Agilent
Trusted Answers

Research Citation Collection

Agilent Custom SurePrint Oligonucleotide Libraries

High-Performance Oligonucleotide Libraries at Your Fingertips!

The Agilent SurePrint Oligonucleotide Library Synthesis platform delivers advanced and reliable array-based DNA synthesis. We have harnessed our SurePrint platform to refine the DNA printing process, creating the ability to generate longer oligo libraries for your important research. Consistent synthesis of high-complexity oligo pools ensures libraries with superior fidelity and full representation, which improves functional results and reduces screening time. The fully custom, ultrahigh-quality oligonucleotide libraries are compatible with any application or experimental approach.

In this compendium, we have selected ten peer-reviewed publications to illustrate the utility and versatility of the SurePrint Oligonucleotide Libraries in a variety of applications. Learn how researchers were able to easily incorporate oligo libraries into their workflows and successfully move their research forward.

Table of Contents

Infectious disease research

De Novo Design of Picomolar SARS-CoV-2 Miniprotein Inhibitors 4

Unbiased Screens Show CD8⁺ T Cells of COVID-19 Patients Recognize Shared Epitopes in SARS-CoV-2 that Largely Reside Outside the Spike Protein 5

Genome-Scale Identification of SARS-CoV-2 and Pan-Coronavirus Host Factor Networks 6

Functional Interrogation of a SARS-CoV-2 Host Protein Interactome Identifies Unique and Shared Coronavirus Host Factors 7

Massively Parallel reporter assays

Dissection of c-AMP Response Element Architecture by Using Genomic and Episomal Massively Parallel Reporter Assays 8

Candidate Silencer Elements for the Human and Mouse Genomes 9

A Deep Neural Network for Predicting and Engineering Alternative Polyadenylation 10

Enhancers research

Mutational Bias and the Protein Code Shape the Evolution of Splicing Enhancers 11

CRISPR research

Functional Genomics of the Rapidly Replicating Bacterium *Vibrio Natriegens* by CRISPRi 12

Gene assembly (DropSynth)

DropSynth 2.0: High-Fidelity Multiplexed Gene Synthesis in Emulsions 13

De Novo Design of Picomolar SARS-CoV-2 Miniprotein Inhibitors

Science, **2020**.
doi: http://doi.org/10.1126/science.abd9909

Authors

Longxing Cao, Inna Goreshnik, Brian Coventry, James Brett Case, Lauren Miller, Lisa Kozodoy, Rita E. Chen, Lauren Carter, Alexandra C. Walls, Young-Jun Park, Eva-Maria Strauch, Lance Stewart, Michael S. Diamond, David Veesler, and David Baker

Abstract

Targeting the interaction between the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike protein and the human angiotensin-converting enzyme 2 (ACE2) receptor is a promising therapeutic strategy. We designed inhibitors using two *de novo* design approaches. Computer-generated scaffolds were either built around an ACE2 helix that interacts with the spike receptor binding domain (RBD) or docked against the RBD to identify new binding modes, and their amino acid sequences were designed to optimize target binding, folding, and stability. Ten designs bound the RBD, with affinities ranging from 100 picoMolar to 10 nanomolar, and blocked SARS-CoV-2 infection of Vero E6 cells with median inhibitory concentration (IC_{50}) values between 24 picoMolar and 35 nanomolar. The most potent, with new binding modes, are 56- and 64-residue proteins (IC_{50} approximately 0.16 nanograms per milliliter). Cryo-electron microscopy structures of these minibinders in complex with the SARS-CoV-2 spike ectodomain trimer with all three RBDs bound are nearly identical to the computational models. These hyperstable minibinders provide starting points for SARS-CoV-2 therapeutics.

Longxing Cao, Inna Goreshnik, Brian Coventry, James Brett Case, Lauren Miller, Lisa Kozodoy, Rita E. Chen, Lauren Carter, Alexandra C. Walls, Young-Jun Park, Eva-Maria Strauch, Lance Stewart, Michael S. Diamond, David Veesler, David Baker. De novo design of picoMolar SARS-CoV-2 miniprotein inhibitors. *Science*, **2020**, Oct 23;*370*(6515):426-431. doi: http://doi.org/10.1126/science.abd9909

Unbiased Screens Show CD8+ T Cells of COVID-19 Patients Recognize Shared Epitopes in SARS-CoV-2 that Largely Reside Outside the Spike Protein

Immunity, **2020**
doi: https://doi.org/10.1016/j.immuni.2020.10.006

Authors

Andrew P. Ferretti, Tomasz Kula, Yifan Wang, Dalena M.V. Nguyen, Adam Weinheimer, Garrett S. Dunlap, Qikai Xu, Nancy Nabilsi, Candace R. Perullo, Alexander W. Cristofaro, Holly J. Whitton, Amy Virbasius, Kenneth J. Olivier, Jr., Lyndsey R. Buckner, Angela T. Alistar, Eric D. Whitman, Sarah A. Bertino, Shrikanta Chattopadhyay, and Gavin MacBeath

Abstract

Developing effective strategies to prevent or treat coronavirus disease 2019 (COVID-19) requires understanding the natural immune response to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). We used an unbiased, genome-wide screening technology to determine the precise peptide sequences in SARS-CoV-2 that are recognized by the memory CD8+ T cells of COVID-19 patients. In total, we identified 3–8 epitopes for each of the six most prevalent human leukocyte antigen (HLA) types. These epitopes were broadly shared across patients and located in regions of the virus that are not subject to mutational variation. Notably, only 3 of the 29 shared epitopes were located in the spike protein, whereas most epitopes were located in ORF1ab or the nucleocapsid protein. We also found that CD8+ T cells generally do not cross-react with epitopes in the four seasonal coronaviruses that cause the common cold. Overall, these findings can inform development of next-generation vaccines that better recapitulate natural CD8+ T cell immunity to SARS-CoV-2.

Andrew P. Ferretti, Tomasz Kula, Yifan Wang, Dalena M.V. Nguyen, Adam Weinheimer, Garrett S. Dunlap, Qikai Xu, Nancy Nabilsi, Candace R. Perullo, Alexander W. Cristofaro, Holly J. Whitton, Amy Virbasius, Kenneth J. Olivier, Jr., Lyndsey R. Buckner, Angela T. Alistar, Eric D. Whitman, Sarah A. Bertino, Shrikanta Chattopadhyay, Gavin MacBeath. Unbiased Screens Show CD8+ T Cells of COVID-19 Patients Recognize Shared Epitopes in SARS-CoV-2 that Largely Reside outside the Spike Protein. *Immunity* (**2020**), *53*, 1095–1107. doi: https://doi.org/10.1016/j.immuni.2020.10.006

Genome-Scale Identification of SARS-CoV-2 and Pan-Coronavirus Host Factor Networks

Cell, **2021**

doi: https://doi.org/10.1016/j.cell.2020.12.006

Authors

William M. Schneider, Joseph M. Luna, H.-Heinrich Hoffmann, Francisco J. Sánchez-Rivera, Andrew A. Leal, Alison W. Ashbrook, Jérémie Le Pen, Inna Ricardo-Lax, Eleftherios Michailidis, Avery Peace, Ansgar F. Stenzel, Scott W. Lowe, Margaret R. MacDonald, Charles M. Rice, and John T. Poirier

Abstract

The coronavirus disease 2019 (COVID-19) pandemic has claimed the lives of over one million people worldwide. The causative agent, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is a member of the *Coronaviridae family* of viruses that can cause respiratory infections of varying severity. The cellular host factors and pathways co-opted during SARS-CoV-2 and related coronavirus life cycles remain ill defined. To address this gap, we performed genome-scale CRISPR knockout screens during infection by SARS-CoV-2 and three seasonal coronaviruses (HCoV-OC43, HCoV-NL63, and HCoV-229E). These screens uncovered host factors and pathways with pan-coronavirus and virus-specific functional roles, including major dependency on glycosaminoglycan biosynthesis, sterol regulatory element-binding protein (SREBP) signaling, bone morphogenetic protein (BMP) signaling, and glycosylphosphatidylinositol biosynthesis, as well as a requirement for several poorly characterized proteins. We identified an absolute requirement for the VMP1, TMEM41, and TMEM64 (VTT) domain-containing protein transmembrane protein 41B (TMEM41B) for infection by SARS-CoV-2 and three seasonal coronaviruses. This human coronavirus host factor compendium represents a rich resource to develop new therapeutic strategies for acute COVID-19 and potential future coronavirus pandemics.

William M. Schneider, Joseph M. Luna, H.-Heinrich Hoffmann, Francisco J. Sánchez-Rivera, Andrew A. Leal, Alison W. Ashbrook, Jérémie Le Pen, Inna Ricardo-Lax, Eleftherios Michailidis, Avery Peace, Ansgar F. Stenzel, Scott W. Lowe, Margaret R. MacDonald, Charles M. Rice, John T. Poirier. Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. *Cell* **2021**, *184*, 120-132.
doi: https://doi.org/10.1016/j.cell.2020.12.006

Functional Interrogation of a SARS-CoV-2 Host Protein Interactome Identifies Unique and Shared Coronavirus Host Factors

Cell Host & Microbe, **2021**
doi: https://doi.org/10.1016/j.chom.2020.12.009

Authors

H. Heinrich Hoffmann, Francisco J. Sánchez-Rivera, William M. Schneider, Joseph M. Luna, Yadira M. Soto-Feliciano, Alison W. Ashbrook, Jérémie Le Pen, Andrew A. Leal, Inna Ricardo-Lax, Eleftherios Michailidis, Yuan Hao, Ansgar F. Stenzel, Avery Peace, Johannes Zuber, C. David Allis, Scott W. Lowe, Margaret R. MacDonald, John T. Poirier, and Charles M. Rice

Abstract

The ongoing severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic has devastated the global economy and claimed more than 1.7 million lives, presenting an urgent global health crisis. To identify host factors required for infection by SARS-CoV-2 and seasonal coronaviruses, we designed a focused high-coverage CRISPR-Cas9 library targeting 332 members of a recently published SARS-CoV-2 protein interactome. We leveraged the compact nature of this library to systematically screen SARS-CoV-2 at two physiologically relevant temperatures along with three related coronaviruses (human coronavirus 229E [HCoV-229E], HCoV-NL63, and HCoV-OC43), allowing us to probe this interactome at a much higher resolution than genome-scale studies. This approach yielded several insights, including potential virus-specific differences in Rab GTPase requirements and glycosylphosphatidylinositol (GPI) anchor biosynthesis, as well as identification of multiple pan-coronavirus factors involved in cholesterol homeostasis. This coronavirus essentiality catalog could inform ongoing drug development efforts aimed at intercepting and treating coronavirus disease 2019 (COVID-19) and help prepare for future coronavirus outbreaks.

H. Heinrich Hoffmann, William M. Schneider, Francisco J. Sánchez-Rivera, Joseph M. Luna, Alison W. Ashbrook, Yadira M. Soto-Feliciano, Andrew A. Leal, Jérémie Le Pen, Inna Ricardo-Lax, Eleftherios Michailidis, Yuan Hao, Ansgar F. Stenzel, Avery Peace, C. David Allis, Scott W. Lowe, Margaret R. MacDonald, John T. Poirier, Charles M. Rice. Functional interrogation of a SARS-CoV-2 host protein interactome identifies unique and shared coronavirus host factors. *Cell Host & Microbe* **2021,** *29*(2), 267-280. doi: https://doi.org/10.1101/2020.09.11.291716

Dissection of c-AMP Response Element Architecture by Using Genomic and Episomal Massively Parallel Reporter Assays

Cell Systems, **2020**
doi: https://doi.org/10.1016/j.cels.2020.05.011

Authors

Jessica E. Davis, Kimberly D. Insigne, Eric M. Jones, Quinn A. Hastings, W. Clifford Boldridge, and Sriram Kosuri

Abstract

In eukaryotes, transcription factors (TFs) orchestrate gene expression by binding to TF-binding sites (TFBSs) and localizing transcriptional co-regulators and RNA polymerase II to cis-regulatory elements. However, we lack a basic understanding of the relationship between TFBS composition and their quantitative transcriptional responses. Here, we measured expression driven by 17,406 synthetic cis-regulatory elements with varied compositions of a model TFBS, the c-AMP response element (CRE) by using massively parallel reporter assays (MPRAs). We find CRE number, affinity, and promoter proximity largely determines expression. In addition, we observe expression modulation based on the spacing between CREs and CRE distance to the promoter, where expression follows a helical periodicity. Finally, we compare library expression between an episomal MPRA and a genomically integrated MPRA, where a single cis-regulatory element is assayed per cell at a defined locus. These assays largely recapitulate each other, although weaker, non-canonical CREs exhibit greater activity in a genomic context.

Davis et al., Dissection of c-AMP Response Element Architecture by Using Genomic and Episomal Massively Parallel Reporter Assays. *Cell Systems*, **2020** *11*, 75–85. Doi: https://doi.org/10.1016/j.cels.2020.05.011

Candidate Silencer Elements for the Human and Mouse Genomes

Nature Communications, **2020**
doi: https://doi.org/10.1038/s41467-020-14853-5

Authors

Naresh Doni Jayavelu, Ajay Jajodia, Arpit Mishra and R. David Hawkins

Abstract

The study of gene regulation is dominated by a focus on the control of gene activation or increase in the level of expression. Just as critical is the process of gene repression or silencing. Chromatin signatures have identified enhancers, however, genome-wide identification of silencers by computational or experimental approaches are lacking. Here, we first define uncharacterized cis-regulatory elements likely containing silencers and find that 41.5% of ~7500 tested elements show silencer activity using massively parallel reporter assay (MPRA). We trained a support vector machine classifier based on MPRA data to predict candidate silencers in over 100 human and mouse cell or tissue types. The predicted candidate silencers exhibit characteristics expected of silencers. Leveraging promoter-capture HiC data, we find that over 50% of silencers are interacting with gene promoters having very low to no expression. Our results suggest a general strategy for genome-wide identification and characterization of silencer elements.

Doni Jayavelu, N., Jajodia, A., Mishra, A. et al. Candidate silencer elements for the human and mouse genomes. *Nat Commun*, **2020**, *11*, 1061.
Doi: https://doi.org/10.1038/s41467-020-14853-5

A Deep Neural Network for Predicting and Engineering Alternative Polyadenylation

Cell, **2019**
doi: https://doi.org/10.1016/j.cell.2019.04.046

Authors

Nicholas Bogard, Johannes Linder, Alexander B. Rosenberg, and Georg Seelig

Abstract

Alternative polyadenylation (APA) is a major driver of transcriptome diversity in human cells. Here, we use deep learning to predict APA from DNA sequence alone. We trained our model (APARENT, APA REgression NeT) on isoform expression data from over 3 million APA reporters. APARENT's predictions are highly accurate when tasked with inferring APA in synthetic and human 3'UTRs. Visualizing features learned across all network layers reveals that APARENT recognizes sequence motifs known to recruit APA regulators, discovers previously unknown sequence determinants of 3' end processing, and integrates these features into a comprehensive, interpretable, *cis*-regulatory code. We apply APARENT to forward engineer functional polyadenylation signals with precisely defined cleavage position and isoform usage and validate predictions experimentally. Finally, we use APARENT to quantify the impact of genetic variants on APA. Our approach detects pathogenic variants in a wide range of disease contexts, expanding our understanding of the genetic origins of disease.

Nicholas Bogard, Johannes Linder, Alexander B. Rosenberg, Georg Seelig. A Deep Neural Network for Predicting and Engineering Alternative Polyadenylation. *Cell*, **2019**, Volume 178, Issue 1, Pages *91-106*.e23. doi: https://doi.org/10.1016/j.cell.2019.04.046.

Mutational Bias and the Protein Code Shape the Evolution of Splicing Enhancers

Nature Communications, **2020**
doi: https://doi.org/10.1038/s41467-020-16673-z

Authors

Stephen Rong, Luke Buerer, Christy L. Rhine, Jing Wang, Kamil J. Cygan, and William G. Fairbrother

Abstract

Exonic splicing enhancers (ESEs) are enriched in exons relative to introns and bind splicing activators. This study considers a fundamental question of co-evolution: How did ESE motifs become enriched in exons prior to the evolution of ESE recognition? We hypothesize that the high exon to intron motif ratios necessary for ESE function were created by mutational bias coupled with purifying selection on the protein code. These two forces retain certain coding motifs in exons while passively depleting them from introns. Through the use of simulations, genomic analyses, and high throughput splicing assays, we confirm the key predictions of this hypothesis, including an overlap between protein and splicing information in ESEs. We discuss the implications of mutational bias as an evolutionary driver in other *cis*-regulatory systems.

Rong, S., Buerer, L., Rhine, C.L. et al. Mutational bias and the protein code shape the evolution of splicing enhancers. *Nat Commun*, **2020**, *11*, 2845.
doi: https://doi.org/10.1038/s41467-020-16673-z

Functional Genomics of the Rapidly Replicating Bacterium *Vibrio Natriegens* by CRISPRi

Nature Microbiology. **2019**
doi: https://doi.org/10.1038/s41564-019-0423-8

Authors

Henry H. Lee, Nili Ostrov, Brandon G. Wong, Michaela A. Gold, Ahmad S. Khalil, and George M. Church

Abstract

The fast-growing Gram-negative bacterium *Vibrio natriegens* is an attractive microbial system for molecular biology and biotechnology due to its remarkably short generation time and metabolic prowess. However, efforts to uncover and utilize the mechanisms underlying its rapid growth are hampered by the scarcity of functional genomic data. Here, we develop a pooled genome-wide clustered regularly interspaced short palindromic repeats (CRISPR) interference (CRISPRi) screen to identify a minimal set of genes required for rapid wild-type growth. Targeting 4,565 (99.7%) of predicted protein-coding genes, our screen uncovered core genes comprising putative essential and growth-supporting genes that are enriched for respiratory pathways. We found that 96% of core genes were located on the larger chromosome 1, with growth-neutral duplicates of core genes located primarily on chromosome 2. Our screen also refines metabolic pathway annotations by distinguishing functional biosynthetic enzymes from those predicted on the basis of comparative genomics. Taken together, this work provides a broadly applicable platform for high-throughput functional genomics to accelerate biological studies and engineering of *V. natriegens*.

Lee HH, Ostrov N, Wong BG, Gold MA, Khalil AS, Church GM. Functional genomics of the rapidly replicating bacterium *Vibrio natriegens* by CRISPRi. *Nat Microbiol*. **2019**, Jul;*4*(7):1105-1113. PMID: 30962569. doi: https://doi.org/10.1038/s41564-019-0423-8.

DropSynth 2.0: High-Fidelity Multiplexed Gene Synthesis in Emulsions

Nucleic Acids Research, **2020**
doi: https://doi.org/10.1093/nar/gkaa600

Authors

Angus M. Sidore, Calin Plesa, Joyce A. Samson, Nathan B. Lubock, and Sriram Kosuri

Abstract

Multiplexed assays allow functional testing of large synthetic libraries of genetic elements, but are limited by the designability, length, fidelity and scale of the input DNA. Here, we improve DropSynth, a low-cost, multiplexed method that builds gene libraries by compartmentalizing and assembling microarray-derived oligonucleotides in vortexed emulsions. By optimizing enzyme choice, adding enzymatic error correction and increasing scale, we show that DropSynth can build thousands of gene-length fragments at >20% fidelity.

Angus M Sidore, Calin Plesa, Joyce A Samson, Nathan B Lubock, Sriram Kosuri. DropSynth 2.0: high-fidelity multiplexed gene synthesis in emulsions, *Nucleic Acids Research*, **2020**, Volume 48, Issue 16, 18, Page *e95*. Doi: https://doi.org/10.1093/nar/gkaa600

Agilent products are NOT authorized for COVID-19 testing, diagnosis, treatment, or mitigation. Agilent has not validated a product to detect the novel coronavirus.

For Research Use Only. Not for use in diagnostic procedures.
PR7000-2675

This information is subject to change without notice.

© Agilent Technologies, Inc. 2021
Published in the USA, July 1, 2021
5994-2946EN