AWARD NUMBER: W81XWH-16-1-0237

TITLE: Analysis of Gastric Adenocarcinoma Data in a Pan-GI Context to Reveal Genes, Pathways, and Interactions that Yield Novel Therapeutic Advantages

PRINCIPAL INVESTIGATOR: Rehan Akbani

CONTRACTING ORGANIZATION: The University of Texas MD Anderson Cancer Center
Houston, TX 77030

REPORT DATE: September 2018

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
Sept 2018	Annual	1 Sept 2017 - 31 Aug 2018

4. TITLE AND SUBTITLE
Analysis of Gastric Adenocarcinoma Data in a Pan-GI Context to Reveal Genes, Pathways, and Interactions that Yield Novel Therapeutic Advantages

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-16-1-0237

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Rehan Akbani, Jaffer Ajani

E-Mail: rakbani@mdanderson.org, jajani@mdanderson.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The University of Texas MD Anderson Cancer Center

Houston, TX 77030

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The subject of this research is the study of gastric cancer, where the purpose is to reveal new insights into the biology of the disease that could potentially have therapeutic implications. Specifically, the scope of the study is based on 3 broad objectives: (i) identification of dysregulated and susceptible pathways, as well as their novel inter-relationships, in gastric adenocarcinoma (GAC); (ii) Pan-Cancer comparison of GAC with other cancers to leverage therapeutic target information across cancers; (iii) Identification of novel therapeutic targets, both with and without currently known drugs that target them. We have identified novel interactions amongst pathways in stomach and other cancers, where we have identified certain sub-groups of stomach cancer patients where those pathways may be exceptionally abnormal and lead to worse survival. The interactions and corresponding sets of genes may be targetable by existing drugs and/or drugs under development for treatment of that sub-group of stomach cancer patients. Other sub-groups have other interactions and genes that may also be targeted using different drugs. In that way, we can potentially give customized regimens of drugs to specific patients whose cancers exhibit targetable characteristics.

15. SUBJECT TERMS
Gastric cancer, disrupted pathways, targetable genes, PanCancer, Pan-gastrointestinal

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	16	**19b. TELEPHONE NUMBER** *(include area code)*

TABLE OF CONTENTS

INTRODUCTION

The subject of this research is the study of gastric cancer, where the purpose is to reveal new insights into the biology of the disease that could potentially have therapeutic implications. Specifically, the scope of the study is based on 3 broad objectives: (i) identification of dysregulated and susceptible pathways, as well as their novel inter-relationships, in gastric adenocarcinoma (GAC); (ii) Pan-Cancer comparison of GAC with other cancers to leverage therapeutic target information across cancers; (iii) Identification of novel therapeutic targets, both with and without currently known drugs that target them. We have identified novel interactions amongst pathways in stomach and other cancers, where we have identified certain sub-groups of stomach cancer patients where those pathways may be exceptionally abnormal and lead to worse survival. The interactions and corresponding sets of genes may be targetable by existing drugs and/or drugs under development for treatment of that sub-group of stomach cancer patients. Other sub-groups have other interactions and genes that may also be targeted using different drugs. In that way, we can potentially give customized regimens of drugs to specific patients whose cancers exhibit targetable characteristics. The objective of this research is to find those targets and improve our understanding of the biology of stomach cancer.

KEYWORDS

Gastric cancer, stomach cancer, gastrointestinal cancers, pathway aberrations, gastric cancer therapeutic targets, Pan-Cancer, dysregulated pathways, gastric cancer subtypes.

ACCOMPLISHMENTS

What were the major goals of the project?

The major goals of the project and their breakdown into milestones (as stated in the updated and approved SOW) are shown below. Also shown is the percentage of completion for the milestones, to date. Numbers that have changed since the last progress report are highlighted in yellow.

Specific Aim 2	Timeline (months)	Percentage completed
Major Task 4: Acquisition and Quality Control of Pan-GI data, in preparation for computational analysis		
Acquire TCGA Pan-GI data and convert them into a "standardized" format suitable for computational analysis	1	100%
Assess and remove (if needed) batch effects from the data, and improve the quality of the data	2	100%
Milestone(s) Achieved: "Cleaned up" TCGA Pan-GI data ready for computational analysis	2	100%

Major Task 5: Computational analysis of the Pan-GI data sets		
Cluster the data sets and study the results	3	100%
Generate pathway activity scores for various pathways across multiple data types, and determine which ones are likely disrupted	4-5	100%
Correlate disrupted pathways across multiple data types (e.g. transcriptomic, proteomic, genomic, epigenomic) and across clinical variables (e.g. histology, stage, grade, outcome) via statistical analysis	6-7	100%
Compare gastric with other Pan-GI cancers and look for similarities and differences	8	100%
Milestone(s) Achieved: First round computational analysis for Pan-GI cancers completed	8	100%
Major Task 6: Publish Pan-GI cancer results		
Discuss results with collaborators and perform any follow up analysis	9-10	100%
Write one or more manuscripts with input from designated mentor and collaborators	11-12	100%
Submit manuscript(s) and wait for reviews	13-14	100%
Respond to reviewers and resubmit. May repeat submission/resubmission process with multiple journals depending on where the paper(s) end up being published. Present results at conferences.	15-16	100%
Milestone(s) Achieved: Pan-GI manuscript(s) published	17	100%
Total time for Specific Aim 2	**17**	
Specific Aim 3 (interspersed timeline)		
Major Task 7: Identification and publication of potential therapeutic targets in gastric cancer		
Identify potential genes and/or pathways in gastric cancer for targeted therapy, using gastric data only from Aim 1	26-27	

Identify potential genes and/or pathways in gastric cancer for targeted therapy, using cross-tumor information from Pan-GI cancers from Aim 2	9-10	100%
Integrate results into manuscripts for Specific Aims 1 and 2. Present results at conferences.	11-17	60%
Milestone(s) Achieved: Potential therapeutic targets identified and published	27	75%
Total time for Specific Aim 3 (interspersed with other aims; not consecutive months)	**27**	
Specific Aim 1		
Major Task 1: Acquisition and Quality Control of gastric cancer data, in preparation for computational analysis	**Months**	
Acquire TCGA gastric cancer data, and in-house MD Anderson data (after procuring necessary approvals)	17.5	50%
Convert all acquired data into a "standardized" format suitable for computational analysis	18	50%
Assess and remove (if needed) batch effects from within TCGA and within MD Anderson data, and improve the overall quality of each data set individually.	19	50%
Merge TCGA data with MD Anderson data, removing batch effects across both data sets.	19.5	50%
Re-assess the quality of the overall data and iterate back to previous steps, if needed, until data are satisfactory	20	50%
Milestone(s) Achieved: "Cleaned up" gastric cancer data from TCGA and MD Anderson ready for computational analysis	20	50%
Major Task 2: Computational analysis of the gastric cancer data sets		
Cluster the data sets and study the results	21	50%
Generate pathway activity scores for various pathways across multiple data types, and determine which ones are likely disrupted	22-23	75%
Correlate disrupted pathways across multiple data types (e.g. transcriptomic, proteomic, genomic, epigenomic) and across clinical variables (e.g. histology, stage, grade, outcome) via statistical analysis	24-25	50%
Milestone(s) Achieved: Computational analysis for gastric cancer completed	25	50%

Major Task 3: Publish gastric cancer results		
Discuss results with collaborators and perform any follow up analysis	26-27	50%
Write one or more manuscript(s) with input from designated mentor and collaborators	28-29	
Submit manuscript(s) and wait for reviews.	30-31	
Respond to reviewers and resubmit. May repeat submission/resubmission process with multiple journals depending on where the paper(s) end up being published. Present results at conferences.	32-36	
Milestone(s) Achieved: Manuscript(s) published	**36**	
Total time for Specific Aim 1	**36**	

What was accomplished under these goals?

1) Major activities since the last reporting period:

i) Fully completed specific aim 2 (100%)
ii) Completed the majority of specific aim 3 (75%)
iii) Completed half of specific aim 1 (50%)
iv) Published six papers related to this grant (including one with the PI as lead corresponding author)
v) One manuscript submitted (genomic aberrations in the TP53 pathway)
vi) One manuscript under preparation (integrated omics analysis of the TP53 pathway)

2) Specific objectives:

Specific Aim 2:

i) We completed the generation of pathway activity scores for 14 pathways across not just the Pan-GI data set, but across all 33 of the TCGA tumor types with 9125 samples. The 14 pathways and the data type used to generate them are:
   a. RTK (protein)
   b. PI3K/AKT (protein)
   c. TSC/mTOR (protein)
   d. Breast reactive (protein)
   e. RAS/MAPK (protein)
   f. Hormone receptor (protein)
   g. Hormone signaling (protein)
   h. DNA damage response (protein)
   i. Apoptosis (protein)
   j. Cell Cycle (protein)
   k. EMT (mRNA)
   l. TGF-β (mRNA)
   m. TP53 (mRNA)
   n. T-cell immune infiltration score (DNA methylation)

ii) We compared several of those pathways across the Pan-GI cancers to see which pathways were disrupted in gastric cancer.
iii) We correlated disrupted pathways across multiple data types (e.g. transcriptomic, proteomic, genomic, epigenomic) to determine the likely source of the disruption.
iv) We published the Pan-GI analysis results in Cancer Cell in April, 2018.

Specific Aim 3:
v) We published details of disruptions in gastric cancer (and PanCancer) across the following pathways: TGF-β, PI3K/AKT, RAS, DNA damage response, and immune response.
vi) Genomically aberrant genes within many of those pathways were identified as potential targets for therapy.

Specific Aim 1:
vii) We cleaned up all the TCGA data and completed analysis on them.
viii) Many hypotheses have been generated using TCGA data, and many potential pathways and aberrant genes have been identified. The hypotheses remain to be validated within the independent MD Anderson gastric samples in the last year of the project.

3) Significant results and key outcomes

There were many significant results published in our papers. Selected updated significant results and key outcomes since the last progress report are shown below.

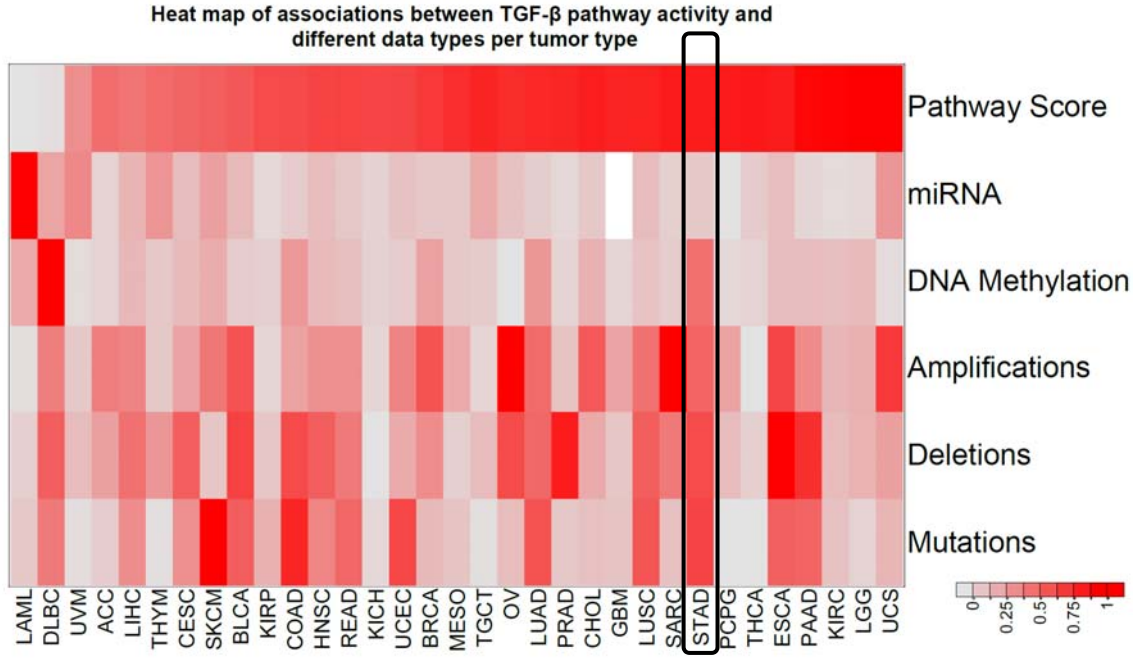

Figure 1. TGF-β pathway activity across 33 types of cancer in TCGA and the roles of different regulatory data types in them. The pathway activity was defined using mRNA levels of 43 pathway genes. Red color refers to high values, gray refers to low values. Gastric cancer (STAD) is highlighted by the black box.

Conclusion: Gastric cancer (STAD) has relatively high TGF-β pathway activity compared to other tumor types, especially colorectal cancer (COAD and READ), making the pathway a potential target for therapy. STAD has high levels of DNA methylation, CNVs and mutations, but not high levels of miRNAs that regulate the pathway.

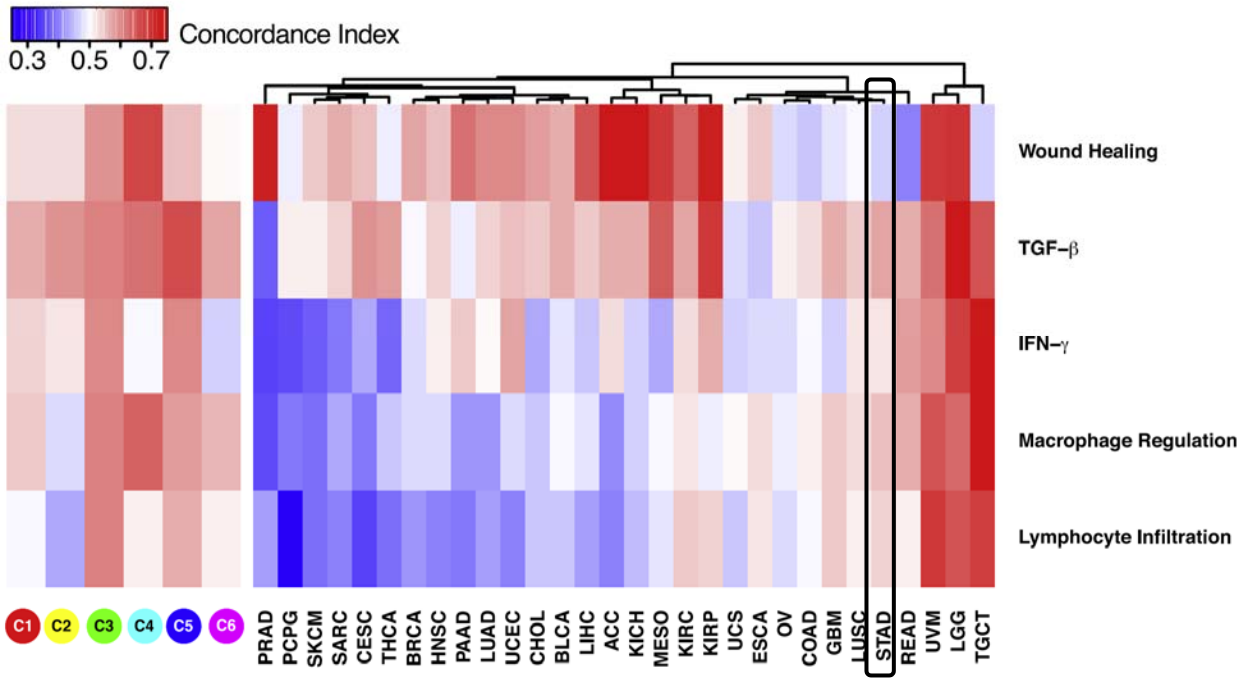

Figure 2. Immune response. Concordance index (CI) for five characteristic immune expression signature scores for immune subtypes in TCGA tumor types. Red denotes higher and blue lower risk, with an increase in the signature score. (From PMID: 29628290)

Conclusion: STAD has relatively high immune activity compared to other tumor types, but similar activity to colorectal cancer. Wound healing signature is low, but macrophage regulation, lymphocyte infiltration and TGF-β activity (as previously mentioned) are high. That raises interesting possibilities for using immunotherapy in gastric cancer.

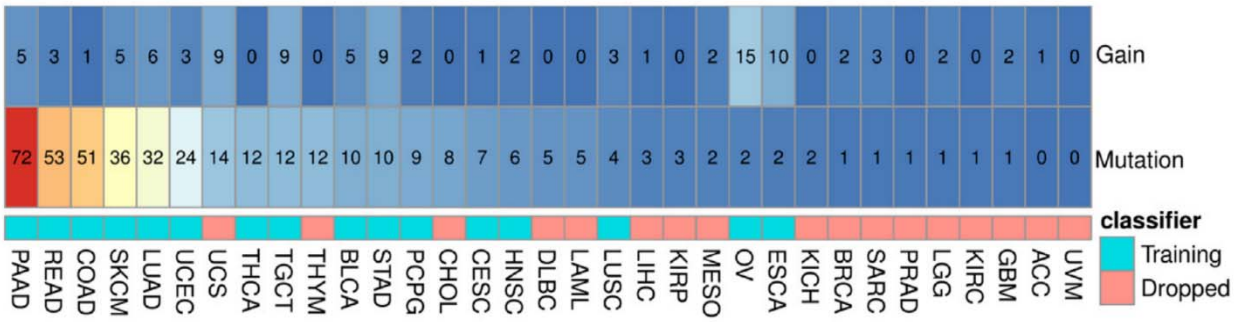

Figure 3. Cancer-type-specific percentages of Ras aberration by copy number gain and deleterious mutation in KRAS, HRAS, or NRAS. (From PMID: 29617658)

Conclusion: In TCGA gastric cancer samples, 10% of the samples had a mutation and 9% had amplifications in KRAS, HRAS and/or NRAS. Those patients may potentially benefit from RAS targeting drugs.

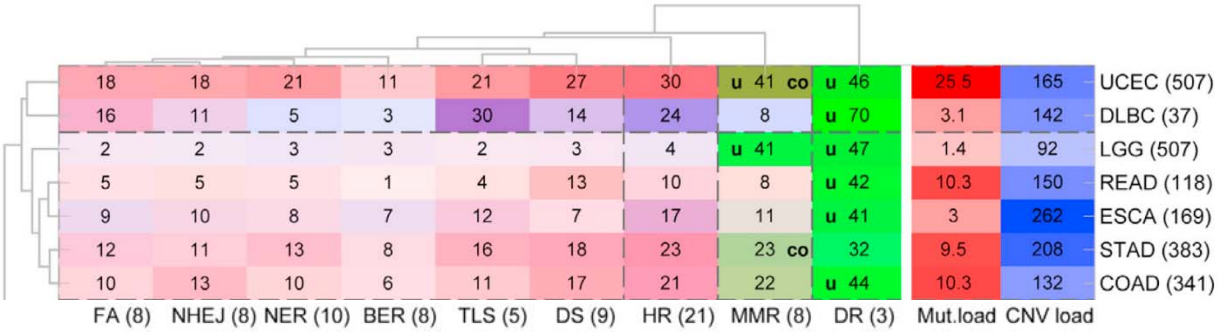

Figure 4. DDR gene alterations are frequent and non-uniformly distributed by type and frequency across cancer types. Clustered heatmap indicates the percentage (%) of samples in a cancer type (rows, with cancer types listed right, number of samples between parentheses) altered for at least one core gene in a given DDR pathway (columns, with core gene numbers indicated in parentheses for each pathway, bottom). Color intensity indicates the percentage altered, with the percentage given as a number in each cell. RGB color indicates mutations (red), deep deletions (blue), or epigenetic silencing through methylation (green). Gray scale indicates equal contribution from all three alteration types. A ''u'' symbol in cells indicates a statistically significant enrichment (FDR [false discovery rate] < 10%, difference in alteration percentages > 2%) in alterations. A ''co'' or ''me'' symbol in cells indicates a statistically significant (FDR < 10%) co-occurrence or mutual exclusivity of samples altered by mutation, deep deletion, or silencing. Only ''co'' relations were observed. The two rightmost columns, Mut. load and SCNA load, indicate average mutation frequency (non-silent mutations/Mb) and copy-number burden (number of copy-number segments) by cancer type. (From PMID: 29617664)

Conclusion: Gastric cancer has a high percentage of samples with alterations many DNA damage repair pathways, including direct damage reversal/repair (DR, 32%), mismatch repair (MMR, 23%), and homology-dependent repair (HR, 23%) pathways. Those pathways could potentially be targeted for therapy.

Stated goals not met
All goals were met in this reporting period.

What opportunities for training and professional development has the project provided?

Rehan Akbani, PhD (PI): The project has allowed Dr. Akbani several opportunities for professional development. He works with his mentor, Dr. Jaffer Ajani on a regular basis and updates him on progress. Dr. Ajani, in turn, guides Dr. Akbani's research. Dr. Ajani has also setup a larger group with about a dozen researchers who are working on gastric and other GI cancers. Dr. Akbani is part of the group and he has been afforded the opportunity to collaborate with those experts. The larger group meets on a monthly basis where researchers take turns in presenting their work and receive feedback. Dr. Akbani has also presented in that group.

Dr. Akbani is also one of the central members of The Cancer Genome Atlas (TCGA) project. TCGA funding ended in July 2016, but the consortium continued to work on PanCanAtlas and other projects through the end of September 2018. The funding from this grant has allowed Dr. Akbani to continue collaborating with TCGA on projects related to this grant, such as Pan-GI cancers, PI3K/AKT pathway disruptions, TGF-beta pathway disruptions, etc. In some of them (e.g. TGF-beta) Dr. Akbani plays a leading role. That would not have been possible without the kind of funding provided by this grant.

Dr. Akbani presented his research in the final TCGA symposium in September 2018 (poster sessions). He also attended the AACR 2018 conference in Chicago. As more of his research is published, he plans to present his results at more conferences and symposia.

Andre Schultz, PhD (postdoctoral fellow): Dr. Schultz is Dr. Akbani's postdoctoral fellow. Under Dr. Akbani's guidance and supervision, Dr. Schultz has gained hands-on experience in working with the TCGA data. He worked on many projects in TCGA, including the TGF-beta pathway project where he is co-author, and continues to work on more projects, like the TP53 pathway project. This grant has supported the training of Dr. Schultz.

How were the results disseminated to communities of interest?

Some of the research has already been published in renowned journals (see Products section). Other manuscripts are either currently under review, or in preparation. Besides publications, Dr. Akbani has presented the results of his research in a TCGA symposium in September 2018. As the research matures further, Dr. Akbani plans to attend more conferences and symposia to present his results.

What do you plan to do during the next reporting period to accomplish the goals?

1) Complete the remaining tasks and specific aims by using in-house gastric cancer data to validate the findings from the TCGA gastric cancer data set.
2) Publish the remaining results of the research in renowned journals.
3) Present the results at widely attended conferences and symposia.
4) Continue to participate in cancer conferences like AACR to improve knowledge of gastric cancer.
5) Collaborate with Dr. Ajani and his team of experts, and solicit regular feedback about the research.

IMPACT

What was the impact on the development of the principal discipline(s) of the project?

The research has resulted in the identification of many genes and pathways that could potentially be targeted for therapy in gastric cancer. Since clinical trials are expensive, the project has provided a focused short-list of promising targets to investigate further. Even though most of the publications resulting from this grant are less than one year old, they have cumulatively been cited 117 times already, according to Google Scholar. It is expected that the publications will continue to be highly cited and have a major impact on the field.

What was the impact on other disciplines?

Nothing to report.

What was the impact on technology transfer?

Nothing to report.

What was the impact on society beyond science and technology?

Nothing to report.

CHANGES/PROBLEMS

Nothing to report.

PRODUCTS

Publications, conference papers, and presentations

Journal publications (see appendix for details):

1) Liu Y, Sethi NS, Hinoue T, Schneider BG, Cherniack AD, Sanchez-Vega F, Seoane JA, Farshidfar F, Bowlby R, Islam M, Kim J, Chatila W, **Akbani R**, Kanchi RS, Rabkin CS, Willis JE, Wang KK, McCall SJ, Mishra L, Ojesina AI, Bullman S, Pedamallu CS, Lazar AJ, Sakai R; Cancer Genome Atlas Research Network, Thorsson V, Bass AJ, Laird PW. Comparative Molecular Analysis of Gastrointestinal Adenocarcinomas. Cancer Cell. 2018 Apr 9; 33(4):721-735. PMID: 29622466

2) Ding L, Bailey MH, Porta-Pardo E, Thorsson V, Colaprico A, Bertrand D, Gibbs DL, Weerasinghe A, Huang KL, Tokheim C, Cortés-Ciriano I, Jayasinghe R, Chen F, Yu L, Sun S, Olsen C, Kim J, Taylor AM, Cherniack AD, **Akbani R**, Suphavilai C, Nagarajan N, Stuart JM, Mills GB, Wyczalkowski MA, Vincent BG, Hutter CM, Zenklusen JC, Hoadley KA, Wendl MC, Shmulevich L, Lazar AJ, Wheeler DA, Getz G; Cancer Genome Atlas Research Network. Perspective on Oncogenic Processes at the End of the Beginning of Cancer Genomics. Cell. 2018 Apr 5;173(2):305-320. PMID: 29625049.

3) Anil Korkut, Sobia Zaidi, Rupa S. Kanchi, Shuyun Rao, Nancy R. Gough, Andre Schultz, Xubin Li, Philip L. Lorenzi, Ashton C. Berger, Gordon Robertson, Lawrence N Kwong, Mike Datto, Jason Roszik, Shiyun Ling, Visweswaran Ravikumar, Ganiraju Manyam, Arvind Rao, Simon Shelley, Yuexin Liu, Zhenlin Ju, Donna Hansel, Guillermo de Velasco, Arjun Pennathur, Jesper B. Andersen, Colm J. O'Rourke, Kazufumi Ohshiro, Wilma Jogunoori, Bao Ngyen, Shulin Li, Hatice U. Osmanbeyoglu, Jaffer A. Ajani, Sendurai A. Mani, Andres Houseman, Maciej Wiznerowicz, Jian Chen, Shoujun Gu, Wencai Ma, Jiexin Zhang, Pan Tong, Andrew D. Cherniack, Chuxia Deng, Linda Resar, The Cancer Genome Atlas Research Network, John N. Weinstein, Lopa Mishra, **Rehan Akbani**. A Pan-Cancer Analysis Reveals High-Frequency Genetic Alterations in Mediators of Signaling by the TGF-β Superfamily. Cell Systems. Sept. 26, 2018, e-pub ahead of print.

4) Way GP, Sanchez-Vega F, La K, Armenia J, Chatila WK, Luna A, Sander C, Cherniack AD, Mina M, Ciriello G, Schultz N; Cancer Genome Atlas Research Network (including **Akbani R**), Sanchez Y, Greene CS. Machine Learning Detects Pan-cancer Ras Pathway

Activation in The Cancer Genome Atlas. Cell Rep. 2018 Apr 3;23(1):172-180. PMID: 29617658

5) Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, Ziv E, Culhane AC, Paull EO, Sivakumar IKA, Gentles AJ, Malhotra R, Farshidfar F, Colaprico A, Parker JS, Mose LE, Vo NS, Liu J, Liu Y, Rader J, Dhankani V, Reynolds SM, Bowlby R, Califano A, Cherniack AD, Anastassiou D, Bedognetti D, Rao A, Chen K, Krasnitz A, Hu H, Malta TM, Noushmehr H, Pedamallu CS, Bullman S, Ojesina AI, Lamb A, Zhou W, Shen H, Choueiri TK, Weinstein JN, Guinney J, Saltz J, Holt RA, Rabkin CE; Cancer Genome Atlas Research Network (including **Akbani R**), Lazar AJ, Serody JS, Demicco EG, Disis ML, Vincent BG, Shmulevich L. The Immune Landscape of Cancer. Immunity. 2018 Apr 17;48(4):812-830. PMID: 29628290

6) Knijnenburg TA, Wang L, Zimmermann MT, Chambwe N, Gao GF, Cherniack AD, Fan H, Shen H, Way GP, Greene CS, Liu Y, **Akbani R**, Feng B, Donehower LA, Miller C, Shen Y, Karimi M, Chen H, Kim P, Jia P, Shinbrot E, Zhang S, Liu J, Hu H, Bailey MH, Yau C, Wolf D, Zhao Z, Weinstein JN, Li L, Ding L, Mills GB, Laird PW, Wheeler DA, Shmulevich I; Cancer Genome Atlas Research Network, Monnat RJ Jr, Xiao Y, Wang C. Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. Cell Rep. 2018 Apr 3;23(1):239-254. PMID: 29617664

Conference presentations:

1) A Pan-Cancer Analysis Reveals High-Frequency Genetic Alterations in Mediators of Signaling by the TGF-β Superfamily, poster presented by Rehan Akbani. TCGA Legacy symposium, September 27-29, Washington DC, USA.

PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Name	Rehan Akbani
Project Role	PI
Research Identifier (e.g. ORCID ID)	
Nearest person month worked	3.6
Contribution to Project	Led the project as PI. Performed analysis. Supervised the work of others.
Funding Support	

Name	Andre Schultz
Project Role	Postdoctoral Fellow
Research Identifier (e.g. ORCID ID)	
Nearest person month worked	6
Contribution to Project	Performed analysis under the direction of the PI
Funding Support	

Name	Yuexin Liu
Project Role	Co-Investigator

Research Identifier (e.g. ORCID ID)	
Nearest person month worked	3
Contribution to Project	Performed analysis under the direction of the PI
Funding Support	

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

No.

What other organizations were involved as partners?

Organization Name: The Cancer Genome Atlas (TCGA)
Location of Organization: NCI/NIH, Washington DC
Partner's contribution to the project: Collaboration (please note that funding from TCGA completed on 7/31/2017, before the reporting period started for this grant, so no funding was provided for this work by TCGA during the reporting period).