Dear Interns, Family, and Friends,

We can’t believe it is time to commemorate another successful summer of KEYS and our 2020 cohort of 49 extraordinary high school interns from across Arizona!

We are proud of each of them, and inspired by their accomplishments over the past seven weeks. This year’s virtual research experiences spanned a wide variety of scientific emphasis areas including data analysis and computational modeling related to cancer biology, neurodegenerative and cardiac diseases, biomedical engineering, environmental impact, toxicology, and COVID-19. Some interns will continue their research as incoming college undergraduates, and others will take all they have learned back to their high schools.

The contribution of volunteered time, talent, and treasure by research faculty at the University of Arizona is essential to the success of KEYS. They welcome interns into their labs and provide opportunities for them to explore and experience the excitement of scientific research and discovery. Additionally, our interns learn so much from the exceptional post docs, graduate and undergraduate students, and KEYS crew and staff who mentor them throughout KEYS and beyond. Over 100 people work together to make KEYS a success each year.

We sincerely appreciate the ongoing support of our generous program sponsors and donors. Without you, there would be no KEYS, and therefore, no opportunity like this for high schoolers to participate in world-class University of Arizona research before their college experience even begins. With our commitment to advancing next generation science and scientists, we are gratified to now count 526 students as KEYS alumni!

Congratulations to our 2020 KEYS class, and best wishes for what we know will be a bright future!

Sincerely,

Marti Lindsey, Ph.D.
Director, Community Engagement
Southwest Environmental Health Sciences Center

Brooke Moreno
Outreach Program Coordinator,
Public Affairs, BIOS Institute

Lisa Romero
Senior Director,
Communications and Public Affairs, BIOS Institute

Kelle Hyland
Outreach Program Coordinator,
Public Affairs, BIOS Institute
**Shirin Adang**  
BASIS Tucson North  
**PI - Dr. Michael Brown**

**Soft Matter Interactions with G-Protein Coupled Receptor Rhodopsin**  
G-protein-coupled receptors (GPCRs) are signaling proteins that play key roles in many biological processes. Dr. Brown’s lab analyzes the role of soft matter on the activation of GPCR rhodopsin, found in the retina. By exposing rhodopsin to varying osmotic pressures using different osmolytes, we observe hydration effects on rhodopsin’s favored conformation state. Molecular Dynamics simulations are used to visualize rhodopsin in its physiological conditions and UV-visible spectroscopy demonstrates whether the osmolyte favored the inactive Meta-I, or active Meta-II state of rhodopsin. By analyzing the response of rhodopsin to these osmolytes, we found larger osmolytes dehydrate rhodopsin, favoring its Meta-I state while smaller osmolytes favor the Meta-II state. We anticipate these results leading to new insights into the role of water in general GPCR activation mechanisms.

---

**Zarif Ahsan**  
BASIS Chandler  
**PI - Dr. Elizabeth Hutchinson**

**Developing a Digital Atlas of the Human Brainstem Cranial Nerves: A Diffusion-Based Approach**  
Vital for autonomic functions such as respiration, the human brainstem houses entry points and nuclei for ten of the twelve cranial nerves. These cranial nerve systems are affected in multiple congenital facial weakness disorders, yet the precise arrangement of their white matter tracts remains unknown. As such, we sought to create a digital atlas of the human brainstem cranial nerves using two post-mortem diffusion tensor images. We segmented fifteen regions of interest (ROIs) and performed tractography using ITKSnap and MRtrix software, respectively; doing so, we were able to create 3D models of the cranial nerve nuclei and fiber tracts. Additionally, we compared scalar metrics with descriptive statistics. We hope future researchers may use these results as both a reference and a means to detect abnormalities.

---

**Zachary Anderson**  
Paradise Valley High School  
**PI - Dr. Kaveh Laksari**

**Creating a Blood Mimic for Ultrasound Sonography**  
We created a carotid bifurcation phantom, an essential part of which, is a novel blood mimic. The goal in developing the blood mimic was to make it more accessible and affordable than other Doppler fluids and blood mimics. To be deemed viable, the blood mimic must match the International Electromechanical Commission’s standard properties for blood mimicking fluids, these being: the attenuation coefficient, viscosity, speed of sound, and density. Based on preliminary imaging results, we propose a blood mimic consisting of 90% water and 10% glycerol by percent weight. We need to conduct more tests, but previous literature indicates that this recipe will meet the IEC’s standards and its components are readily available and more cost-effective than other blood mimics or Doppler Fluids.
**Lauren Badinger**  
Palo Verde High Magnet School  

*The Role of Fatty Acid Elongase 5 (ELOVL5) in Human Cytomegalovirus (HCMV) Replication*

Human Cytomegalovirus (HCMV) is a herpesvirus that asymptomatically affects most people. In some cases, HCMV causes severe disease and birth defects. There is no cure or vaccine. The Purdy Lab studies how HCMV replication depends on lipids with very long chain fatty acid tails (VLCFAs). VLCFAs are made by one of seven human fatty acid elongases (ELOVL1-7). We hypothesize that ELOVL5 is important for the production of lipids with polyunsaturated VLCFAs induced by HCMV infection. To test this hypothesis, we used lipidomics to compare lipids in infected CRISPR/Cas9 ELOVL5-knockout cells and non-targeting (NT) control cells. We found that the production of VLCFAs significantly decreased in HCMV-infected ELOVL5-knockout cells. We conclude that HCMV induces ELOVL5 synthesis of lipids with VLCFAs to support virus replication.

**Avaneesh Bellampalli**  
BASIS Ahwatukee  

*Identifying Preliminary SARS-CoV-2 Epitopes by Analysis of Conservation From SARS-CoV-1*

In wake of the COVID-19 global pandemic, the development of a vaccine against SARS-CoV-2 is critical; identifying epitope regions within SARS-CoV-2 proteins is foundational to this. Our investigation leverages SARS-CoV-2’s compositional homology to its relative SARS-CoV-1. Using existing information regarding SARS-CoV-1 along with publicly-available databases, we compiled an index of 11 B-cell and 44 T-cell epitope candidates derived from SARS-CoV-1 and identically conserved within the SARS-CoV-2 spike glycoprotein. All have indicated a reactive immune response confirmed via positive B-cell, T-cell, or MHC assays. We mapped, with highlights, which residues are epitopic, revealing the densest regions of supposed-immunogenicity to be the focus of further investigation. This list aims to serve the efforts toward COVID-19 vaccine development and other investigations into conservation and cross-reactivity in related-viruses.

**Surajinder Bharaj**  
BASIS Ahwatukee  

*Visualizing Trends in SARS-CoV-2 Testing Data*

SARS-CoV-2 is a coronavirus strain causing the current COVID-19 pandemic. PCR is used to test for current infection and Serology for previous infection. Recent increased testing has produced substantial data. This study examines “per-person” and overall testing data in Arizona to identify testing strategy changes. I used R to create reproducible analysis on current testing data and generated visualizations of testing trends. Results show a clear difference in the pattern of “number of tests” when adjusting for multiple tests on the same person. Positivity rates were impacted when adjusting for multiple tests over the entire period of testing (through July 7, 2020). Positivity rates did not change for trends over the weeks. Therefore, most cases tested multiple times occurred longer than one week apart.
Megan Bime  
Catalina Foothills High School  
**PI - Dr. Frank Duca**

**Alterations in Small Intestinal Metabolomics Following Oligofructose Treatment in Obese Rats**

Type II diabetes is a metabolic disorder characterized by consistently high blood glucose and insulin resistance. The gut microbiome is key to regulating metabolism in the gastrointestinal tract and better understanding the development of type-2 diabetes and obesity. Metabolites, which are intermediate or end molecules formed in a metabolic process, were examined from the small intestine of rats for significant change between a high-fat diet, chow diet, and the addition of the prebiotic oligofructose. Through literary searches, metabolites that showed the most significant changes were further studied for their involvement in the metabolic process of the disease. This research can provide new hypotheses in future studies targeting specific metabolites towards potential novel therapies for the treatment of obesity and type-2 diabetes.

---

Natalia Bojorquez  
Nogales High School  
**PI - Dr. Pascale Charest**

**Study of the correlation between TBL1XR1 and mTORC2 expression and activity in Non-Small Cell Lung Cancer**

Non-small cell lung cancer (NSCLC) is the most deadly cancer type in the world. Once it metastasizes, or migrates to a secondary site, there are few treatments. Of interest in directed cell migration is the mechanistic Target of Rapamycin Complex 2 (mTORC2), a serine/threonine kinase. Transducin (β)-like 1 X-linked receptor 1 (TBL1XR1) has also been associated with metastasis in different types of cancers and is overexpressed in NSCLC tissue. I compared mTORC2 gene expression with TBL1XR1 gene expression in NSCLC patient samples to determine a role for TBL1XR1 in NSCLC cell migration. My analysis showed no statistical significance between mTORC2/TBL1XR1 expressions. My analysis could mean that TBL1XR1 is not involved in mTORC2 signaling or that it plays a role outside of transcriptional regulation.

---

Ivan Carrillo  
Nogales High School  
**PI - Dr. Elizabeth Hutchinson**

**Segmentation and Analysis of MRI Microscopy Data in the Bonnet Macaque Brain**

The Bonnet Macaque is an excellent non-human primate model. Previous research suggests that one technique utilized a histologically determined methodology in identical monkeys, in an in vivo (alive) MRI group template. Others demonstrated that the human and Macaque brain have similar structures, such as the hippocampus. Changes in microstructure could be associated with each specimen and can give knowledge regarding what a "standard" brain’s anatomical alterations are. My research segments regions of interests, creates 3D models, and identifies tracts in Macaque brains for qualitative and statistical analysis using two software programs called ITKSNAP and MRVIEW (MRtrix3). The desired outcome is to have structural differences in the diffusion tensor imaging when contrasting the specimens.
Sensor noise comparison to improve low signal detection of fluorescing nanoparticle

Nanophotonics is the study of how light interacts with nanometer scale objects. My lab studies how nanoparticles (particles on the order of nanometers) can improve the imaging capabilities of lens-free microscopes. Most imaging systems cannot image features smaller than 250 nm but many important biological features have sizes smaller than 250 nm. Imaging these features are important for disease detection and further understanding processes within living cells. Currently, the lab needs a camera sensor sensitive enough to image a test object, a single 200 nm fluorescing particle. To determine which camera to buy I used MATLAB to simulate images from different camera sensors and calculated each sensor’s signal-to-noise ratio (SNR). The camera with the highest SNR will be purchased and added onto a lens-free microscope.

Improving ESTIRA’s Graphical User Interface: Making Additions to the ESTIRA GUI using Tkinter

ESTIRA is an open-source, cost-effective multimodal stretching stage that is designed to characterize the mechanical and electrical properties of biomedical devices. ESTIRA will provide the scientific community with an accessible tool with which to develop biomedical devices. Through this project, the graphical user interface (GUI) of ESTIRA was improved by adding new features such as a new data entry field and real time data graphing. These changes were implemented by editing the existing ESTIRA GUI code using the Python coding language, existing Python code libraries and the ESTIRA GUI code scripts.

The Effect of Sleep Disruption on Dietary Constituents in Cafeteria diet-fed Rats

Obesity and sleep disruption (SD) affects women more than men. SD reduces micronutrient intake in adults, which may influence metabolic health. We hypothesized that SD would cause a greater decrease in micronutrient intake in females. Cafeteria diet-fed male and female Sprague-Dawley rats slept undisturbed or received SD. Body weight and calorie intake were determined daily. Micronutrients, amino acids (AA) and other dietary components were calculated and corrected for total calorie intake. SD increased sugar intake and overall, females ate more unhealthy components (branched chain AA and anti-nutrients) and empty (without nutrients) calories than males while sleep disrupted males ate fewer micronutrients and AA. Thus, SD had a sex specific effect on dietary items, which may have contributed to greater weight gain in sleep disrupted females.
Effect of Hydration on Rhodopsin Activation
Rhodopsin, responsible for vision under dim light, is a prominent member of the G Protein Coupled Receptor (GPCR) protein class which is targeted by a third of all pharmaceuticals. Two states of rhodopsin exist in physiological equilibrium: active and inactive. We hypothesized that hydration plays a crucial role in the shift to the active conformation. The fraction of active rhodopsin is quantified via UV-visible absorbance spectroscopy upon light exposure in the presence of various dehydrating osmolytes that generate osmotic pressure. We found that water does play an influential role in rhodopsin activation and that dehydration with osmotic stress favors the inactive state. As countless physiological processes involve GPCRs, greater insight into the relationship between water and rhodopsin activation enables treatments of GPCR dysfunction across human anatomy.

Identification of Alzheimer’s Disease-related pathology using post-mortem MRI maps
The purpose of this project was to identify Alzheimer’s disease (AD) pathology from post-mortem brain tissue scans via complex Magnetic Resonance Imaging (MRI) techniques. Prior AD pathology research shows a positive correlation between AD progression and involvement of the Locus Coeruleus (LC), the primary focus of this study. Employing the software ITKSNAP, we created 3D segmentations of the LC, Superior Cerebellar Peduncle, and Periaqueductal gray matter from a normal and AD scan. The segmentations were analyzed quantitatively and qualitatively, finding notable differences in the size and composition of the LC. With a deepened understanding of how MRI can identify AD pathology, our knowledge of the disease is heightened, bringing us closer to a cure.

Proteomic Analysis of sex-defined genome sequencing in Migraine
A migraine is a neurological condition which can cause numerous symptoms such as intense headaches, nausea, and vomiting. This project investigates the genes of both humans and rodents (male and female) and analyzes them by same sex/different species and same species/different sex. My hypothesis is as follows: The proteome differ both by species and between sexes. To obtain the data needed, an undergraduate student collected data experimentally and shared it via Excel. Data was sorted by deleting duplicate genes so that each gene appeared only once per category which was compared to determine overlap in the genes. Any overlap could help determine how migraines can be different between sex and species, but also how they could be similar.
**Aidan Gee**
Catalina Foothills High School

**PI - Dr. Daniela Zarnescu**

**Characterizing the Biochemical Properties of UG-Rich RNA in a Drosophila Model of ALS**

The Zarnescu lab is researching Amyotrophic Lateral Sclerosis (ALS), a neurodegenerative disorder that inhibits locomotor function causing paralysis. The lab uses Drosophila melanogaster (fruit fly) for their similar proteinaceous orthologs to humans. TDP-43 (TAR DNA binding protein) is researched due to its linkage to over 97% of ALS cases. TDP-43 aggregation can cause neuron death which ultimately leads to asphyxiation or heart failure. We aim to understand how specific sequences of DNA/RNA affect its biological properties in a Drosophila model of ALS. Results from RNA sequencing data sets indicate that uracil-guanine (UG) richness determines transcript solubility in a model of TDP-43 proteinopathy. Additionally, evidence suggests an association between TDP-43 enriched genes and ribosomes, leading us to formulate hypotheses about disease mechanisms underlying TDP-43 toxicity.

---

**Prasiddhi Gyawali**
Sonoran Science Academy

**PI - Dr. Ramona Walls**

**An R-Shiny app for trait data processing**

The Functional Resource for Environmental Studies (FuTRES) project works with vertebrate trait data from individuals with a goal to make data more accessible and interoperable. Since data are collected by different people using different methods, it can be hard for everyone to understand exactly what the data mean. To solve this problem, FuTRES works to write programs that can be used to easily standardize datasets. Standardization is done by getting the data from the CyVerse Discovery Environment and transferring them into RStudio where functions are applied. Code is written for a specific dataset then generalized so that it functions with other datasets. Code can always be extended or further generalized to fit another need. An R-Shiny app contains the code, providing a user-friendly interface.

---

**Rebecca Han**
Catalina Foothills High School

**PI - Dr. Dongkyun (DK) Kang**

**Analyzing Full Width at Half Maximum of Light Profile Microscopy Images to Determine Resolution**

Light profile microscopy (LPM) is a method that is able to experimentally calculate the resolution of images that other microscopes take. Dr. Kang’s lab developed a LPM device and used it to image pork belly samples under various illumination numerical apertures to see how thin the light sheet became, as this correlates with the resolution. I worked with these images to build a program that analyzes the light sheet width and found that as the light promulgated deeper into the tissue, the full width at half maximum (FWHM) followed the expected curvature. Features from the sample interfered with the FWHM program, and our efforts to minimize it were not as successful as anticipated. We are planning on finding alternative methods to decrease the background noise.
Gene Expression Analysis Shows Patterns Behind Immune Aging

Thymic involution has shown to be an important factor of age-related immunodeficiency. We wanted to describe genes that might be associated with age differences in thymic activity. Using reproducible methods in R, we improved our ability to understand gene expression patterns over time. An exploratory data analysis first evaluated the impact of the limit of detection (LOD), or the lowest analyte concentration that is reliably detected. When comparing genes to their LOD, there were two extreme differences: ADIPOQ had larger expressions within adults while HPRT showed larger expressions within young donors. In 18 genes with >60% of values above their LOD, our data showed significant decline in expression with age. These patterns may be faltering thymic activity, but more research needs to be done.

Hailey Johnson
Cienega High School

Role of PHLDA1 and FOXC2 in mTORC2 Signaling in Clinical Invasive Breast Carcinoma

The mechanistic Target of Rapamycin Complex 2 (mTORC2) promotes cell migration, cancer metastasis, and is involved in multiple signaling pathways. These components may inform strategies for detection and prevention of metastasis. Using publicly available databases, I identified PHLDA1 and FOXC2, two genes correlated with RICTOR expression - a component of mTORC2 - in invasive breast carcinoma. I investigated how their mRNA expression correlated with levels of RICTOR, along with AktpS473, and PKC-alpha, downstream components of mTORC2 signaling. I used Excel, and Prism to gather information and analyze data to see if my genes had statistical significance relating to mTORC2. I concluded that the only significance was that FOXC2 was correlated to RICTOR. It is possible the genes are involved in signaling pathways indirectly correlated to mTORC2.

Grace Harrington
Cactus Shadows High School

PI - Dr. Pascale Charest

Hailey Johnson
Cienega High School

PI - Dr. Bonnie LaFleur

Eshaan Joshi
BASIS Tucson North

PI - Dr. Tyson Swetnam

Abstract Unavailable
Automated physiological analysis of Engineered Heart Tissues from Human Induced Pluripotent Cells

Cardiovascular disease is responsible for one in every four deaths. To understand these deaths, cardiomyocytes can be generated from a patient’s blood by first converting blood into a stem cell state called induced pluripotent stem cells (iPSCs). Cardiomyocyte derived iPSCs (hiPSC-CMs) can be assembled into 3D constructs that resemble the heart called engineered heart tissues (EHTs). While it is possible to perform cardiovascular disease modeling in a dish, the tools to analyze the function of hiPSC-CMs remains laborious. To meet this need, our lab developed a software (MATLAB based) to quantify cardiac physiological parameters (peak amplitude, peak duration, and faster rising rate) in an automated and highly accurate manner. Further deployment of our software will be used to understand how mutations lead to cardiovascular disease.

Eli Lefkowitz
Catalina Foothills High School

PI - Dr. Jared Churko

Exploration of Neutralizing Antibodies Against COVID-19

The S protein of SARS-CoV-2 facilitates entry into cells by binding to the angiotensin-converting enzyme 2 (ACE2) receptor on the target membrane. The high similarity between proteins of SARS-CoV-2 and SARS-CoV, another coronavirus, makes it possible for SARS-CoV-specific antibodies to bind to SARS-CoV-2 proteins. This inquiry aimed to determine which antibodies neutralize SARS-CoV-2, and how they do so. Literature review was employed to compile antibodies, and epitope analysis and protein modelling to understand their neutralization mechanisms. The results of this inquiry showed that most of the antibodies targeted the receptor binding domain of the S protein. Not all of these antibodies blocked S-protein binding to ACE2. It also was shown that cross-neutralization between SARS-CoV and SARS-CoV-2 is rare, despite the close relation between them.

Esha Mathur
University High School

PI - Dr. Bonnie Hurwitz

Exploring the diversity of prophages associated with the cyanobacteria Nostoc

Phages, viruses infecting bacteria, are the most abundant organisms on Earth, and impacts all ecosystems. Phages in terrestrial environments are currently understudied. My project addresses this gap, exploring the diversity and biological functions of phages integrated Nostoc genomes (called prophages). Nostoc is a photosynthetic cyanobacteria found free-living in soil and freshwater or in various symbiotic relationships in lichens and mosses. Using dedicated bioinformatics approaches, we identified more than 500 prophages, the majority of which are novel. A surprising finding in our work was that these prophages carry genes involved in a variety of cellular functions that may help their hosts in their environment. We demonstrate that there is still much to be explored to increase on the diversity and functions of prophages within terrestrial ecosystems.
**Nik Mathur**  
University High School  

**PI - Dr. Pascale Charest**

**ZDHHC11 and its correlation with mTORC2 controlled metastasis of Non-Small Cell Lung Cancer**  
The Charest lab aims to understand how cells migrate in response to external signals involved in guiding cancer cells to spread during metastasis. We study a protein complex, mTORC2 that was implicated as a key signaling step in metastasis. My research focused on studying other proteins that could correlate with the role of mTORC2 in Non-Small Cell Lung Cancer (NSCLC) cell migration. Using publicly available cancer genetic tissue data, I determined biostatistical relationships between the gene ZDHHC11, the mTORC2 complex, and cell migration, as measured by cancer staging. My analysis yielded no significance between ZDHHC11, mTORC2 component RICTOR nor NSCLC staging. However, ZDHHC11 was statistically correlated to the AKT S473 phosphorylation pathway which is associated with Rictor as mTORC2 is the primary AKT S473 kinase.

---

**Sonia Mathur**  
University High School  

**PI - Dr. Leif Abrell**

**What's In Your Rainwater? Inorganic and Organic Contaminants Measured in Roof- Harvested Rainwater**  
Rainwater harvesting is a vital component of water resources for many, but there is concern about its safety for human use. Project Harvest was co-created from a needs assessment that revealed this concern in Arizona communities. Using rainwater samples collected by citizen scientists around Arizona and a Home Description Survey (HDS) about construction materials at participants' homes, Project Harvest investigates the contaminants found in harvested rainwater to understand what influences the quality of harvested rainwater, and if it is safe for human, plant, and soil health. Using Excel and Rstudio to analyze data for lead, zinc, and 4-Nonylphenol, we constructed visualizations to show the influence the sampling season, sampling year, and sampling community has on the concentration of contaminants in harvested rainwater and its quality.

---

**Matthew Mayer**  
Walden Grove High School  

**PI - Dr. Melanie Culver**

**Mesocarnivore Ecology in the Santa Rita Mountains**  
My study examined a diverse group of smaller predators; mesocarnivores, in the Sky Islands. Hosting a wide range of species hailing from both the Rocky Mountains in the north and tropical climates to the south. The goal of my project is to analyze remote trail camera photographs taken in 2015-2016 as part of a Jaguar and Ocelot detection and monitoring project that began in 2011. I then analyzed the data for mesocarnivore occurrences. I expect my data to reveal behavioral patterns that reduce competition, and intraguild predation - predation of competitors, and that some species of mesocarnivores may prefer different habitats. This new information can assist range managers and wildlife conservationists in the Sky Islands with management and conservation actions of these species.
Using Biocuration to Validate Novel Predicted Disease-Pathway Associations and Improve Statistical Power for Dysregulated Pathway Analysis

Identifying associations between dysregulated pathways and disease phenotypes can establish targets for precision medicine treatments. The lack of optimal antifungal treatment for valley fever gives rise to the necessity for identification of novel biomolecule targets. The existence of predicted relationships between valley fever and various pathways were validated via curation of evidence from literature. Consensus curation and interrater agreement statistics were used to authenticate these results. The N-of-1-pathways MixEnrich framework is a model that uses transcriptome data to identify dysregulated genetic pathways related to phenotypic change in a single patient. Metabolite identifiers from various databases were curated to extend this framework to predict dysregulated metabolic pathways based on serum metabolite data. These can then be targeted for potential personalized therapeutic treatments.

Pulmonary Arterial Hypertension and Right Ventricular Function: R Shiny App Development

Pulmonary Arterial Hypertension (PAH) is a disease that is defined by high blood pressure in the arteries of the lungs. Right Ventricular (RV) function is a determinant of mortality in PAH. Our research aims to identify new markers of RV dysfunction. We hypothesized that the creation of an R Shiny App would improve the interpretation of RV hemodynamics. MATLAB was used to analyze cardiac waveforms, and new markers and pressures were recorded. Similar analysis was run for multiple conditions. R Shiny was used to interactively plot this data from a REDCap report. New plots were made to see how subjects’ pressures varied with different parameters. This app will efficiently aid in further interpretation of hemodynamics, expanding knowledge on how PAH and therapeutics impact RV function.

Genetic Association Analysis in the Development of Diabetes

Diabetes is a disease in which the body’s capability of reacting to the hormone insulin is defective. Previous complex disease research suggests that genetic elements are vital in their development; thus, the purpose of this research is to estimate the genetic component in the development of diabetes in order to better understand the disease and advance its treatments. We calculated genetic relationship matrices (GRMs) of UK Biobank data via Genome-wide Complex Trait Analysis (GCTA) to demonstrate the genetic relationship of each pair of individuals in our sample. Following the analysis of these phenotypes, their heritability was estimated before adjusting the results for the population prevalence. With the knowledge gained from this research, the goal of improved treatment for this disease becomes well within reach.
**Margot Nicholson**  
*Catalina Foothills High School*  
**PI - Dr. Michael Marty**

**Automating Data Visualization Processes for UniDec**

Native mass spectrometry (MS) is an important tool for characterizing proteins. However, data from native MS can be difficult to interpret, which discourages non-specialists from engaging with literature of the field. UniDec is an application that analyzes MS data. My project aimed to make a system in Unidec that allows users to quickly create data visualizations, increasing throughput and engagement with non-specialist audiences. To create this system, I used Python to edit the UniDec source code. I made individual plotting scripts, and then integrated them into UniDec. The resulting system had two styles of pre-formatted plots that could be created by clicking on a data file. The project succeeded in producing a plotting system for UniDec, helping users share their findings more quickly and easily.

---

**Christina Niyigena**  
*Sahuaro High School*  
**PI - Dr. Michael Marty**

**Automating a User-Friendly Interface for Faster Deconvolution of large MS Data Sets**

Native mass spectrometry (MS) has become a useful tool for studying protein structure and protein-ligand interactions. In native MS a sample is ionized and converted into gas-phase ions, while retaining non-covalent interaction, these ions are then separated according to their mass to charge (m/z), which produces a complex mass spectrum. The Marty lab created the UniDec software to lower the complexity of the mass spectrum by converting m/z to mass. However, UniDec lacks a high throughput method of importing mass spectra, which would be useful to analyze large data sets. A functional script was created that imported data from excel files with a user-friendly interface that can extract the needed values.

---

**Sreevatsa Nukala**  
*BASIS Chandler*  
**PI - Dr. Ali Akoglu**

**Overcoming the Limitations Posed by TCR-beta Repertoire Modeling through In-Silico DNA Recombination Algorithm**

The immune system detects and neutralizes antigens using T-cell receptors (TCRs). The V(D)J recombination is the primary mechanism for generating a diverse TCR repertoire essential for detecting a variety of antigens. We model the entire mouse TCRß repertoire at 10^15 scale using two GPU implementations, which counts the number of unique pathways to create each in-vivo sequence. The first implementation reduces the time scale of TCR synthesis from 50 weeks scale on a general-purpose processor to ~1.7 days, and its memory-optimized version reduces this further by ~2.3x. Our preliminary pathway profiling demonstrates that there is a high degree of correlation between n-nucleotide length and the number of pathways. This motivates further research into understanding the relationship between V, D, J sequences, and n-nucleotides.
Elissa Ornelas  
Walden Grove High School  
PI - Dr. Michael Worobey

Impact of ‘Essential Services’ Maintained in Rural Communities on COVID-19 Disease Caseloads

The RNA virus, ‘Severe Acute Respiratory Syndrome Coronavirus 2’ (SARS-CoV-2) which results in the COVID-19 disease, has disseminated amongst ‘essential services’. The lab hypothesizes that: meat processing facilities have no direct ramifications on COVID-19 caseloads within rural communities, and there is no statistical difference in caseloads based on the animal that the facility processes. ‘A counties’ with facilities and ‘B sister counties’ without facilities are analyzed simultaneously in Excel. Results do not suggest that an increase in population is correlated with elevated case documentation. In various ‘A county’ data, there was a statistically valid increase in mean caseloads within ‘A counties’ as compared to ‘B counties’. Ultimately, meat processing facilities appear to influence the spread of the virus in extra-urban counties dependent on ‘essential services’.

Karen Sanchez Perez  
Walden Grove High School  
PI - Dr. Anne Wertheimer

Exploring the Role of Cytokines as Biomarkers through the Establishment of their Standard Concentrations in Healthy Individuals

Cytokine is the general term for a group of secreted molecules involved in cell signaling during immune responses. Cytokines, including interleukins and interferons, possess various pro and anti-inflammatory effects, as well as several other functions through numerous biochemical pathways and interactions. The purpose of this research is to explore the role of cytokines as biomarkers by determining what a normal concentration should be in a healthy patient. This will allow further research to explore cytokine-disease associations for diagnostic purposes. I focused on IL-6, IL23 and IL17 concentrations as they correlate to diseases such as Osteoporosis, Encephalitis and Breast Cancer. The blood samples obtained from healthy subjects over time are analyzed in the Multiplex Assay system in order to identify and quantify the cytokines present.

Suraj Puvvadi  
BASIS Scottsdale  
PI - Dr. Yves Lussier

Validation of Predictions From a Novel Small-Sample Statistic of Valley Fever and Dysregulated Pathways and Increasing Statistical Power of Metabolite Identifiers Through Biocuration

Precision medicine is an approach to patient care that optimizes the efficiency and personalizes treatments based on transcriptomic, genomic, and metabolomic profiling. Rigorous curation of high throughput data is the first step for precision medicine. This project focuses on comparing a chronic form of valley fever, disseminated coccidioidomycosis (DCM), to unbiased molecular pathways through biocuration techniques. Through biocuration of metabolite databases and medical literature, we’ve heuristically evaluated and annotated the knowledge-based component of the N-of-1 pathways tool that consists of metabolites and metabolomic pathways. Also, we validated predictions from a novel small-sample statistic by unveiling significant dysregulated pathways related to DCM through medical curation. Overall, this project emphasizes the importance of biocuration techniques and is the first step in producing individualized treatments to heal humanity.
Sienna Rhoda  
Palo Verde Magnet High School  
**PI - Dr. Ted Trouard**

**Understand Focused Ultrasound in Alzheimer’s Disease Mouse Models**
This project’s focus is on possible treatments of Alzheimer’s (AD). The biggest obstacle for scientists in treating AD is delivering therapeutics to the brain, through the blood brain barrier (BBB). In this study the barrier was forcefully and safely opened through a combination of Magnetic Resonance guided Focused Ultrasound (MRgFUS), and Microbubbles in AD mouse models. With the help of a contrast agent, Gadolinium, the breaks in the barrier can be studied through MRI. Medical imaging software uses these images to determine the volume and intensity of the BBB openings. This study was successful and shows consistent openings of the BBB in all 11 AD mouse models. Through the data collected the lab is able to further determine the efficiency of MRgFUS.

Isabel Ross  
Cienega High School  
**PI - Dr. Laura Meredith**

**Green Infrastructure Impacts on Carbon Cycling: Evaluating Changes in Soil Microbial Composition and Function**
Green Infrastructure (GI) redirects water into the soil, affecting soil microbial diversity, decomposition, and stabilization of carbon. Soil microbes drive decomposition and carbon stabilization, making them important for carbon cycling. Decomposition by soil microbes releases carbon dioxide, a greenhouse gas, into the soil and atmosphere which is important for climate change. This project focuses on how GI systems change soil microbial community composition and function to understand how GI affects the carbon cycle. We used a phylogenetic gene marker and FAPROTAX to identify bacteria taxa and function. We observed changes in microbial composition and key functions with rainwater harvesting techniques. Although decomposition also creates more nutrients for carbon-absorbing plants, increased decomposition observed in this study could lead to additional sources of greenhouse gases.

Kelly Rushford  
Walden Grove High School  
**PI - Dr. Laura Meredith and Malak Tfaily**

**Volatile Organic Compounds: The Missing Link to the Metabolome**
Are volatile organic compounds missing from data of soil metabolism? Soil bacteria are necessary for plant growth and nutrient cycling foundational to ecosystem health. They release volatile organic compounds, or ‘VOCs’, which they use as communication signals and energy reservoirs to enhance the stability of their community. The functional relationship between soil organisms and VOCs is opportune for noninvasively profiling the soil community by measuring all of the molecules in a soil sample. In these “-Omics” sciences, data provide comprehensive insight into the soil environment, but VOCs may be too complex to be represented by the tools used to detect metabolites. The challenge of finding VOCs in the metabolome is analyzed using NMR, GC-MS, and FTICR on a peatland sample.
Mykaela Salvacion  
BASIS Oro Valley  

**Effects of Radiation Therapy on Calcium Signaling in the Salivary Gland**  
In 2017, head and neck cancers affected over 53,000 people and led to 10,800 deaths in the United States. While radiation therapy can improve patient survival, it can lead to loss of salivary gland function, resulting in expensive and painful dental issues. Little is known regarding exactly how radiation therapy affects the metabolic pathways in salivary glands, specifically calcium signaling involved in salivary excretion. Using MetaboAnalyst, I performed a metabolomic analysis of metabolite data extracted from untreated and irradiated mice to determine metabolite profile and pathway differences. My results suggest a downregulation in calcium signaling following radiation and warrant further investigation into this pathway and significantly enriched pathways that may provide targets to restore salivary function and improve patient quality of life.

Kyra Singh  
BASIS Tucson North  

**Using a 4-State Model to Simulate Altered Calcium Binding Properties of HCM-Associated cTnT Mutations**  
Hypertrophic cardiomyopathy (HCM) is a genetic disease that affects the heart’s muscle and inhibits its ability to relax well. It affects 1/500 people worldwide, and is the leading cause of sudden cardiac death in young adults. HCM can be caused by mutations that alter the ability of Troponin C (TnC) to interact with calcium, and a common mechanism thought to explain this involves altered interactions between Troponin I (TnI) and TnC. This project applies a mathematical 4-state model of TnC activation to simulate the altered calcium binding properties of healthy and diseased cardiac thin filaments. The results successfully simulated mutation-specific differences in calcium dissociation by modulating the effective concentration of TnI. However, to model changes in calcium sensitivity, another variable must change.

Yasmine Sleiman  
Gilbert Classical Academy  

**Large-Scale Data Analysis of Diagnostic Expression in Children with Autism Spectrum Disorder (ASD) from Electronic Health Records (EHRs)**  
While 1 in 54 children are diagnosed with Autism Spectrum Disorder (ASD), the CDC found the diagnostic process for ASD to be increasingly delayed in the U.S. We are looking to see if females exhibit similar diagnostic criteria trends as males and how these trends vary in children with and without ASD. Using Python, the diagnostic criteria frequencies are calculated. Manual matching of electronic health record (EHR) descriptions to the twelve criteria of the Diagnostic and Statistical Manual of Mental Disorders (DSM) identify the specific criteria children with ASD fall under. Children with ASD have 59.91% higher criteria frequencies regardless of subgroup. Females held significantly lower frequencies. The program depicts how various factors affect the prevalence of ASD much more accurately and efficiently.
Analyzing Cytokine Levels with Reproducible Statistical Methods

Cytokines are proteins found in the blood’s matrix (any form of plasma or serum). Previously, our lab studied 21 cytokines and determined that out of five matrices, Leucosep plasma showed cytokine levels most similar to serum, which is the only matrix not collected with anticoagulants. In my project, I continued to study the levels of five cytokines (GM-CSF, IFN-γ, IL-1ß, IL-6, and TNF-α) in Leucosep and Serum at more timepoints. My project aimed to determine how Quality Control Charts compare with Bland-Altman analysis for agreement and if the equivalence between Leucosep and Serum hold over time in all cytokines. The cytokine data was obtained through multiplex immunoassay technology. The data was analyzed using RStudio to make Quality Control graphs, Bland-Altman plots, and scatterplots.

Evaluation of Focused Ultrasound-Enhanced Treatment for Parkinson’s Disease

Parkinson’s Disease is the second most prevalent neurodegenerative disorder in the world, affecting more than 10 million people. It is characterized by a progressive loss of dopaminergic neurons due to a toxic buildup of the α-synuclein protein in the brain. While therapies for Parkinson’s exist, treatment for the disease remains a challenge as the protective barrier that surrounds the brain prevents therapeutic agents from entering. MRI-guided focused ultrasound therapy is a minimally invasive technique that temporarily opens the barrier for treatment. Using image analysis software, my project evaluated focused ultrasound therapy in model mice by quantifying the volumes and intensities of the brain areas opened. My results demonstrated that this low-risk technique overcomes a major challenge for treating Parkinson’s.

Modeling 2D Material Structures

This lab, Modeling 2D Material Structures, studies properties in 2D materials by exploring plasmonic devices, analyzing the SPPs, surface plasmon polaritons, for how the light propagates from metal waveguides. This summer, I focused on the problem of short propagation lengths. The hypothesis is that metal waveguides will produce a longer propagation length if they are thinner. I used a software called Lumerical, a computer program for high-performance photonic simulation, to simulates lasers on different gold samples and test that hypothesis. The results I’ve reached so far don’t show enough of a pattern to claim if I’ve reached the goal of the theory or not. However, I do have hope from some of the simulations that this hypothesis is correct, but continued research is required.
Douglas Swango  
BASIS Tucson North

**Investigating the Role of NIPBL in the mTORC2 Signaling Pathway of Breast Cancer Cells**

In my research, I studied the mechanistic target of rapamycin complex 2 (mTORC2) signaling pathway in breast cancer by analyzing if certain proteins influence metastasis. Prior research has implicated mTORC2 in cell migration in both unicellular organisms and mammals, leading researchers to believe there may be relationships between this pathway and cancer spread. For my research, I searched through databases and primary literature, looking for genes with potential connections to cancer metastasis. I chose the cohesin loading factor nipped-B-like protein (NIPBL) to study, and then analyzed various bioinformatics in the data using statistical tests to determine if correlations existed. I found that NIPBL expression was strongly correlated with the expression of an mTORC2 component, but not with cancer spread or mTORC2 activity.

Logan Taber  
Palo Verde Magnet High School

**Evaluating Essential Services in the Spread of COVID-19**

The virus SARS-CoV-2, which causes the disease known as COVID-19, is the coronavirus responsible for the global pandemic we are currently facing. As a coronavirus, COVID-19 mainly attacks the upper and lower respiratory tracts. COVID-19 is a serious threat, but the more we know and understand about it, the better we can control the virus and protect ourselves from it. That is why the Worobey lab posed the question: Does heavy-reliance on “essential services” in rural counties affect the spread of COVID-19 there? We compared “sister” counties with populations of less than 150,000 people. One county has meat processing facilities, the other does not. Using basic statistical analysis on the data we collected, we determined that, yes, essential services affect the spread of COVID-19.

Nicole Dennis Talley  
BASIS Oro Valley

**Study of Small Molecules Targeting Superoxide Dismutase (SOD1) in Familial Amyotrophic Lateral Sclerosis (ALS) Using Chemical Bioactivity and Protein Databases**

ALS (amyotrophic lateral sclerosis) is a neurodegenerative disease that results in muscle atrophy and eventual death of the patient with life expectancy from a few months to 2 years. One of the main causes of ALS pathology are mutations in different proteins leading to protein aggregation and loss of function in neurons. Over the summer, my laboratory and I focused on studying small molecules targeting mutant superoxide dismutase 1 (SOD1) related toxicity. We studied structure activity relationships in small molecules that were tested for effect on SOD1 indifferent bioassays. We identified that a group of antibiotics called cephalosporins and also, small molecules arylphenyl pyrazolones can be used to identify a lead pharmacophore for further development and testing.
Texture Analysis of Fluorescence Images for Classification and Prediction of Ovarian Cancer Risk

Ovarian Cancer is the deadliest gynecological disease. In my lab, I performed research on images of the fallopian tube fimbria, where ovarian cancer is believed to start. Hopefully, a trend or connection between the fluorescence properties of tissues and their patient categories can be seen, to diagnose cancer earlier. This is the first specific research being done. We are questioning, using fluorescence imaging and texture analysis, can we see if cells are cancerous and diagnose the patient non-invasively. We are using ImageJ to create stencils of the fluorescing tissue, MATLAB to conduct texture analysis on the tissue using the stencils, and Excel to display the results in a table and a line graph, to show how texture measures may differ in health and disease.
2020 ACKNOWLEDGEMENTS

KEYS PERSONNEL

Dr. Marti Lindsey
Brooke Moreno
Dr. Nadja Anderson
Cindy Bujanda

KEYS CREW

Lisa Romero
Kelle Hyland
Jordan Pilch
Marissa Romero

Ryan Hunt
Robyn Pratt
Uwe Hilgert
2020 KEYS PRINCIPAL INVESTIGATORS

Dr. Leif Abrell
Dr. Ali Akoglu
Dr. Jennifer Barton
Dr. Michael Brown
Dr. Pascale Charest
Dr. Jared Churko
Dr. Melanie Culver
Dr. Frank Duca
Dr. Vijay Gokhale
Dr. Philipp Gutruf
Dr. Bonnie Hurwitz
Dr. Elizabeth Hutchinson

Dr. Dongkyun Kang
Dr. Bonnie LaFleur
Dr. Kaveh Laksari
Dr. Tally Largent-Milnes
Dr. Gondy Leroy
Dr. Kirsten Limesand
Dr. Yves Lussier
Dr. Michael Marty
Dr. Euan McLeod
Dr. Laura Meredith
Dr. John Purdy
Dr. John Schaibley

Dr. Tyson Swetnam
Dr. Jil Tardiff
Dr. Jen Teske
Dr. Malak Tfaily
Dr. Ted Trouard
Dr. Rebecca Vanderpool
Dr. Romona Walls
Dr. Anne Wertheimer
Dr. Michael Worobey
Dr. Zhongguo Xiong
Dr. Daniela Zarnescu
Dr. Jin Zhou

2020 KEYS MENTORS

Dillion Aberasturi
Maryam Baker
Meghan Balk
Ryan Bartelme
Kai Blumberg
Vanessa Buzzard
Yuliang Chen
Youens Clark
Courtney Comrie
Laurel Dieckhaus
Zachary Frankman
Steven Fried
Yang Gu
Md Sahil Hassan
Linnea Honeker
Ally Janowski

Colleen Kenost
Matthew Klass
Ian Kline
Marius Kostelic
Nirmal Kumbhare
Erik Lehmkuhl
Jianrong Li
Joshua Mack
Susan Malusa
Helen Mann
Dan Meade
Lauren Meeks
Lauren Melcher
Nicholas Mortimore
Devin Murphy
Patrick O'Neal

Kunal Palawat
Haerin Park
Roberto Peralta
Ingrid Peterson
Alise Ponsero
Nima Pouladi
Ricky Cordova
Neeka Sewnath
Harrison Thurgood
Ricky Cordova
Thomas Watts
Alyssa Werner
Mollie Wiegand
Liam Wilson
Yuecheng Xi
Samir Rachid Zaim
2020 KEYS SUPPORT TEAM

Ken Addison
Nadja Anderson
Sean Cadin
Ricky Cordova
Melanie De Sa
Xing Ding

Elliott Cheu
Tom Keating
Julie Ledford
Kirsten Limesand
Bob Mittan
Regina Ozaeta

Robyn Pratt
Dana Rothstein
Sarah Smallhouse
Brenda Steeber
Patricia Wheeler

SPECIAL THANKS TO 2020 KEYS DONORS

Kristen Alicea-Jorgenson
The American Online Giving Foundation
Anonymous Foundation
Arizona Foundation for Educational Advancement
Arizona Women's Forum
Jennifer K. Barton
Carlton Benson
Biotech Calendar
Thomas R. Brown Foundation
Brian M. Ellerman and Laurie Eccard-Ellerman
Fidelity Charitable Gift Fund
Kevin Hall
Uwe Hilgert
Connie Hillman Family Foundation
Megan Hughes
Ryan Hunt
Daniel and Lisa Jurkowitz

Mr. and Mrs. Tom W. Keating
Juanito Manaloto
Shelley McLaughlin
Laura Meredith
John Moore
Kelly Myrdal
Mr. and Mrs. Jimmy W. Nokes
Dan Nguyen
Susan Petrus
Peter Powers-Lake
Anish Raju
Marissa Romero
Mark and Riki Sadosky
Anthony M. Webber
Helen Wertheim
Gena Wilson
Liam Wilson

2020 KEYS SPONSORS

If you are interested in supporting the KEYS Research Internship Program, contact keys@bio5.org or 626-2465.

KEYS.ARIZONA.EDU
KEYS 2020

49 INTERNS
260 APPLICANTS

3.93 AVG GPA UNWEIGHTED

25 HIGH SCHOOLS ACROSS ARIZONA