACMG evidence levels
  - Develop CIViC Assertions
- Support ClinVar submission
ClinGen MVLD format was mapped to CIViC
Adapting crowdsourced clinical cancer curation in CLiViC to the ClinGen minimum variant level data community-driven standards

Arpad M. Danos1* iD | Deborah I. Ritter2* | Alex H. Wagner1 iD | Kilannin Krysiak1 | Dmitriy Sonkin3 iD | Christine Micheel4 | Matthew McCoy5 | Shruti Rao5 | Gordana Raca6 | Simina M. Boca5 | Angshumoy Roy2 | Erica K. Barnell1 | Joshua F. McMichael1 | Susanna Kiwala1 | Adam C. Coffman1 | Lynzey Kujan1 | Shashikant Kulkarni2,7,8 | Malachi Griffith1 iD | Subha Madhavan5 | Obi L. Griffith1 iD | on behalf of The Clinical Genome Resource Somatic Working Group and Clinical Interpretation of Variants in Cancer team members

1 McDonnell Genome Institute, Washington University School of Medicine, Saint Louis, Missouri
2 Baylor College of Medicine, Houston, Texas
3 Biometric Research Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Rockville, Maryland
4 Vanderbilt-Ingram Cancer Center, Nashville, Tennessee
5 Georgetown Lombardi Comprehensive Cancer Center, Washington, District of Columbia
6 Keck School of Medicine, University of Southern California, Los Angeles, California
7 Baylor Genetics, Houston, Texas
8 Dan L. Duncan Cancer Center, Houston, Texas
Organizations and sub-organizations allow ClinGen working group and task teams to track activity

<table>
<thead>
<tr>
<th>Organization</th>
<th>Actions</th>
<th>Last Action</th>
<th>Members</th>
</tr>
</thead>
<tbody>
<tr>
<td>The McDonnell Genome Institute</td>
<td>28,172</td>
<td>7 days ago</td>
<td>23</td>
</tr>
<tr>
<td><a href="http://genome.wustl.edu/">http://genome.wustl.edu/</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illumina</td>
<td>2,709</td>
<td>2 years ago</td>
<td>1</td>
</tr>
<tr>
<td><a href="https://www.illumina.com/">https://www.illumina.com/</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>University Health Network (Toronto)</td>
<td>1,960</td>
<td>about a month</td>
<td>8</td>
</tr>
<tr>
<td><a href="http://www.uhn.ca/">http://www.uhn.ca/</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Charité Comprehensive Cancer Center</td>
<td>1,215</td>
<td>3 days ago</td>
<td>2</td>
</tr>
<tr>
<td><a href="https://cccc.charite.de/en/">https://cccc.charite.de/en/</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ClinGen</td>
<td>204</td>
<td>29 days ago</td>
<td>28</td>
</tr>
<tr>
<td><a href="https://www.clinicalgenome.org/">https://www.clinicalgenome.org/</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pediatric Cancer Task Force at ClinGen</td>
<td>159</td>
<td>8 days ago</td>
<td>9</td>
</tr>
<tr>
<td><a href="https://www.clinicalgenome.org/working-groups/somatic/non-small-cell-lung-cancer-nsc1c/pediatric/">https://www.clinicalgenome.org/working-groups/somatic/non-small-cell-lung-cancer-nsc1c/pediatric/</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCCA (POGS)</td>
<td>134</td>
<td>4 months ago</td>
<td>3</td>
</tr>
<tr>
<td><a href="http://www.bcgsc.ca/project/pog">http://www.bcgsc.ca/project/pog</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treehouse Childhood Cancer Initiative</td>
<td>64</td>
<td>8 days ago</td>
<td>3</td>
</tr>
<tr>
<td><a href="https://treehousegenomics.ucsc.edu/">https://treehousegenomics.ucsc.edu/</a></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CIViC Assertions allow multiple evidence items to be summarized into a single "state-of-the-field" statement.
Somatic WG adopted CIViC as curation platform

- 34 ClinGen Somatic WG curators in CIViC
- 238 Evidence items entered and reviewed
- 3 Somatic assertions using AMP guidelines
  - MVLD pre-curation of somatic assertion
  - Curators enter PMID evidence items
  - Somatic WG “sign off” assertions
- Pilot ClinVar submission in progress
Partnership between ClinGen, CIViC and Cancer Genetics journal creates a pipeline for rapid dissemination of clinically relevant cancer variants

- **Cancer Genetics Journal**
  - Simple submission
  - Rapid review and publication

- **CIViC**
  - Expert review
  - Data formatting

- **ClinVar**
  - Transfer interface from Civic
  - Variant scoring system
CIViC-ClinGen Hackathon and curation jamboree at Scripps

- 50 participants
- 25 institutions
- 5 countries, 4 continents
Curators discussed incentivization schemes, best practices, data model, evidence codes, and worked on curation.

https://github.com/griffithlab/civic-meeting/issues/
### Hackers focused on CIViC integration

<table>
<thead>
<tr>
<th>#</th>
<th>Title</th>
<th>Author</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>CIViC networks available in NDEx</td>
<td>agary-ucsd</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Strategies for Variant Normalization &amp; Representation</td>
<td>susannasiebert</td>
<td>3</td>
</tr>
<tr>
<td>38</td>
<td>ClinGen VCI &lt;-&gt; Evidence Registry &lt;-&gt; CIViC Integration (VHL collaboration)</td>
<td>malachig</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td>Improve CIViC WikiData integration</td>
<td>malachig</td>
<td>8</td>
</tr>
<tr>
<td>33</td>
<td>Integrate CIViC with genome browsers (igv.js, JBrowse, ...)</td>
<td>jrobinso</td>
<td>4</td>
</tr>
<tr>
<td>32</td>
<td>Integration of CRAVAT and CIViC</td>
<td>malachig</td>
<td>3</td>
</tr>
<tr>
<td>31</td>
<td>Write a VEP plugin that annotates a VCF with CIViC data</td>
<td>susannasiebert</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Exporting CIViC data</td>
<td>susannasiebert</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Using the CIViCpy module to query data</td>
<td>ahwagner</td>
<td>1</td>
</tr>
</tbody>
</table>

[GitHub Link](https://github.com/griffithlab/civic-meeting/issues/)
Worked through a diagnostic assertion as a group

Group CIViC training - evidence submission, moderation and summarization
Identified specific needs for pediatric tumor variant interpretation

1. Disease ontology
   a. Adding pediatric cancer types
   b. Lynn Schriml
   c. ClinGen Pediatric Task Force

2. Human Phenotype Ontology
   a. New Term: Pediatric onset
      i. Infantile (<12 mo)
      ii. Childhood (1-5 years)
      iii. Juvenile (5-15 years)
CIViC Integration with open-sourced software

IGV Integration

CRAVAT Integration

JBrowse Integration
NDEx integration

Gene-Variant Associations

Nodes: 1216  Edges: 2944
PUBLIC  Read Only
@context: view namespaces
Owner: CIVIC Database
Created: Jul 6, 2018 6:17:06 PM
Last Modified: Apr 4, 2019 7:00:30 PM
UUID: b30b57c3-8172-11e8-a4bf-0ac135e8bacf
Format: Unknown
Description: Obtained using Clinical Evidence Summaries.

Legend:
- gene node = dark green
- Variant node = light green
- Edge colors map to "evidence_level" (A through E) and match the CIVIC color code.
- Edge thickness maps to "rating" (thinner = 1, thicker = 5)

Version: 2019-04-05
Acknowledgments

The CIViC Community

McDonnell Genome Institute

Griffith Lab (and CIViC team)

Obi Griffith
Malachi Griffith
Ben Ainscough
Erica Barnell
Katie Campbell
Kaitlin Clark
Adam Coffman
Kelsy Cotto

Arpad Danos
Yan-Yang Feng
Felicia Gomez
Jasreeta Hundal
Susanna Kiwala
Kilannin Krysiak
Lynzey Kujan
Jason Kunisaki

Josh McMichael
Cody Ramirez
Zach Skidmore
Nick Spies
Lee Trani
Alex Wagner
Jason Walker