ANNUAL REPORT 2021

July 1, 2020 - June 30, 2021

UNIVERSITY OF CALIFORNIA
SANTA CRUZ | Genomics Institute

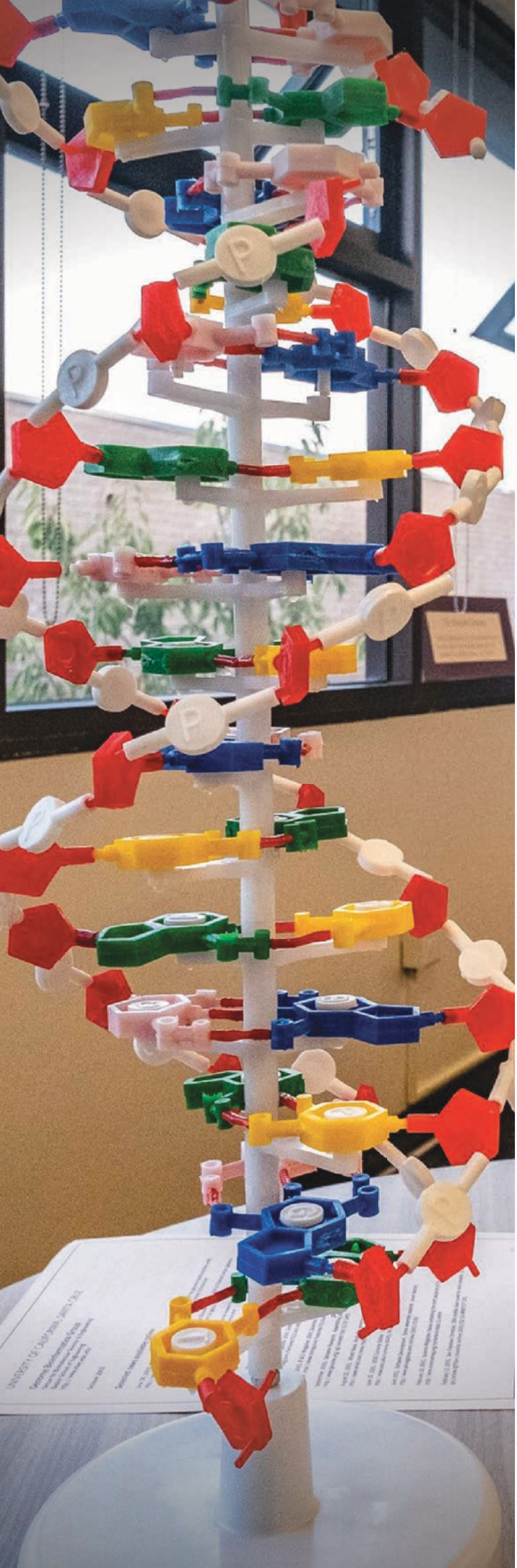


TABLE OF CONTENTS

CELEBRATING 20 YEARS OF GENOMICS AT UC SANTA CRUZ

DIRECTOR'S LETTER

The promise of genomics to effect positive change has never been greater. The Genomics Institute is proud to be on the frontline of innovation.

We are responding to the pandemic and preparing for future pathogen threats, applying our computational genomics expertise to sequence the virus and develop a public health-led bioinformatic platform for the State of California, with utility for national and international sharing of pathogen data. Our virus browser, which grew from our internationally recognized UCSC Genome Browser, is now tracking the COVID-19 virus sequence as it transmits and mutates, thereby helping public health, medical and science experts accelerate testing, vaccine research and treatments to control further spread. The UCSC Genome Browser, now celebrating its 20th year, has become the most accessed browser in the world and a flagship to help researchers reveal life's code.

Twenty years ago, as part of the Human Genome Project, UCSC scientists succeeded in publishing the first draft human genome sequence to the Internet. Our tradition of promoting data sharing is unwavering, not only in pathogen genomics but as a thread tying our work together: Through our largest project, the Computational Genomics Lab & Platform, we are connecting the world's biomedical data into a cohesive data biosphere.

With the Treehouse Childhood Cancer Initiative, PanCancer Consortium leadership, and the BRCA Exchange, we are leveraging genomics and data sharing to attack cancer; we are now expanding our cancer efforts with further types of data and diagnostic analysis.

Live Cell Genomics is our newest project and passion. By connecting traditional wet lab biology with automated systems on the Internet of Things to enable distributed experimentation from around the world, we are developing new avenues to understand rare diseases and undiagnosed diseases, advancing cancer research, and bringing dramatic new capabilities to neuroscience. Long-read nanopore DNA and RNA sequencing is one of the many new technological advances fueling our discoveries. Our conservation genomics scientists are using ancient DNA to understand extinction and protect species; they are using environmental DNA to monitor biodiversity in ecosystems to better understand how communities respond to impacts including agriculture, urbanization, and fire.

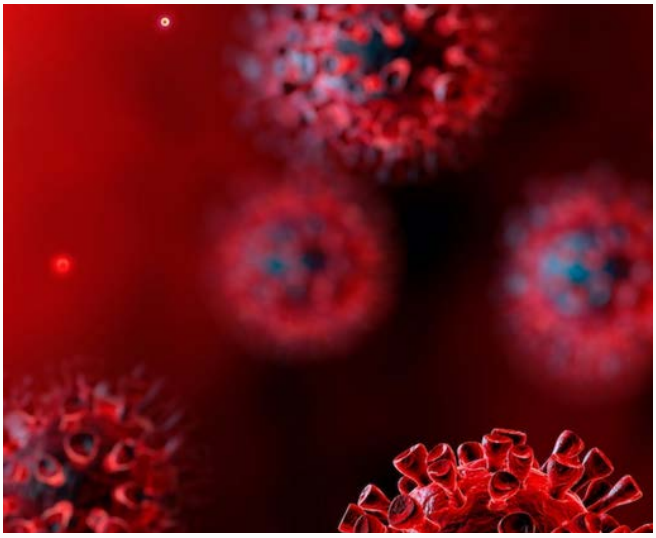
Mindful of the need to address inequities in traditional research, we are leading an effort to create a new pan-human genome reference, one that better represents human diversity. We are pioneering technologies to allow students in schools with less access to resources to learn science through remote experiments they can control using a cell phone. Our scientific commitment to genomic diversity is matched by our interest in furthering diversity, equity and inclusion at all levels of our institution. In our work today and in our strategy for the future, we embrace the work of improving educational opportunity, amplifying diverse voices, challenging what is known and revealing what is unknown

DAVID HAUSSLER
Scientific Director

TOP STORIES

AT THE FOREFRONT OF THE COVID RESPONSE: SEQUENCING EFFORTS CULMINATE IN PATHOGEN GENOMICS INITIATIVE

Our championing of testing and sequencing as critical tools to help end the pandemic has culminated in a new California public health initiative on pathogen genomics.

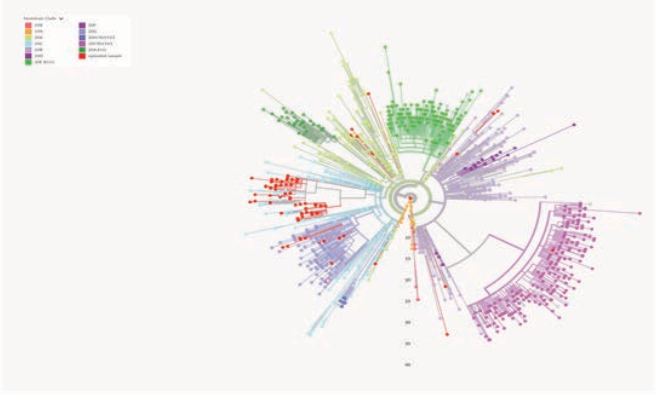


Throughout the pandemic, we have made our position on sequencing clear: The U.S. needs coordinated genomics sequencing programs to track COVID variants.

Addressing sequencing needs at a local level, we have been routinely sequencing positive virus samples that come from the Colligan Clinical Diagnostic Lab, a campus COVID-19 testing lab started in May 2020 to meet local testing needs.

In a February 2021 editorial in Bloomberg, Scientific Director David Haussler joined leading researchers pushing for a more aggressive and more concerted nationwide sequencing effort. Haussler went on record in Nature with other scientists calling for "fully open sharing of coronavirus genome data."

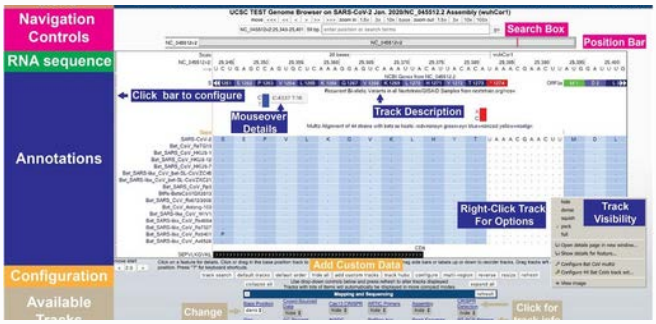
The Genome Browser's Angie Hinrichs's statement in Science decrying access and transparency issues with key coronavirus sequences databases was then followed up by a Nature editorial from Haussler, Max Haeussler, Hinrichs, Russ Corbett-Detig, and Isabel Bjork asking for freedom to share Covid-19 sequence data.



Our championing of testing and sequencing as critical tools to help end the pandemic has culminated in a new California public health initiative on pathogen genomics. Watch our website for developing news on this initiative.

20 YEARS ON, UCSC GENOME BROWSER IS CRITICAL TO PANDEMIC-ERA RESEARCH

On July 7, 2020, we celebrated a significant anniversary: It has been 20 years since UC Santa Cruz published the assembled draft human genome to the Internet. Reflecting back on the global impact of the Human Genome Project, we commemorated the significant moment when UC Santa Cruz ensured the genome would be free and available to everyone, everywhere, forever.



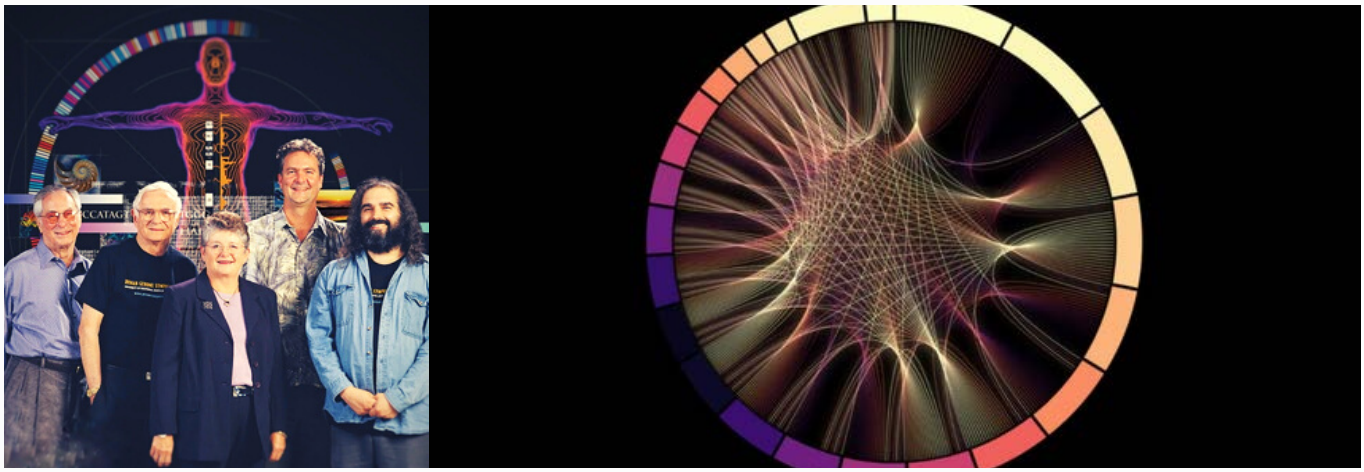
MILLION-COVID-GENOME TREE COULD BE A FIRST

The Genome Browser is now home to the first phylogenetic or evolutionary tree to ever connect more than one million genetic relatives. A UCSC team that includes (LtoR) Angie Hinrichs, Yatish Turakhia and Russell Corbett-Detig tackled this computational puzzle in order to create a dynamic evolutionary tree that would better enable real-time genomic contact tracing. As of fiscal 2020 year-end, the Virus Browser has grown to roughly three million samples.

HOW DNA REVEALED THE WOOLLY MAMMOTH'S FATE – AND WHAT IT TEACHES US TODAY

Speaking at a session at the World Economic Forum in Davos 2020, UCSC genomicist Beth Shapiro shared how genomics techniques that help us understand how ancient species became extinct can guide today's conservation efforts today. Extracting DNA from a frozen Alaskan island lake, Shapiro's team uncovered how a population of woolly mammoths died out. This pointed to habitat connectivity as a critical factor in helping endangered species survive, enabling them to escape dangers like drought or fire.

SCIENTIFIC COMMUNITY CELEBRATES THE HUMAN GENOME PROJECT ANNIVERSARY



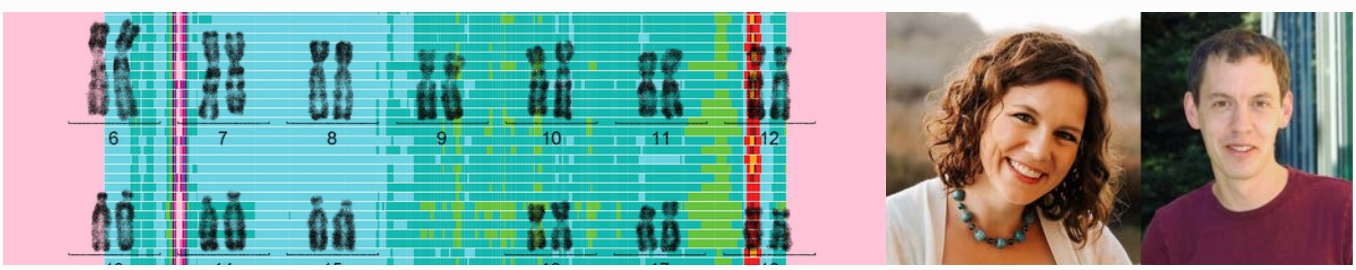
July 7, 2020 was the 20th anniversary of the magical day UC Santa Cruz researchers published the draft human genome assembly to the internet, playing a critical role in the Human Genome Project and ensuring the human genome would be free for everyone, forever. Pictured in 2000 at the time of the UC Santa Cruz Human Genome Symposium are, left to right, UCSC Foundation Trustee Arthur Graham; Chancellor Emeritus Robert Sinsheimer; Chancellor M.R.C. Greenwood; David Haussler and then graduate student Jim Kent

Nature's editors used the occasion of the 20th anniversary of its publication of the Human Genome Project Consortium's research as an opportunity to review two decades of genomics milestones. Among those milestones attributed to UCSC were the release of the Genome Browser; the first draft Neanderthal genome; and just this past year, the first gapless, telomere-to-telomere human chromosome assembly, and first complete human X chromosome assembly.

In a similar vein, Science Magazine highlighted Jim Kent's (pictured at work in his garage in 2000) heroic efforts to create "a key software tool used in the international effort to sequence the human genome." Kent's code pieced together the "millions of fragments of DNA sequence generated at labs around the globe, literally making the human genome."

Among those milestones attributed to UCSC were the release of the Genome Browser; the first draft Neanderthal genome; and just this past year, the first gapless, telomere-to-telomere human chromosome assembly.

COMPLETING THE SEQUENCE OF A HUMAN GENOME



While the Human Genome Project was undoubtedly a tour de force, the truth was that twenty years since UCSC posted the first draft assembly to the internet, about 8 percent of the more than 3-billion-base-pairs in the human genome remained unsequenced. Researchers have come to realize the significance of those gaps in the sequence, which genomicist Karen Miga calls the "final unknown." Now, Miga, Benedict Paten and fellow Genomics Institute researchers are once again leading efforts to complete the human genetic code. Their 2021 preprints detailing the complete sequence of a human genome mark a major milestone, helping ensure that all variants will be discovered and studied.

Twenty years after UCSC posted the first draft assembly to the internet, about 8 percent of the more than 3-billion-base-pairs in the human genome remained unsequenced.

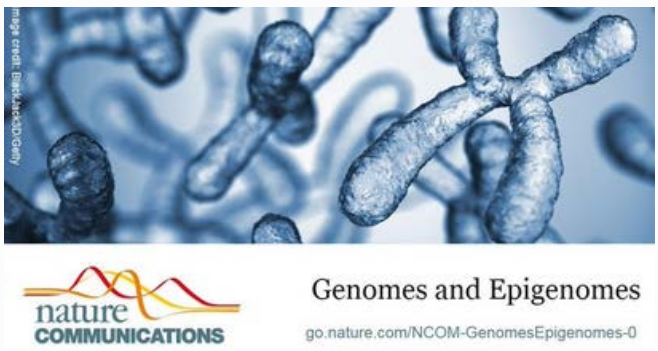
NEW TOOLS ENABLE LARGE-SCALE VERTEBRATE GENOMICS STUDY

In November 2020, our team published a paper in Nature describing their genome alignment tool, Progressive Cactus, which allows computationally efficient alignment of several hundred whole-genome sequences. Progressive Cactus empowers large-scale studies of vertebrate evolution, and coincidentally enabled significant studies of the evolution of birds and mammals published in the same Nature Issue.

In an April 2021 edition of Nature, the UCSC-affiliated Vertebrate Genomes Project (VGP) announced a flagship study for 16 diploid high-quality, near error-free, and near complete vertebrate reference genome assemblies that result from five years of piloting the first phase of the VGP project, an outgrowth of a UCSC meeting in 2009.

NATURE SPOTLIGHTS UCSC XENA DATA

UCSC played a key part in leading the main project of the global scientific consortium Pan-Cancer Analysis of Whole Genomes (PCAWG) to completion. PCAWG generated a vast quantity of whole-genome cancer sequencing resource data. UCSC Xena followed this up with a user guide series intended to help researchers leverage the data to understand the biology of cancers more deeply. Nature Communications editors called Xena's user guide for the online exploration and visualization of PCAWG data "exciting research in genomes and epigenomes."



SCIENTIST WINS HEART HEALTH RESEARCH FELLOWSHIP

Following her work creating the world's largest database of genetic variants in the breast cancer gene BRCA, UCSC genomics scientist Melissa Cline was awarded a fellowship to clarify genetic risk for sudden heart failure. Heritable cardiomyopathies affect 1 in 200-500 individuals in the U.S., many of whom are symptom-free yet at risk. Genetic testing is helping to identify at-risk individuals, but is still limited by our understanding of variants' potential clinical impact.

BRCA EXCHANGE

Genetic testing is helping to identify at-risk individuals, but is still limited by our understanding of variants' potential clinical impact.

NSF ADVANCES 25 PROJECTS TO EXPLORE BOLD IDEAS FOR TRANSFORMATIVE RESEARCH

Among the winners chosen from 800 entries for this significant NSF recognition was affiliate Richard (Ed) Green for "The evolution and diversity of the human brain." Green will use brain organoids to shed light on the role of recently evolved human-specific genetic variants. Green plans to discover variants using innovative computational analysis comparing the human genome with those of our closest, extinct ancestors – Neanderthals and Denisovans.

A TEAM RICH IN TOP-CITED RESEARCHERS

A number of our team made a list of Clarivate's 2020 Highly Cited Researchers. The list includes global research scientists who have demonstrated exceptional influence through their publications.

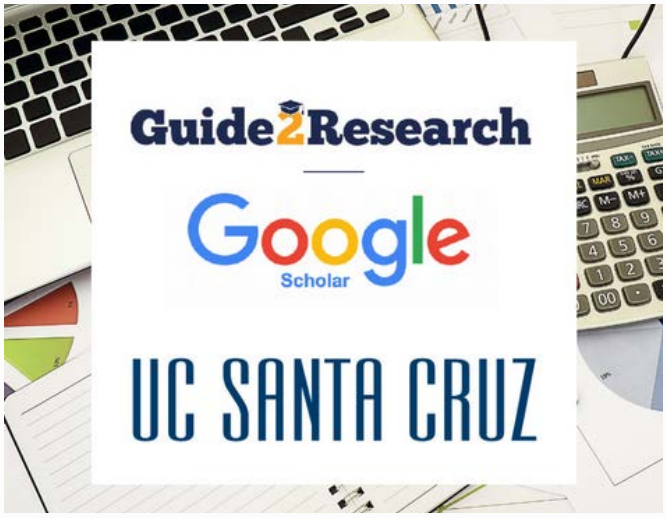
As in previous years, the list includes Mark Diekhans, James Kent, and nine more Genome Browser team members: Galt Barber, Hiram Clawson, Max Haeussler, Angie Hinrichs, Donna Karolchik, Robert Kuhn, Brian Raney, Kate Rosenbloom, & Ann Zweig.

The list includes...scientists who have demonstrated exceptional influence through their publications.

SCIENTIFIC DIRECTOR RANKS NEAR TOP WORLDWIDE IN COMPUTER SCIENCE RESEARCH

We are very proud that Scientific Director David Haussler ranks #6 in the world according to Guide2Research's 2021 Top Scientists in the field of Computer Science and Electronics. The ranking is based on the H-Index, which attempts to measure productivity and citation impact.

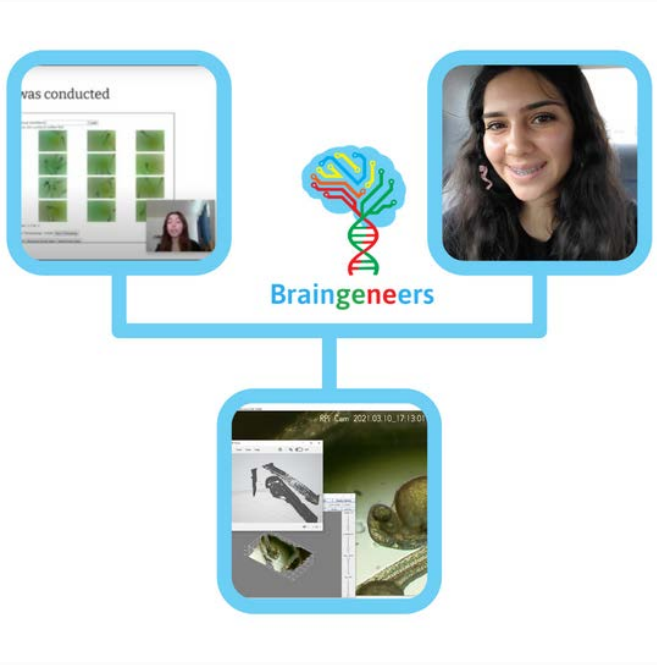
Scientific Director David Haussler ranks #6 in the world according to Guide2Research's 2021 Top Scientists.



INVENTION & COLLABORATION TAKES RESEARCH BEYOND BASICS

An affordable, DIY science platform invented by UCSC & UCSF grad students called the Picroscope inspired an ambitious collaboration with Salinas's Alisal High School AP Bio class. Taking basic research into students' homes during the pandemic, the collaboration set up a course built around a life-sciences experiment and tools for remote interaction and observation from students' homes. Building on their local success, the group launched the course in seven Latin American countries simultaneously, where students there used the Picroscope to understand the side effects of chemicals used in COVID-19 treatment.

Building on local success, the group launched the course in seven Latin American countries simultaneously.



FIVE-YEAR PLAN

LOOKING FORWARD

We are now in our second year of our 5-year plan, which identifies our primary areas of focus in genomics research for the fiscal years between July 2019 and June 2024. These are: (1) defeating diseases and pathogens; (2) furthering data sharing and open compute platforms; (3) advancing scientific discovery by use of novel sequencing and other biotechnologies; (4) protecting the web of life through comparative and conservation genomics; (5) furthering diversity and inclusion in genomics research and its applications, and (6) expanding research in society and genomics. In this section, we define our strategic objectives, we identify progress toward those and identify areas of future exploration and development.

1. COMBATING DISEASE AND PATHOGENS

This year, we were honored to use our computational genomics skills to further research in pathogens, and to help pave the way for an integrated virus data platform that is changing how virus disease information is accessed and used across the globe. As we prioritize informing public health and fellow research science and medical experts, we have maintained focus, too, on the important work we spearhead in using genomics to eradicate diseases.

We continue to apply big data comparative genomics to discover treatment options for children and adults with rare, undiagnosed or refractory diseases. We do this work in partnership with leading research hospitals. To expand our impact, we successfully advocated to build a diagnostic lab that is now testing for COVID-19, and will transition into a pediatric genomics diagnostic lab that will be part of a larger live cell biotechnology complex. Our focus now is on building a live cell and biotechnology research complex that will offer integrated basic science, sequencing and diagnostics. The vanguard application for this lab is centered on developing advanced ex vivo biotechnology using stem cell derived human brain tissues for neuroscience.

These technologies will rapidly spread to other biomedical areas. We are accelerating the development of new technology and new systems to grow brain tissue on the internet of things, interacting with deep learning, using the computational expertise on which we have built our reputation.

The dedicated, live-cell genomics sequencing center will offer a full-cycle, integrated wet lab facility where researchers can grow tissues with live cells, sequence them and diagnose disease. Digitization of the wet lab will achieve more than a move from analog to a networked, scalable research facility: Digitization will achieve exponential, reproducible improvements in speed of therapeutic discoveries and potential for improved care. It will make these improvements more accessible through lower costs and the disaggregation of research and clinical care.

2. FURTHERING GENOMIC DATA SHARING AND OPEN COMPUTE PLATFORMS

Our commitment to data sharing, symbolized by our publishing of the first draft of the human genome to the internet, remains the cornerstone of our identity. We are furthering data sharing with our UCSC Genome Browser and in particular the SARS-CoV-2 Genome Browser, which is designed to help researchers track and analyze coronavirus genome data, its variants and related vital data donated by labs around the world. The Browser includes the positions in the virus genome used by all the different "PCR" methods employed to detect the virus. It includes the positions in the virus genome where experimental research by labs around the world has indicated that mutations are likely to produce dangerous new strains and tracks the positions in the virus genome that antibodies produced by infected patients use to recognize the virus. We are the only public entity that is actively collecting and constantly integrating these types of information into a single, online resource for tracking new mutations and their molecular biology. Thousands of researchers worldwide now rely on the SARS-CoV-2 Genome Browser.

We will continue to develop our virus browser to freely and openly share this information. Going forward, we are improving our browser's impact by expanding its audience, by developing tools and practices that are accessible to persons who are not themselves experts in genomics. In this way, scientific discovery is advanced because access to sequencing technologies is democratized.

In parallel, the Computational Genomics Platform (CGP) team is hard at work helping to create a Data Biosphere. The Data Biosphere is an open, modular, community focused and standards based ecosystem of multiple cloud-based platforms. These platforms now include the National Human Genome Research Institute's AnVIL, the National Heart, Lung and Blood Institute's BioData Catalyst and the Human Cell Atlas Data Coordination Platform, each designed to hold rich genomic datasets and enable research for large communities of researchers. The CGP develops software, services and educational training materials for each effort. Connecting these platforms is Dockstore, a UCSC co-led effort developed by the CGP to increase the interoperability of genomics and biomedical computation.

3. ADVANCING DISCOVERY WITH NOVEL GENOMIC SEQUENCING & OTHER BIOTECHNOLOGIES

We are using novel technologies to sequence the SARS CoV-2 virus genome, to add to the data available on the UCSC Browser, and to assist in the understanding, treatment and cure of undiagnosed and rare diseases. New technology, such as nanopore sequencing -- developed here on the UC Santa Cruz campus -- offers long read sequencing which provides more and different genomic data than was previously available.

For 20 years, the world had a single individual's genome for a reference. We were at the forefront, as part of a worldwide consortium, of the effort to develop and publish the human genome. Yet while sequencing whole genomes has become affordable, a significant cost of using genomes in medicine is the cost of looking up all the genetic variants in a patient's genome in the reference genome - a complicated procedure that looks for approximate matches.

The procedure is approximate because the standard reference genome does not represent much of the genetic variation that exists in the human population. We are now advancing the human genome reference. A UCSC-led consortium to improve the reference genome is now in its second year, and is dedicated to using novel genomic sequencing technology to change how we think about genomics. Previously the human genome reference was unable to represent the genetic diversity in our species. Many different reference genomes, from individuals representing groups from all parts of the world, offers a fair and unbiased resource for global genomic medicine. We refer to this as our pangenome work because it diversifies the reference genome. Going forward, we are using novel sequencing technologies also to complete the human genome, which means sequencing and recording the information on parts of the genome that were previously unexplored. This will be a focus of work over the next 3 years of this 5-year plan, with our goal being to better understand the biological role of the complex, repeat-rich and largely unexplored parts of our genome: their genetic and epigenetic organization, how these sequences change over time, and how they contribute to human disease.

4. COMPARATIVE AND CONSERVATION GENOMICS TO PROTECT THE WEB OF LIFE

Comparative and conservation genomics is an area of particular growth for the Genomics Institute, and as part of this growth, we have refined and expanded our role. We use genomics to protect populations, communities, and ecosystems. Some applications of genomics for conservation include: establishing genomic resources for non-model organisms such that conservation practitioners can use these resources to identify populations that may be suffering from inbreeding depression, link particular mutations to maladaptive traits, and identify candidate individuals for translocation for genetic rescue; perform comparative genomic analyses to detect genomic signatures associated with endangerment risk; and explore the use of biotechnologies like cloning and gene editing in future conservation toolkits. We are leading collaborators in projects including the Genome 10K project, which is now part of the Vertebrate Genome Project (VGP) and whose goal is to sequence one genome from each of the 66,000 vertebrate animal species on our planet to better understand and conserve the species, and the California Conservation Genomics Project, which is a multi-million dollar investment by the State of California into generating genomic resources to protect California's biodiversity.

We are also leading development of approaches that use DNA preserved in environmental samples, including water and soil, to track changes in biodiversity both over time and across habitats. This year, the UCSC Paleogenomics Lab entered a three-year agreement with the National Park Service to develop environmental DNA — or eDNA — conservation programs for the agency. This multifaceted collaboration aims to use DNA scattered in the environment to discover what organisms call West Coast parks home and to slow the spread of invasive species. We work at the grassroots level with eDNA: We are key players in CALeDNA, a program where school-age children and the public join university students to collect samples in natural habitats across the state. These samples are processed for eDNA and the data and bioinformatic pipelines are made available to the public for any use. The goal of CALeDNA is to monitor biodiversity in California ecosystems using evidence from the entire web of life to better understand how communities respond to external impacts including agriculture, urbanization, and fire.

To support our biodiversity genomics work, we are planning a conservation genomics program that capitalizes on technological advances in genome sequencing and assembly to develop and distribute resources necessary for managers to make on-the-ground decisions about species conservation and habitat protection. These resources would include portable platforms to survey biodiversity, and virtual platforms designed to enable biodiversity-based assessments or hypothesis testing with predictive models that integrate cutting-edge, remote-sensing environmental data with historical data. Our program will prepare a genomics-enabled economy characterized by biodiversity, and will advise where to put the next protected marine and terrestrial parks.

5. FURTHERING DIVERSITY AND INCLUSION IN GENOMICS

As we move genomic science forward, we never forget the humanity of the individuals who are the cornerstone of our genomics experience: We are committed to racial justice. We are committed to ethics and integrity. We are committed to community action.

We remain on track to grow to projected 25-40% increase in staff and revenue during this 5-year period. We are focusing on growing well —which we define as engaging our teams and enhancing diversity in all forms. New ideas are best wrought by different perspectives; we challenge ourselves daily to change in uncomfortable ways in order to grow and expand. In this spirit, and in the spirit of the social justice movement of 2020, the Genomics Institute established a Diversity Committee last summer and we have already pushed forward proposed policies. For example, we are encouraging goals for diversity, equity and inclusion (DEI) actions in staff annual performance assessments, and developing more equitable and inclusive hiring practices at all levels. Over the next five years, we will hold ourselves accountable to these goals, using metrics to assess and improve our performance.

6. EXPANDING RESEARCH IN GENOMICS AND SOCIETY

It is our goal to explore the societal implications of genomics, with the aim of bringing forward a new vision of genomics that would integrally incorporate the impact of society and justice.
In the past decade genomics have advanced at an exceptional pace, offering the promise to revolutionize medicine and our general understanding of genomic diversity, while also stimulating vigorous discussion of the future of longstanding democratic rights — such as privacy, informed consent, and property. The exploration of the socio-economic, ethical, cultural, legal and community dimensions of genomic research is critical for transforming genomic data into meaningful knowledge that supports the public good. We are committed to implementing a vision of genomics that integrally embraces the impact on society and justice.

This year, faculty and staff from Science and Justice Research Center (SJRC) and from the Genomics Institute have conducted a series of discussions that re-envision a Genomics and Society program. We are planning conversation spaces that facilitate interdisciplinary discussions about fair and just practice in genomics and the societal implications of the work done at the Institute and beyond, with the long-term goal of establishing a fully integrated Genomics and Society Research sub-track at the institute. The initial steps of this initiative will be the creation of a seminar series as well as educational programming for undergrad and graduate students. We further aim to create resources to support and guide future research at the Institute by conceptualizing and developing a Responsible Conduct of Research (RCR) training program. The long term success of this initiative is dependent on the creation of permanent resources including additional employees or staff. This means that in the coming years, major efforts will be to engage in discussion, strategizing and to explore funding opportunities, while we continue to develop a concrete vision of a Genomics and Society program.

IN SUMMARY

We envision an experiment of the future using advanced genomic technology, perhaps one of hundreds of thousands of similar experiments conducted in parallel all over the world to fight a viral outbreak in real time. We imagine a day when new drugs are developed as efficiently as new software. We imagine a world where undiagnosed diseases no longer exist, treatments are targeted, and we understand what makes us human. We are committed to creating the foundation for this future in a responsible and ethical manner, with explicit attention to community engagement, inclusion, and fair representation. This is the world that we are set on achieving.

DIVERSITY, EQUITY & INCLUSION

A significant effort was made this year to enhance our efforts to be more diverse, equitable, and inclusive in our genomics-related research, scholarship, and personnel. The advancements in DEI made this year were a combination of initiatives led by staff, academic researchers and faculty affiliates of the Genomics Institute. Here, we provide an overview of this important work, and focus on impact and next steps.

Faculty Director Angela Brooks organized a Genomics Institute Diversity Committee in June 2020, which meets biweekly to plan and prioritize organization-wide initiatives designed to promote diversity, equity and inclusion at the Genomics Institute. Purposely built to include representatives from all categories of organization members from undergraduates to faculty, the committee created an opportunity to increase connectivity between the groups and incorporate a wider variety of perspectives.

Selected Committee-initiated accomplishments included:

- Scheduling two four-week, free UCSC Bioinformatics & Coding Short Courses to meet the needs of students without prior programming experience, open to non-UCSC-student participants — both to be held during the Summer 2021 period.
- Training managers and achieving 61% staff participation in DEI action goal setting for annual performance reviews
- Highlighting campus policy and resources in a Fair Hiring summary and similarly spotlighting resources to address conflict, harassment, bias and discrimination.
- Initiating an Institute-wide conversation regarding feelings of inclusion

Research scientist Mary Goldman initiated and leads an organization-wide racial justice learning and action group to explore racism in the profession of science, the profession of engineering, and in scientific research itself. An example of a collective action the group has taken is a Wiki-a-thon, where they contributed new Wikipedia entries for underrepresented scientists. This activity attracted attendees not only from the Institute but also from the campus community. This group has met biweekly since July 2020.

The Wiki-a-thon attracted attendees from the Institute and also from the campus community. Goldman's group has met biweekly since July 2020.

Research scientist Holly Beale meanwhile spearheaded The Treehouse Undergraduate Bioinformatics immersion (TUBI) program, which provides an opportunity for undergraduate students to gain real world work experience as researchers in a supportive environment, contributing to Treehouse bioinformatic projects. Beale started TUBI in Winter Quarter 2020 and has served 80 students to date.

Mary Goldman (left)
Holly Beale (right)

Beale started TUBI in Winter 2020 and has served 80 students to date.

The Genomics Institute Office of Diversity (GIOD) is focused on driving forward DEI work that directly impacts students and research. Through the Office of Diversity, we continue our long-standing undergraduate research training program, the Research Mentoring Research Internship Program (RMI). Established in 2003, RMI has continued to operate for the intervening 18 years with funding from the National Human Genome Research Institute. The RMI supports and empowers first generation and minority-identified students in the field of genomics with scholarships, hands-on research training, and career development resources.

Our 2021 RMI graduates went on to careers in biotech and joined prestigious graduate programs at UC San Francisco and Harvard. In recognition of UCSC's status as an Hispanic Serving Institution, the GIOD continued to look for opportunities to be more inclusive of our Latinx community.

In April of 2021 we hosted our first-ever bilingual Spanish-English DNA Day, which featured a website complete with teaching resources. The virtual event featured activities such as strawberry DNA extraction and research presentations in English and Spanish.

Our RMI undergraduate research training program was established in 2003 and has continued to operate for the intervening 18 years with funding from the National Human Genome Research Institute.

Extramurally, the Office of Diversity actively reaches out to academic organizations such as the San Francisco Unified School District Biotech Academy, the Office of Undergraduate Research at California State University Monterey Bay, and the California Forum for Diversity in Graduate Education. We connect with student populations nationally through the Society for Advancement of Chicanos and Native Americans in Science (SACNAS), and the Annual Biomedical Research Conference for Minority Students (ABRCMS). It was through this outreach that the Office became the facilitator for our Picroscope collaboration with Salinas's Alisal High School AP Bio class See Top Stories, p. 9, for more.

INTERDISCIPLINARY PROGRAM SUPPORT

INTERDISCIPLINARY GENOMICS PROGRAMS

The Genomics Institute is committed to supporting interdisciplinary projects in genomics that align with its mission, vision and strategic plan. We feel that by doing so, we can positively impact academic and community engagement and inclusion in genomics research. Such projects are often presented for consideration by research affiliates with proposals explaining the use of funds, expected outcome in advancement of genomics, and year-end reporting.

This year, we made commitments to help establish the Center for Public Philosophy (CPP) and the Earth Futures Institute (EFI); we continued our support for genomics-related activities at the Science & Justice Research Center (SJRC); we renewed our commitment to the Arts Division's Jennifer Parker/OpenLab to extend the Arts + Genomics Initiative; we established the new Double Helix (GIDH) Graduate Fellowship, which awards outstanding applicants to the interdisciplinary Graduate Program in Biomedical Sciences and Engineering; we lent financial support to conservation genomics through a grant extending technical assistance; finally, we increased our financial support for cross-divisional activities led by our Diversity Committee, which includes representatives from Physical Sciences and Engineering.

SUPPORT FOR SCIENCE AND JUSTICE

This year, Science & Justice produced The Pandemicene, a podcast series addressing overlapping topics of institutions of care, genomics, health and society. Their Just Biomedicine research cluster's chapter in Counterpoints: A San Francisco Bay Area Atlas of Displacement and Resistance (PM Press, 2021) examined visions for health and healthcare that have been imagined and practiced in San Francisco's biotech hub. Support also allowed a graduate student with an Institute-affiliated faculty mentor to complete the Science & Justice Training Program, as well as the continuation of the Theorizing Race After Race research cluster, providing two more Institute grad students with summer fellowships.

SUPPORT FOR ART & GENOMICS

Art Professors Jennifer Parker and Karolina Karlic completed a limited edition box set of 100 fine art photographic prints for its Intersecting Data Fields: An Art & Genomics Collaboration, an exhibition about the diversity of the people, work, and impact of the Institute. These are now a part of the Special Collections archives at the McHenry Library. Professor Parker led a collaboration with the UCSC Sesnon Gallery on a virtual exhibition, highlighting UCSC genomics research. The exhibition opening featured an international panel of artists, and thanks to former UCSC graduate student Dr. José Carlos Espinel, a Spanish language version, ¿Qué nos hace humanos?, was made available through the Complutense University of Madrid, at ucm.es/arte-co/what-makes-us-human.

FINANCES

RESEARCH DEVELOPMENT REPORT

In fiscal year 2021 (FY21), the Genomics Institute submitted 48 proposals for 13 principal investigators (PIs) requesting $120.4M in new funds ($92.5M in direct costs, $27.9M indirect) for projects ranging from one to five years in duration. The average request was $829K per year in new funding. Of those proposals, 17 have been awarded to date (26 are still pending), resulting in approximately $31.8M in new funding ($23M direct, $8.8M indirect) most of which starts in FY22.

In addition, the GI brought in $28.5M in research funds for FY21, including $21.7M direct, $6.7M indirect, and $163K in gift fees for 34 awards, 11 of which were first year projects while the remainder were continuations. Of note, >$3.3M was awarded for projects focused on SARS-CoV-2, and $444K for diversity-aimed projects, in addition to the gift and operating funds for these purposes mentioned elsewhere.

Federal funding comprised 66% of FY21 research funds; followed by 13%, 11%, and 10% from foundations, state, and research gift sources, respectively, while research funds from UC/UCSC totaled <1%. A summary follows in the table and chart below.

>$3.3M was awarded for projects focused on SARS-CoV-2...and $444K for diversity-aimed projects...

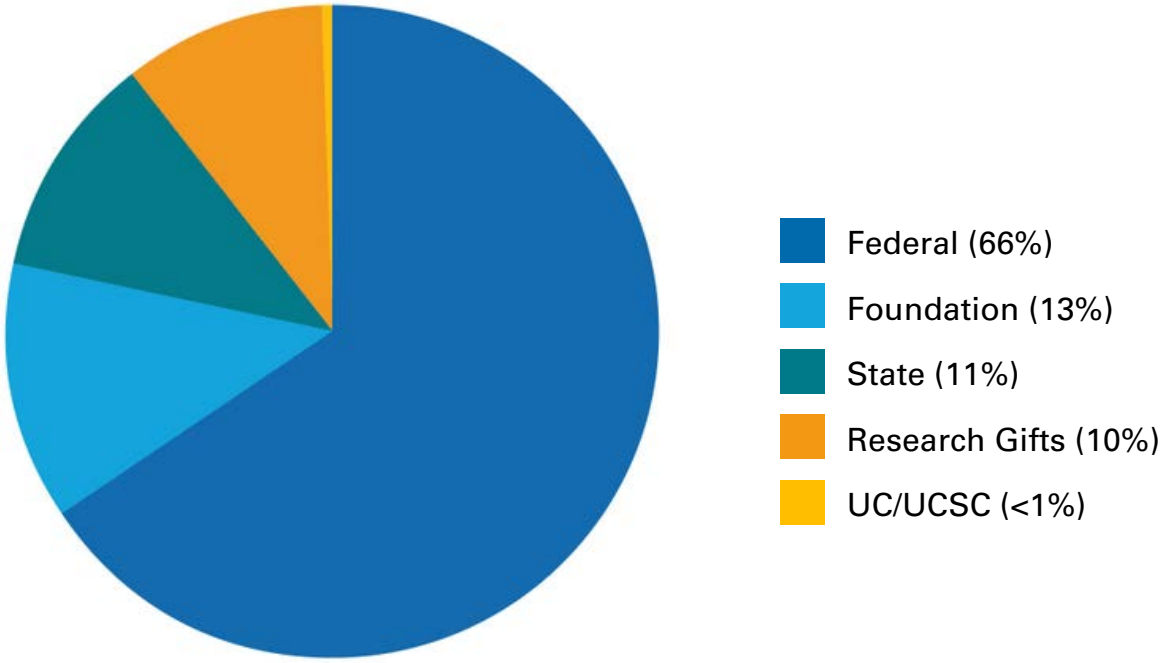
GENOMICS INSTITUTE AWARDS

Funding Category	Direct Costs	Indirect Costs (or Gift Fee)	Total Direct + Indirect	% of Total Award
Federal	$13,441,189	$5,453,308	$18,894,497	66
Foundation	3,303,299	362,359	3,665,558	13
State	2,376,671	713,001	3,089,672	11
Research Gifts*	2,556,800	163,200	2,720,000	10
UC/UCSC	116,943	33,054	149,997	<1
Grand Total	**$21,794,832**	**$6,724,922**	**$28,519,724**	

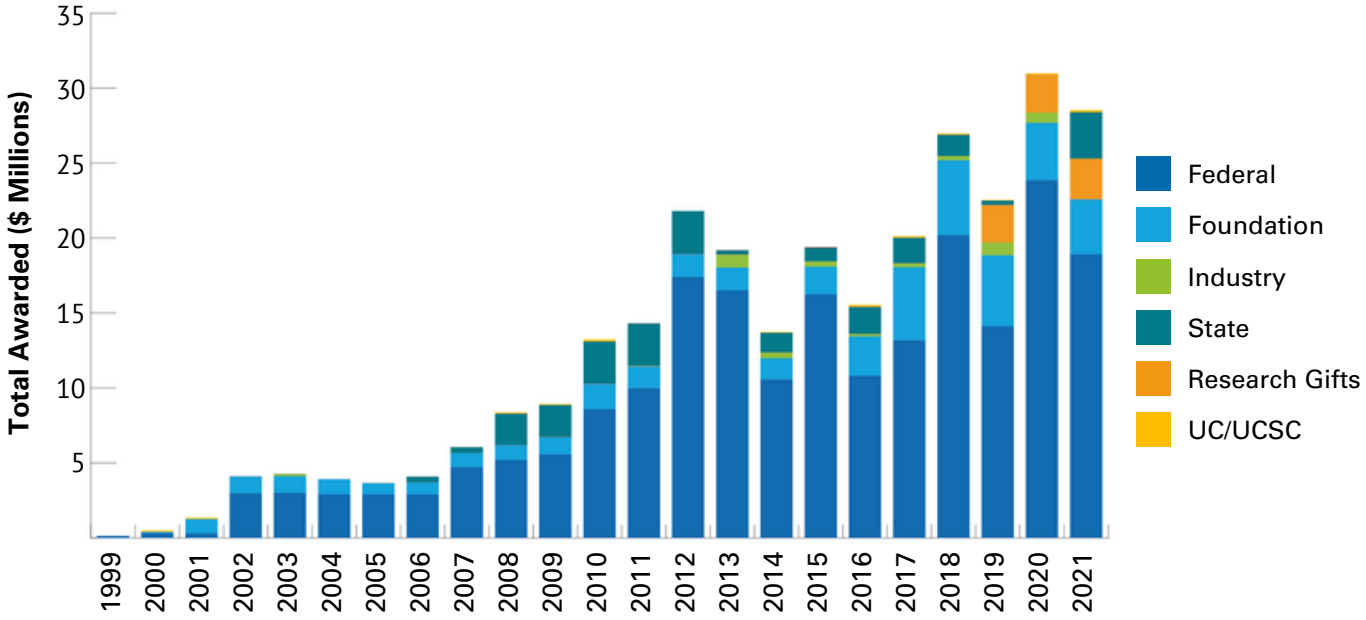
** Awarded by Schmidt Futures, a venture facility to turn private philanthropic capital into public benefit, this research gift was requested through a proposal process and has project reporting requirements similar to a grant.*

CONTRACT & GRANT FUNDING SOURCES

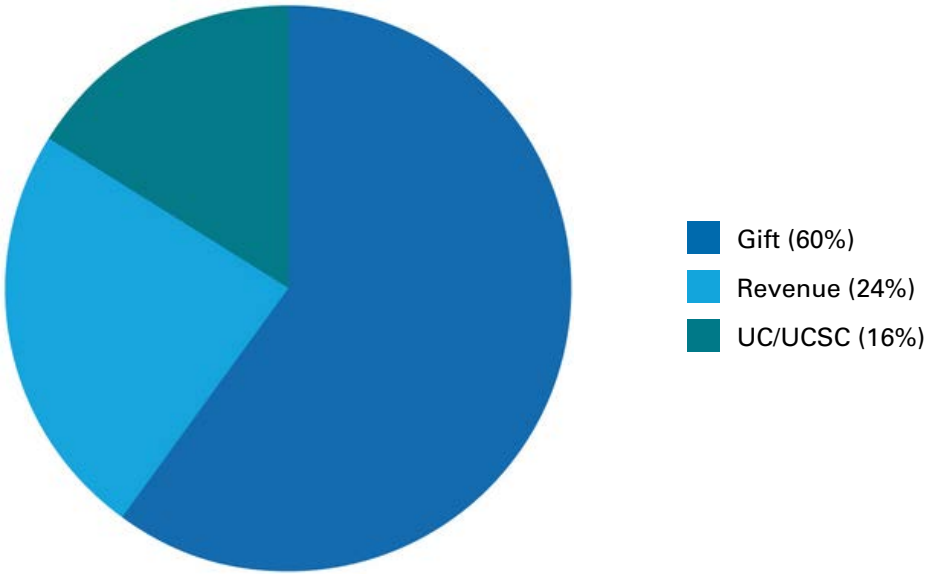
The chart below displays annual research awards received beginning in 1999 from our foundational organization, the CBSE. Despite the interruptions and distractions of the past year, FY21 annual research funding remained within 7% of FY20 funding. We believe that the relative stability afforded by gifts from our generous donors in addition to significant campus support allowed us to maintain the infrastructure to achieve this level of research funding.



ANNUAL RESEARCH AWARDS 1999-2021



OPERATING BUDGET OVERVIEW



Our operating budget supports the administrative, operational and physical infrastructure of the Institute. Total available operating funds in FY 2021 were $5.44M and were composed of $3.27M in gift funds, $1.31M in revenue funding, and $0.86M in UC/UCSC funding.

Gift funds include $850K from a generous private donor for strategic support of Genomic Institute operations (year 2 of a 10-year commitment) and $1.46M from various donors to support our efforts on the SARS-CoV-2 Genome Browser, Covid-19 testing, and sequencing. Numerous smaller funds from various private donors and tech industries provide general operating support for core PIs as well as for named purposes e.g. the Treehouse, Xena, and BRCA Exchange projects and the Genomics Institute Office of Diversity.

Revenue is predominantly composed of Genomics Project Support recharge income which totalled $1,307,682. An additional $18,418 was directed to the institute for its share of Genome Browser Royalty Income for FY22. Other portions of royalty income were distributed to software authors, HHMI and the Chancellor's Office per copyright distribution agreements. The Genomics Institute had no Genome Browser Training Income in FY21.

Total available operating funds in FY 2021 were $5.44M and were composed of $3.27M in gift funds, $1.31M in revenue funding, and $0.86M in UC/UCSC funding.

OPERATING BUDGET REPORT

Category	Carry Forward	Budget Adjustments*	Total Budget	% of Total Budget	Transactions	Unspent Balance
Gift	$1,101,753	$2,168,557	$3,270,310	60%	$1,459,446	**$1,810,864**
Revenue	(14,606)	1,326,100	1,311,494	24%	1,145,004	**166,490**
UC/UCSC	100,260	762,560	862,820	16%	874,340	**(11,520)**
Totals	**$1,187,407**	**$4,257,217**	**5,444,625**	**100%**	**3,478,790**	**1,985,835**

**Includes recharge income and budget transfers for cross-divisional support.*

EXPENSES OVERVIEW

In conjunction with the annual $850K gift mentioned above, we receive $650K annually from UCSC for operating expenses. Other smaller pools of state funding are earmarked for our Office of Diversity and sequencing lab design.

Salary and benefits account for 69% of our operating expenses overall at $2.4M for approximately 21 FTE of staff and student workers in support of our grant, finance, HR, communications, facilities, events and front office operations. Non-capital expenditures account for the next highest spending category at 26% ($0.9M); included are all supplies, non-capital equipment, services, recharges, temp agency, and other expenses for office operations

We are grateful to all of our funders, both inside and outside the university, for allowing us to provide critical services and administrative support to our research teams.

Nearly $120K was distributed to interdisciplinary partners and projects including the Science & Justice Research Center, Conservation Genomics, Arts, the Center for Public Philosophy, and the newly established Genomics Institute Double Helix Fellowship.

as well as research activities funded by gift accounts. Of note, we distributed more than $117K for FY21 to interdisciplinary partners and projects including the Science & Justice Research Center, Conservation Genomics, Arts, the Center for Public Philosophy, and the newly established Genomics Institute Double Helix Fellowship. Sequencing equipment purchases accounted for 5% of our spending while other minor expenses make up less than 1%.

As always, we are grateful to all of our funders, both inside and outside the university, for allowing us to provide critical services and administrative support to our research teams.

RESEARCH OPERATIONS EXPENSES

Expense Category	Gift	Revenue	UC/UCSC	Total Transactions	% of Expenses
Salaries & Benefits	$646,910	$1,168,827	$583,687	$2,399,424	68.97%
Travel	-	196	-	196	0.1
Non-Capital Expenditures	625,300	(9,217)	290,653	906,736	26.06
Capital Expenditures	176,357	-	-	176,357	5.07
Tuition & Fees	10,879	-	-	10,879	0.31
Revenue	-	(19,228)	-	(19,228)	-0.55
Indirect Costs	-	4,426	-	4,426	0.13
Totals	**$1,459,446**	**$1,145,004**	**$874,340**	**$3,478,790**	**100%**

STRUCTURE & PEOPLE

STRUCTURE

The UC Santa Cruz Genomics Institute is a research organization dedicated to effecting positive change through genomics research. The Institute provides the organizing framework, focus, and leadership for the next great leap in the science of genomics and its implications for human health and species conservation. Our Leadership — our Associate Directors, Directors and their teams — have expertise in many but not all areas pertaining to our vision; it is through our Affiliates that we expand our impact and reach in genomics with a wide range of expertise and perspectives. Our Advisory Committee supports our adherence to our vision.

Our Affiliates are a cross-divisional group of more than 30 UC Santa Cruz faculty and research scientists (non- faculty researchers) who share the mission of the Genomics Institute and express a desire to have a relationship with the Genomics Institute on areas of their work that involves genomics research. Affiliates may have a small or large part of their work dedicated to genomics; we aim for a consistent intersection between affiliates' research aims and interests and Genomics Institute efforts.

Our Affiliates are a cross-divisional group of more than 30 UC Santa Cruz faculty and research scientists.

LEADERSHIP

David Haussler, Scientific Director; Benedict Paten, Beth Shapiro, Josh Stuart, Lars Fehren-Schmitz, Associate Directors; Angela Brooks, Faculty Director, Diversity; Ann Pace, Director, Research Development & Finance; Isabel Bjork, Executive Director

ADVISORY COMMITTEE

UCSC MEMBERS: Holger Schmidt, Advisory Committee Chair (Engineering/ECE); Mark Akeson (BME); Mircea Teodorescu (ECE); Olena Vaske (MCDB); Jenny Reardon (SOC); Jennifer Parker (ART); Carol Greider (MCDB); John MacMillan (OR/CHEM).

EXTERNAL MEMBERS: Richard Durbin, Sanger Institute; Erich Jarvis, Rockefeller University, Vertebrate Genome Project; Sheri Spunt, Sanford University; Peter Goodhand, GA4GH

AFFILIATES, ASSOCIATE DIRECTORS UPDATES

(Clockwise from top left) our newest affiliates & associate directors: Nobel Laureate Carol Greider, Prof. of MCD Bio.; Upasna Sharma, Assist. Prof. of MCD Bio.; Ali Shariati, Assist. Prof., BME; Karen Miga, Assist. Prof., BME; Lars Fehren-Schmitz, Assoc., Prof. Anthropology and Vanessa Jonsson, Assist. Prof., Applied Math.

Karen Miga was among those the journal Nature nominated to watch for groundbreaking stories in 2021.

In 2021, the Institute extended a warm welcome to (clockwise from top left) our newest affiliates, Nobel Laureate Carol Greider, Prof. of MCD Bio.; Upasna Sharma, Assist. Prof. of MCD Bio.; Ali Shariati, Assist. Prof., BME; and Vanessa Jonsson, appointed Assist. Prof., Applied Math for FY21. Additionally, we were very happy to welcome new faculty Assoc. Dir. Lars Fehren-Schmitz, Assoc. Prof., Anthropology; and Assoc. Dir. Karen Miga, appointed Engineering faculty for FY21. Miga was among those the journal Nature nominated to watch for groundbreaking stories in 2021. We also want to recognize Josh Stuart, who stepped away after a decade of service in a leadership role as the Institute's Associate Director: Stuart will continue as a Faculty Affiliate, and we look forward to continued collaboration with the Systems Biology Group.

We recognize Josh Stuart, who stepped away after a decade of service in a leadership role as the Institute's Associate Director.

PROJECTS & TEAMS

Our organization is home to more than a dozen key initiatives including Braingeneers, BRCA Exchange, CIRM at UCSC, Computational Genomics Laboratory (CGL) & Computational Genomics Platform (CGP), Genome 10K, Global Alliance for Genomics & Health, QB3 at UCSC, The Telomere-to-Telomere (T2T) Consortium, Treehouse Childhood Cancer Initiative, UCSC Genome Browser, Pathogen Genomics, and UCSC Xena. Visit genomics.ucsc.edu and navigate to the Projects menu to learn more.

Visit genomics.ucsc.edu
and navigate to the Projects menu to learn more.

OUR FACILITIES

Four computational research labs have designated space on the 2nd floor (11,031 sq. ft.) of UCSC's Westside Research Park at 2300 Delaware and three of these labs, led by faculty researchers, have additional space at the UCSC campus in Baskin School of Engineering's Physical Sciences Building (PSB).

Both locations have shared offices (67 to 377 sq. ft.) and open floor plans (481 to 1984 sq. ft.), which are often divided into cubicles. Together, these spaces provide seating for 122 research and technical staff, postdocs, graduate students, telecommuters, and visiting researchers.

CONTACT US

UC Santa Cruz Genomics Institute
Revealing life's code.

1156 High Street, Mail Stop: Genomics Institute
Santa Cruz, CA 95064
genomics.ucsc.edu | +1 (831) 459-1477

UC SANTA CRUZ | Baskin Engineering