Sidera --- A Platform Oncology Drug Discovery and Development Company

-Novel discovery tools to rapidly identify compounds affecting migration of aggressive cells

Oct 2017

The Sidera team has benefited greatly from attracting an advisor through AVX that has refined their R&D plan and is helping them raise angel funding. They are still very interested in additional advisors, and attracting a business lead. This detail from the team:

"Over the past year, Sidera Medicine has advanced its primary program (undisclosed target) with more successful experiments that validate its proprietary approach of identifying targets and lead compounds for aggressive cancers, and developed a business plan towards filing for an IND within two years.

Specifically, new experiments show definite proof that Sidera's discovery platform is able to separate a heterogeneous group of cancer cells into two phenotypes; (1) primarily proliferative cells, and (2) cancer stem cells that a are slow-cycling, less differentiated, invasive, and also drug resistant. This provides a powerful capability to develop targeted therapies against cancer stem cells and synergistic therapies that target both phenotypes.

The Sidera team has also gathered more experimental data in support of its two lead compounds as promising development candidates against aggressive glioblastoma. In vitro experiments with patient derived glioblastoma cells and normal human brain cells (astrocytes) show that the compounds effectively block the migration of glioblastoma cells in a dose-dependent manner, while not affecting normal human brain cells (i.e., migration of these cells is not blocked by the compounds). Furthermore, the ongoing POC in vivo studies in mice xenografted with human glioblastoma demonstrate inhibition of tumor growth after systemic daily administration of low doses of the lead compounds. During this process, no adverse physiological or neurological effects were observed in the animals. Currently, the POC in vivo studies are continuing to quantify inhibition of glioblastoma invasion in brain tissue, and to investigate the effect of the lead compounds on overall survival.

In preliminary experiments to expand the targeted indications, the compounds have also been tested on metastatic melanoma, showing effective inhibition of melanoma-cell migration in vitro.

To formalize a strategy to meet Sidera's near-term goals, a detailed business and financial plan was developed with the help of Sidera's current acting CEO Nolan Sigal. Specifically, the plan addresses the needs for building a company that can advance Sidera's primary program towards filing for an IND within the next two years, and develop a robust pipeline with more indications and more druggable targets of cancer stem cells."

Feb 2017

Sidera brings together leading experts and institutions (Yale, Johns Hopkins, Mayo) forming a multidisciplinary team to take on the challenge of identification of novel small molecules using a proprietary technology with glioblastoma as an initial indication with more to follow. The team has met with several VCs and strategic corporates to gain consensus driven feedback on the data set they need to drive the next round of funding discussions. The proof of concept in rodent models is currently underway.

The team is currently searching for a CEO as one founder, Dr. Onur Kilic, works full-time on driving a development and funding plan. Dr Kilic plans to join the company full-time shaping the scientific strategy. Collectively, the founding team has several decades of deep expertise across synthetic biology, medicinal chemistry, clinical oncology, and material science which led to the ongoing collaborative research focusing on challenging aggressive cancers such as glioblastoma. The specific focus has been on the most invasive cell types which resemble stem cells. They have explored the migratory pattern of the cells and created a micro-patterned lab-on-a-chip solution which differentiates for the most aggressive cell types.

The team felt currently available technology to understand the cell migration was inadequate and built a patented technology to mimic what is going on in vivo. They have demonstrated the prognostic capability of the platform by assessing fast migrating cells which lead to a clinically poor outcome. The molecular studies to single out and identify molecular signatures indicate the cells behave somewhat like stem cells and are not as responsive to chemo drugs. These learnings have directed the identification of small molecules that target the underlying pathway.

The second key to the company's technology is the ability to create custom phosphorylated proteins to activate human kinases drug targets in E.Coli. The genetic control switches have been patented and drives the discovery engine. This is a very challenging technique. Through traditional small molecule screening candidates are identified, however by leveraging Sidara's migration platform technology they can uniquely select the best candidates to advance. This has allowed for previously challenging or nearly impossible drug targets in the signaling pathway to be selected.

The team is well positioned with Dr. Quinones-Hinojosa, the Chair of Neurosurgery at Mayo Clinic to develop and select the best delivery mechanism to cross the blood brain barrier once the leading drug candidate is selected.

In summary, the team has advanced nearly a decade of basic research into a translational strategy that has the potential to address the unmet needs of the most aggressive cancers.

Published in final edited form as:

Science. 2011 August 26; 333(6046): 1151–1154. doi:10.1126/science.1207203.

Expanding the Genetic Code of *Escherichia coli* with Phosphoserine

Hee-Sung Park[1,*,†], **Michael J. Hohn**[1,*], **Takuya Umehara**[1], **Li-Tao Guo**[1], **Edith M. Osborne**[2,‡], **Jack Benner**[2], **Christopher J. Noren**[2,§], **Jesse Rinehart**[3,4,§], and **Dieter Söll**[1,5,§]

[1]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA

[2]New England BioLabs, Ipswich, MA 01938, USA

[3]Department of Cellular and Molecular Physiology, Yale University, New Haven, CT 06520, USA

[4]Systems Biology Institute, Yale University, West Haven, CT 06516, USA

[5]Department of Chemistry, Yale University, New Haven, CT 06520, USA

Abstract

O-Phosphoserine (Sep), the most abundant phosphoamino acid in the eukaryotic phosphoproteome, is not encoded in the genetic code, but synthesized posttranslationally. Here, we present an engineered system for specific cotranslational Sep incorporation (directed by UAG) into any desired position in a protein by an *Escherichia coli* strain that harbors a Sep-accepting transfer RNA ($tRNA^{Sep}$), its cognate Sep–tRNA synthetase (SepRS), and an engineered EF-Tu (EF-Sep). Expanding the genetic code rested on reengineering EF-Tu to relax its quality-control function and permit Sep-$tRNA^{Sep}$ binding. To test our system, we synthesized the activated form of human mitogen-activated ERK activating kinase 1 (MEK1) with either one or two Sep residues cotranslationally inserted in their canonical positions (Sep^{218}, Sep^{222}). This system has general utility in protein engineering, molecular biology, and disease research.

O-Phosphoserine (Sep) was identified 80 years ago as a constituent of phosphoproteins from egg yolk (1). Since then, the extent and importance of the eukaryotic phosphoproteome has been realized and has provided insight into large interconnected networks of kinases and phosphatases (2). Protein kinases represent one of the largest eukaryotic gene families, composing nearly 2% of all human genes (3). Sep is the most abundant phosphoamino acid; based on an analysis of >2000 HeLa cell phosphoproteins, the relative abundance of Sep is

[§]To whom correspondence should be addressed. dieter.soll@yale.edu (D.S.); noren@neb.com (C.J.N.); jesse. rinehart@yale.edu (J.R.).

[*]These authors contributed equally to this work.

[†]Present address: Department of Chemistry, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea.

[‡]Present address: Department of Chemistry, Angelo State University, San Angelo, TX 76909, USA.

Supporting Online Material

www.sciencemag.org/cgi/content/full/333/6046/1151/DC1

Materials and Methods

Figs. S1 and S2

Table S1

References (*27*–*37*)

7.3 and 48 times higher than that of phosphothreonine and phosphotyrosine, respectively (2). A major research limitation is the inability to biosynthesize these phosphoproteins for detailed studies of their enzyme or substrate properties.

The discovery of Sep–tRNA synthetase (SepRS), a unique aminoacyl–tRNA synthetase devoted to Sep-tRNA^{Cys} formation in methanogenic archaea, provided an opportunity to develop our Sep-insertion strategy. The natural role of SepRS is the formation of Sep-tRNA^{Cys}, which is then converted to Cys-tRNA^{Cys} by the enzyme SepCysS in the presence of a sulfur donor (4) (Fig. 1A). Given the high specificity of SepRS for Sep and for tRNA^{Cys} and our knowledge of the identity elements in this tRNA (5), and based on our understanding of the structure of this enzyme and its catalytic site (6, 7), we devised a system to incorporate Sep into proteins directed by the UAG (amber) codon. For this, we chose *Methanocaldococcus jannaschii* (Mj) tRNA^{Cys} and the mesophilic *Methanococcus maripaludis* (Mmp) SepRS as the orthogonal pair [reviewed in (8)] for the synthesis of phosphoserylated amber suppressor tRNA.

We first designed tRNA^{Sep} (Fig. 1B), an amber suppressor tRNA derived from Mj tRNA^{Cys} by two mutations in the anticodon, and an additional C20U change that improves aminoacylation by SepRS (5). In vitro aminoacylation by Mmp SepRS showed (Fig. 1C) that the anticodon change lowered (to about 40%) the ability of tRNA^{Sep} to be aminoacylated when compared to tRNA^{Cys}. In agreement with earlier data (5), total *Escherichia coli* tRNA could not be acylated with Sep (Fig. 1C). On the basis of these in vitro data, Mj tRNA^{Sep} and Mmp SepRS appear to be an orthogonal pair.

Efficient and selective addition of Sep to the *E. coli* genetic repertoire requires exclusive interaction of SepRS with tRNA^{Sep} for Sep-tRNA^{Sep} formation without interfering in the host translation system, as well as a sufficient intracellular concentration of Sep. Because *E. coli* has a Sep-compatible transporter (9), Sep (2 mM) was added to the LB growth medium, and the endogenous *serB* gene encoding phosphoserine phosphatase was deleted in the *E. coli* test strain without affecting growth. To assess whether the Mj tRNA^{Sep}–Mmp SepRS pair is functional and orthogonal in *E. coli* in vivo, we performed a suppression assay that used a gene encoding chloramphenicol acetyltransferase (CAT) with a UAG stop codon at the permissive position 112 (wild-type amino acid: Asp) to produce the CAT enzyme. Cell survival was measured in the presence of Sep and varying amounts of chloramphenicol (Cm) where the different half-maximal inhibitory concentration (IC_{50}) values and the tRNA^{Sep}-dependent CATsynthesis correlate with suppression efficiency (Fig. 2). When only tRNA^{Sep} was expressed (Fig. 2, column B), Cm resistance increased about 3.3-fold over background (Fig. 2, column A). Thus, tRNA^{Sep} can be aminoacylated to a certain degree by an unknown *E. coli* aminoacyl–tRNA synthetase (we later found that Gln is being incorporated at the amber stop codon). In contrast, simultaneous expression of tRNA^{Sep} and SepRS did not provide Cm resistance (Fig. 2, column C). This may indicate that SepRS can outcompete any endogenous aminoacyl–tRNA synthetase and form Sep-tRNA^{Sep}; however, this aminoacyl-tRNA is neither delivered to the ribosome nor accommodated on it. Providing additional EF-Tu did not improve the result (Fig. 2, column D). Coexpression of tRNA^{Sep}, SepRS, and SepCysS should result in formation of Sep-tRNA^{Sep} and subsequent SepCysS-mediated conversion to Cys-tRNA^{Sep} (4). Indeed, a 2.3-fold increase in Cm resistance was

observed (Fig. 2, column E). This further supports the notion that although Sep-tRNA^{Sep} is synthesized, it cannot be used properly by the *E. coli* protein biosynthesis machinery. By contrast, coexpression of tRNA^{Sep} and Mmp CysRS generated a 12.3-fold increase in Cm resistance (Fig 2, column F), demonstrating that Cys-tRNA^{Sep} can be readily used for amber codon suppression in the CAT gene.

Given that EF-Tu is a component of quality control in protein synthesis (10), it is plausible that Sep-tRNA^{Sep} may be rejected by EF-Tu. Chemically synthesized Sep-tRNA^{Gln} was a poor substrate for in vitro protein synthesis (11). tRNAs carrying negatively charged amino acids are bound poorly by EF-Tu (12), and molecular dynamics simulations suggested that Sep-tRNA^{Cys} may not be bound by EF-Tu (13). We tested this assumption in EF-Tu–mediated Sep-tRNA hydrolysis protection experiments (14). Although EF-Tu protected [^{35}S]Cys-tRNA^{Cys} from deacylation at pH 8.2 (fig. S1), [^{14}C]Sep-tRNA^{Cys} was significantly deacylated irrespective of the presence of EF-Tu (Fig. 3B and fig. S1). Thus, insufficient binding of Sep-tRNA^{Sep} to EF-Tu may explain the lack of Sep insertion into protein.

This observation required the generation of EF-Tu variants able to productively bind Sep-tRNA. We were encouraged by reports that EF-Tu variants allow binding of tRNAs charged with certain unusual amino acids (15, 16). Guided by the structure of the *E. coli* EF-Tu:Phe-tRNA^{Phe} complex (17), we selected six residues (His^{67}, Asp^{216}, Glu^{217}, Phe^{219}, Thr^{229}, and Asn^{274}) in the amino acid binding pocket of EF-Tu (Fig. 3A) for complete randomization in order to generate EF-Tu variants that bind Sep-tRNA. Variants that permitted SepRS and tRNA^{Sep}-dependent Sep incorporation were selected in vivo (see SOM). One clone, designated EF-Sep (amino acid variants shown in Fig. 3A), led to a 10-fold increase in Cm resistance (Fig. 2, column H), whereas the combination of SepRS and EF-Sep without tRNA^{Sep} was not active in the CAT suppression assay (Fig. 2, column G). Thus, it appeared that EF-Sep did bind Sep-tRNA^{Sep}, a conclusion that was confirmed in the hydrolysis protection assay (Fig. 3B). This assay also shows that EF-Sep still retained some ability to bind Cys-tRNA (fig. S1).

To prove that the observed suppression is due to Sep incorporation, we expressed myoglobin with an amber codon in the Asp^{127} position (fig. S2A). The expected full-length protein was synthesized only when EF-Sep, SepRS, and tRNA^{Sep} were coexpressed (fig. S2A). Mass spectrometry–time-of-flight (MS-TOF) and MS/MS analysis showed that Sep is present at the position specified by UAG in both the intact and trypsin-digested proteins (Fig. S2, B and C).

Final validation of our strategy was the synthesis of a Sep-containing human protein MEK1 (mitogen-activated ERK activating kinase 1). This key eukaryotic enzyme of the mitogen-activated signaling cascade is crucial for cell proliferation, development, differentiation, cell cycle progression, and oncogenesis (18). Activation of MEK1 requires posttranslational phosphorylation of Ser^{218} and Ser^{222} by MEK activating kinases (e.g., Raf-1, MEKK, or MOS). Substitution of both Ser residues with Glu yields a constitutively active enzyme, albeit with lower activity (19). We generated a clone encoding a MEK1 fusion protein [with the maltose binding protein (MBP) at the N terminus and a His_6 tag at the C terminus] in which Ser^{222} was changed to Glu, and the Ser^{218} codon was replaced by UAG to encode

Sep. After expression in the presence of SepRS, $tRNA^{Sep}$, and EF-Sep, 25 μg of full-length MBP-MEK1(Sep^{218}, Glu^{222}) was isolated from 1 liter of culture. The presence of Sep in MBP-MEK1(Sep^{218},Glu^{222}) protein was demonstrated by its ability to phosphorylate ERK2, which then phosphorylates myelin basic protein. MBP-MEK1(Sep^{218},Glu^{222}) had a 2500-fold higher specific activity than MBP-MEK1(Ser^{218},Ser^{222}) and a 70-fold higher specific activity than the constitutively active MBP-MEK1(Glu^{218},Glu^{222}) (Fig. 4A). MS/MS analysis confirmed the incorporation of Sep at position 218 (Fig. 4B). To determine if our *E. coli* expression system would allow the simultaneous insertion of two Sep residues into the protein, we changed the Ser codons in positions 218 and 222 to UAG. As expected, the expression efficiency of MBP-MEK1(Sep^{218},Sep^{222}) was markedly reduced compared to that of wild-type MBP-MEK1 (only about 1 μg of full-length protein was obtained from 1 liter of culture). The presence of Sep at both active-site positions of MEK1 was tested by Western blot analysis with a monoclonal antibody specific to phosphorylation at these two residues (Fig 4C). Only MBP-MEK1(Sep^{218},Sep^{222}), and to a weaker extent MBP-MEK1(Sep^{218},Ser^{222}), were detected in this experiment, whereas neither MBP-MEK1(Ser^{218},Ser^{222}), MBP-MEK(Sep^{218},Glu^{222}), or MBP-MEK(Glu^{218},Glu^{222}) was recognized by the antibody. This demonstrates that the addition of SepRS, $tRNA^{Sep}$, and EF-Sep endows *E. coli* with the ability to read UAG as a phosphoserine codon.

Our work underscores the key role of EF-Tu in quality control of protein synthesis by ensuring facile delivery of the correct cognate aminoacyl-tRNA to the ribosome (10, 20). Generating an orthogonal aminoacyl–tRNA synthetase:tRNA pair was insufficient to genetically encode Sep. Expansion of the genetic code to include Sep depended critically on reengineering of EF-Tu to bind Sep-$tRNA^{Sep}$. This situation is precisely paralleled in the naturally evolved genetic encoding system for selenocysteine, which requires a specialized elongation factor [SelB in prokaryotes, and EFSec in eukaryotes; reviewed in (21)]. Inspired careful manipulation of components of the protein-synthesizing system will allow further expansion of the genetic code without sacrificing organismal fitness.

Orthogonal aminoacyl–tRNA synthetase:tRNA pairs (in our case, SepRS:$tRNA^{Sep}$) are critical elements for genetic code expansion, whether produced by evolutionary processes and found in nature [e.g., pyrrolysyl–tRNA synthetase:$tRNA^{Pyl}$ (22)] or designed in the laboratory by synthetic biologists [reviewed in (8)]. The tRNA of choice is typically an amber suppressor tRNA; however, a general limitation in product yield is associated with recoding the new amino acid by a stop codon (e.g., UAG), because peptide chain elongation by the designed aminoacylated suppressor tRNA competes with chain termination by the release factor (e.g., RF1). Yet this impediment may soon be removed by the advent of *E. coli* expression strains with genome-wide UAG codon reassignments and an RF1 deletion (23, 24).

The ability to generate physiologically relevant active kinases and stoichiometrically phosphorylated protein domains could reveal new types of kinase inhibitors for drug development, allow systematic dissection of phosphorylation-dependent protein-protein interactions, and expose new structure-function relationships (18, 25, 26).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank P. Dennis, P. O'Donoghue, and J. Ling for enthusiastic discussions. M.J.H. was a Feodor Lynen Postdoctoral Fellow of the Alexander von Humboldt Foundation (Bonn, Germany). H.-S.P. held a postdoctoral fellowship of the Korean Science Foundation. This work was supported by grants from NSF (MCB-0645283 and MCB-0950474) (to D.S.), National Institute of General Medical Sciences (GM 22854) (to D.S.), National Research Foundation of Korea (ABC-20100029737) (to H.-S.P.), and National Institute of Diabetes and Digestive and Kidney Diseases (K01DK089006) (to J.R.). Yale University holds the U.S. Patent 7,723,069 B2: "Site Specific Incorporation of Phosphoserine into Polypeptides Using Phosphoseryl-tRNA Synthetase" by D. Soll and J. Rinehart. Yale University has applied for a patent that covers the engineered EF-Tu described in this manuscript. Reagents are available under a Yale Material Transfer Agreement.

References and Notes

1. Lipmann F. Biochem Z. 1933; 262:3.
2. Olsen JV, et al. Cell. 2006; 127:635. [PubMed: 17081983]
3. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. Science. 2002; 298:1912. [PubMed: 12471243]
4. Sauerwald A, et al. Science. 2005; 307:1969. [PubMed: 15790858]
5. Hohn MJ, Park HS, O'Donoghue P, Schnitzbauer M, Söll D. Proc Natl Acad Sci USA. 2006; 103:18095. [PubMed: 17110438]
6. Kamtekar S, et al. Proc Natl Acad Sci USA. 2007; 104:2620. [PubMed: 17301225]
7. Fukunaga R, Yokoyama S. Nat Struct Mol Biol. 2007; 14:272. [PubMed: 17351629]
8. Liu CC, Schultz PG. Annu Rev Biochem. 2010; 79:413. [PubMed: 20307192]
9. Wanner BL, Metcalf WW. FEMS Microbiol Lett. 1992; 79:133.
10. LaRiviere FJ, Wolfson AD, Uhlenbeck OC. Science. 2001; 294:165. [PubMed: 11588263]
11. Rothman DM, et al. J Am Chem Soc. 2005; 127:846. [PubMed: 15656617]
12. Dale T, Sanderson LE, Uhlenbeck OC. Biochemistry. 2004; 43:6159. [PubMed: 15147200]
13. Eargle J, Black AA, Sethi A, Trabuco LG, Luthey-Schulten Z. J Mol Biol. 2008; 377:1382. [PubMed: 18336835]
14. Ling J, et al. Proc Natl Acad Sci USA. 2007; 104:15299. [PubMed: 17878308]
15. Doi Y, Ohtsuki T, Shimizu Y, Ueda T, Sisido M. J Am Chem Soc. 2007; 129:14458. [PubMed: 17958427]
16. Ohtsuki T, Yamamoto H, Doi Y, Sisido M. J Biochem. 2010; 148:239. [PubMed: 20519322]
17. Nissen P, et al. Science. 1995; 270:1464. [PubMed: 7491491]
18. Sebolt-Leopold JS, Herrera R. Nat Rev Cancer. 2004; 4:937. [PubMed: 15573115]
19. Alessi DR, et al. EMBO J. 1994; 13:1610. [PubMed: 8157000]
20. Schrader JM, Chapman SJ, Uhlenbeck OC. Proc Natl Acad Sci USA. 2011; 108:5215. [PubMed: 21402928]
21. Yoshizawa S, Böck A. Biochim Biophys Acta. 2009; 1790:1404. [PubMed: 19328835]
22. Nozawa K, et al. Nature. 2009; 457:1163. [PubMed: 19118381]
23. Mukai T, et al. Nucleic Acids Res. 2010; 38:8188. [PubMed: 20702426]
24. Isaacs FJ, et al. Science. 2011; 333:348. [PubMed: 21764749]
25. Scott JD, Pawson T. Science. 2009; 326:1220. [PubMed: 19965465]
26. Yaffe MB, et al. Cell. 1997; 91:961. [PubMed: 9428519]

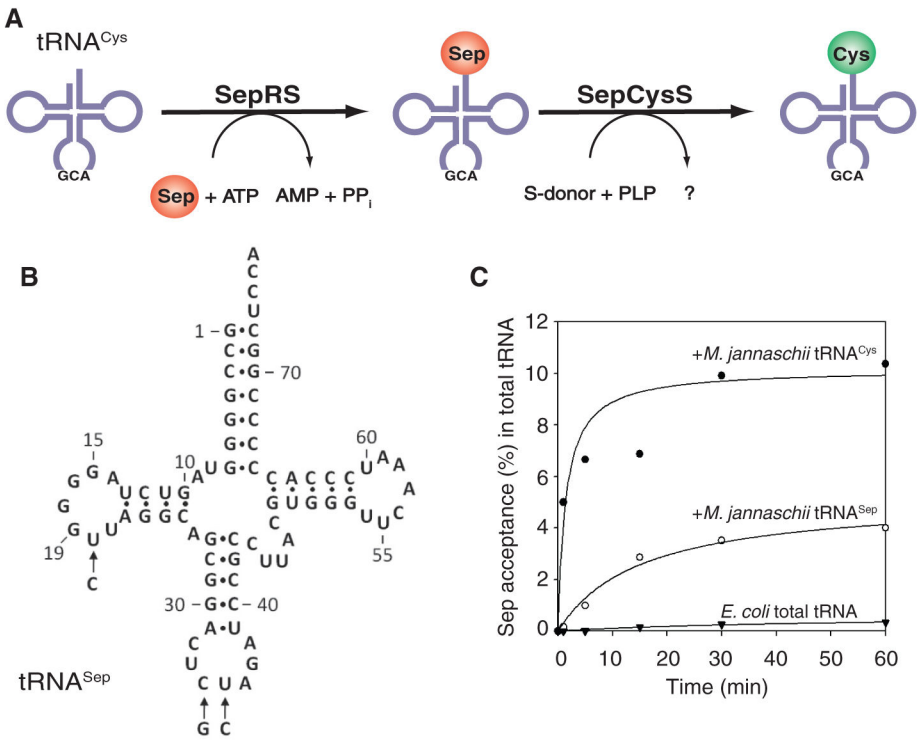


Fig. 1.

Design of $tRNA^{Sep}$ and its aminoacylation by *M. maripaludis* SepRS. (**A**) Pathway of Cys-$tRNA^{Cys}$ formation in *M. maripaludis*. ATP, adenosine 5′-triphosphate; AMP, adenosine 5′-monophosphate; PPi, inorganic pyrophosphate; PLP, pyridoxal phosphate. (**B**) Cloverleaf structure of $tRNA^{Sep}$. Arrows indicate the three nucleotide changes compared to *M. jannaschii* $tRNA^{Cys}$. (**C**) Acylation with Sep of *M. jannaschii* $tRNA^{Cys}$ and $tRNA^{Sep}$ catalyzed by *M. maripaludis* SepRS. Total *E. coli* tRNA (▼), or tRNA from *E. coli* strains expressing the *M. jannaschii* $tRNA^{Cys}$ (●) or the $tRNA^{Sep}$ (○) gene was acylated by *M. maripaludis* SepRS with [^{14}C]Sep (0.1 mM) in the presence of ATP (10 mM).

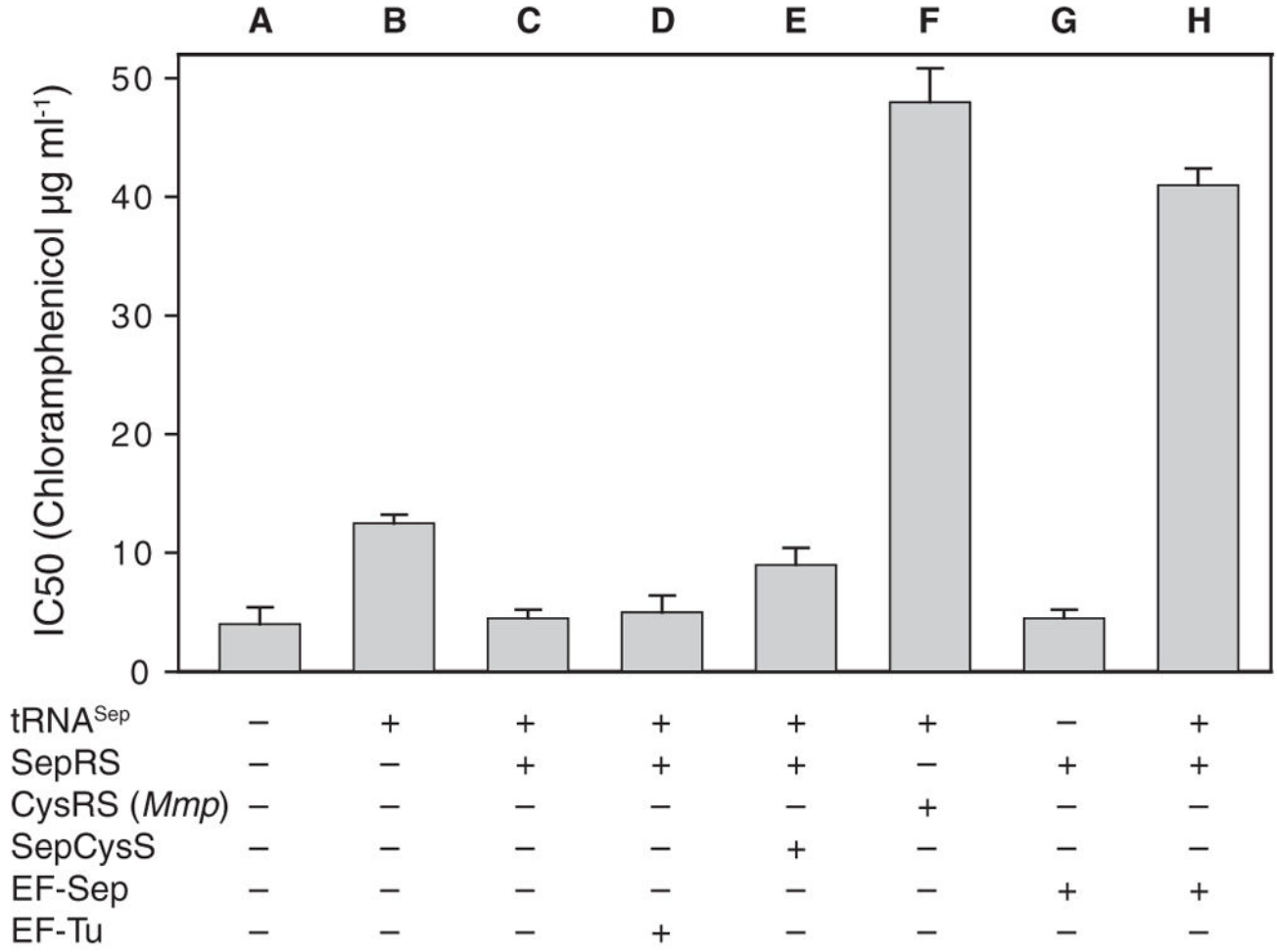


Fig. 2.

In vivo synthesis of chloramphenicol acetyl-transferase (measured by IC_{50} value) by $tRNA^{Sep}$-dependent read-through of an amber mutation in the CAT gene. $tRNA^{Sep}$ was co-expressed in the *E. coli* Top10Δ*serB* strain with the proteins indicated in the figure (SepRS, *M. maripaludis* CysRS, SepCysS, EF-Sep, EF-Tu). Selection was carried out on LB agar plates containing 2 mM Sep and various concentrations of chloramphenicol. Error bars indicate SEM.

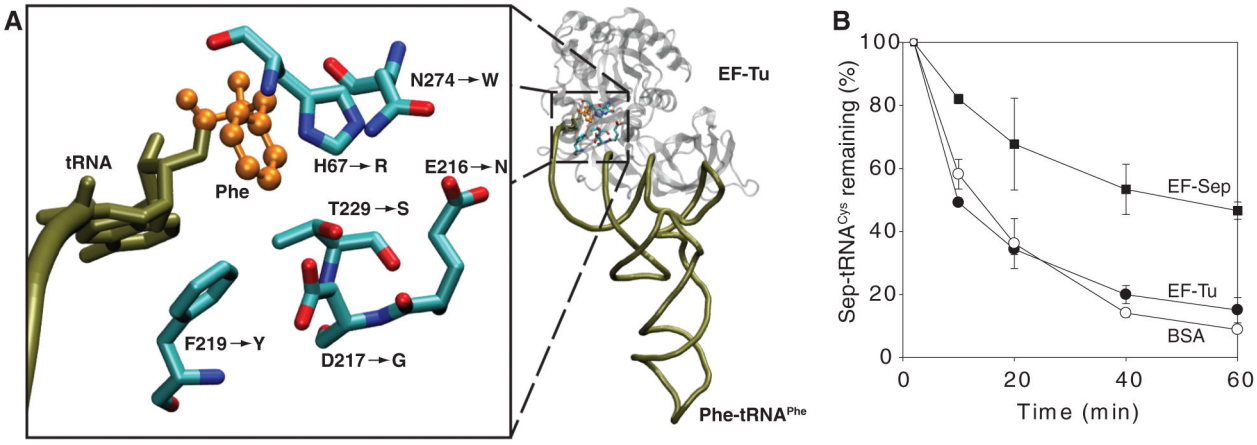


Fig. 3.

Design of EF-Sep. (**A**) Model of the amino acid binding pocket of *E. coli* EF-Tu bound to Phe-tRNA (based on Protein Data Bank structure 1OB2). To accommodate Sep-tRNA, an *E. coli tufB* library was constructed that would allow the six highlighted amino acid residues to change to any of the 20 canonical amino acids. The six mutations in our final EF-Sep are indicated by an arrow. (**B**) EF-Sep protects Sep-tRNA^{Cys} from deacylation. *M. jannaschii* [^{14}C]Sep-tRNA^{Cys} was incubated at pH 8.2 and at room temperature in the presence or absence of wild-type EF-Tu or EF-Sep. Deacylation of Sep-tRNA^{Cys} was measured by acid precipitability. Error bars indicate SEM. Abbreviations for the amino acid residues are as follows: D, Asp; E, Glu; F, Phe; G, Gly; H, His; N, Asn; R, Arg; S, Ser; T, Thr; W, Trp; and Y, Tyr.

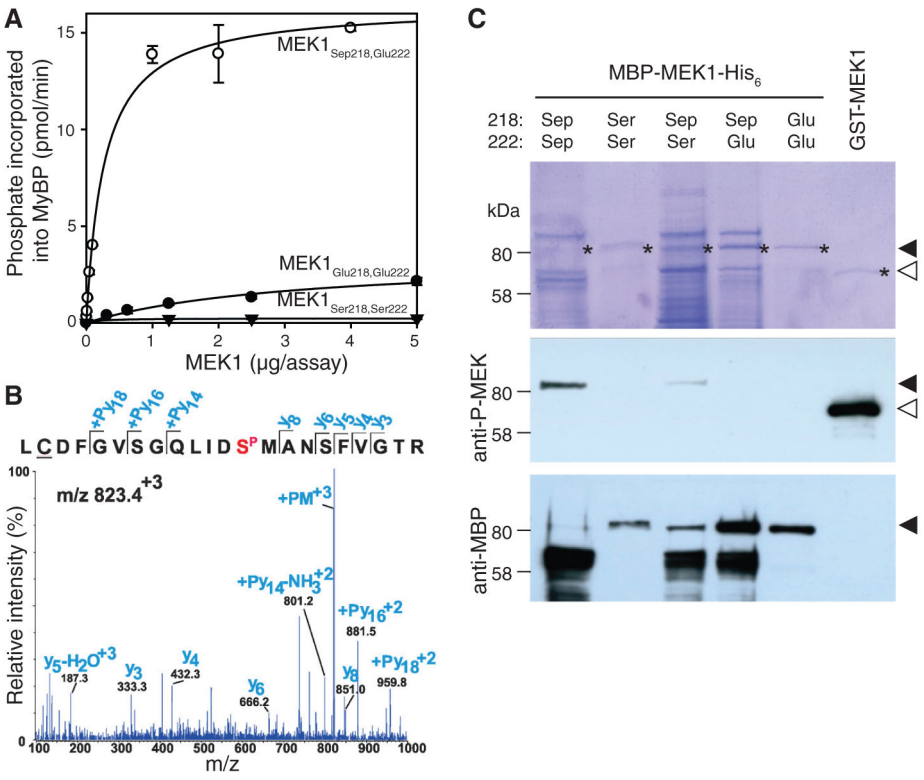


Fig. 4.

Properties of *E. coli*–produced Sep-containing human MEK1. (**A**) Kinase activity assay. MEK1 was produced in *E. coli* as a fusion protein with an N-terminal maltose-binding protein (MBP) tag and a C-terminal His_6 tag. Residues Ser^{218} and Ser^{222}, which are targets to phosphorylation by MEK1 activators, were mutated to either Glu^{218}/Glu^{222} or Sep^{218}/Glu^{222} to produce active MEK1 variants. Various amounts of MBP-MEK1-His_6 with active-site residues Ser^{218}/Ser^{222} (▼), Glu^{218}/Glu^{222} (●), or Sep^{218}/Glu^{222} (○) were used to phosphorylate inactive ERK2 in vitro. ERK2 activity was then measured in a radiometric assay using γ-[^{32}P]ATP and myelin basic protein (MyBP) as substrates. Error bars indicate SEM. (**B**) MS/MS spectrum confirming the presence of Sep^{218} in MBP-MEK1(Sep^{218}/Ser^{222}). *m*/*z*, mass-to-charge ratio. (**C**) MBP-MEK1-His_6 variants with genetically encoded active-site residues Sep^{218}/Sep^{222} (lane 1), Ser^{218}/Ser^{222} (lane 2), Sep^{218}/Ser^{222} (lane 3), Sep^{218}/Glu^{222} (lane 4), and Glu^{218}/Glu^{222} (lane 5) were produced in *E. coli* and partially purified by Ni^{2+} affinity chromatography. Proteins were separated by SDS–polyacrylamide gel electrophoresis and either stained with Coomassie (top) or transferred to a nylon membrane and detected with monoclonal antibodies specific for the phosphorylated active site of human MEK (center) or the MBP tag (bottom). Purchased activated glutathione *S*-transferase (GST)–MEK was used as a control (lane 6). Dark and light arrowheads indicate the positions of MBP-MEK1-His_6 and GST-MEK1, respectively. The strong bands (~70 kD size) are probable truncation products caused by termination at UAG.

Published in final edited form as:

Science. 2013 October 18; 342(6156): 357–360. doi:10.1126/science.1241459.

Genomically Recoded Organisms Expand Biological Functions

Marc J. Lajoie[1,2], **Alexis J. Rovner**[3,4], **Daniel B. Goodman**[1,5], **Hans-Rudolf Aerni**[4,6], **Adrian D. Haimovich**[3,4], **Gleb Kuznetsov**[1], **Jaron A. Mercer**[7], **Harris H. Wang**[8], **Peter A. Carr**[9], **Joshua A. Mosberg**[1,2], **Nadin Rohland**[1], **Peter G. Schultz**[10], **Joseph M. Jacobson**[11,12], **Jesse Rinehart**[4,6], **George M. Church**[1,13,*], and **Farren J. Isaacs**[3,4,*]

[1]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

[2]Program in Chemical Biology, Harvard University, Cambridge, MA 02138, USA

[3]Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06520, USA

[4]Systems Biology Institute, Yale University, West Haven, CT 06516, USA

[5]Program in Medical Engineering and Medical Physics, Harvard–Massachusetts Institute of Technology (MIT) Division of Health Sciences and Technology, Cambridge, MA 02139, USA

[6]Department of Cellular and Molecular Physiology, Yale University, New Haven, CT 06520, USA

[7]Harvard College, Cambridge, MA 02138, USA

[8]Department of Systems Biology, Columbia University, College of Physicians and Surgeons, New York, NY 10032, USA

[9]MIT Lincoln Laboratory, Lexington, MA 02420, USA

[10]Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037, USA

[11]Center for Bits and Atoms, MIT, Cambridge, MA 02139, USA

[12]MIT Media Lab, MIT, Cambridge, MA 02139, USA

[13]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA

Abstract

We describe the construction and characterization of a genomically recoded organism (GRO). We replaced all known UAG stop codons in *Escherichia coli* MG1655 with synonymous UAA codons, which permitted the deletion of release factor 1 and reassignment of UAG translation function. This GRO exhibited improved properties for incorporation of nonstandard amino acids that expand the chemical diversity of proteins in vivo. The GRO also exhibited increased resistance to T7 bacteriophage, demonstrating that new genetic codes could enable increased viral resistance.

*Corresponding author. farren.isaacs@yale.edu (F.J.I.); gchurch@genetics.med.harvard.edu (G.M.C.).

Supplementary Materials

www.sciencemag.org/content/342/6156/357/suppl/DC1

Materials and Methods

Figs. S1 to S22

Tables S1 to S37

References (26–70)

The conservation of the genetic code permits organisms to share beneficial traits through horizontal gene transfer (1) and enables the accurate expression of heterologous genes in nonnative organisms (2). However, the common genetic code also allows viruses to hijack host translation machinery (3) and compromise cell viability. Additionally, genetically modified organisms (GMOs) can release functional DNA into the environment (4). Virus resistance (5) and biosafety (6) are among today's major unsolved problems in biotechnology, and no general strategy exists to create genetically isolated or virus-resistant organisms. Furthermore, biotechnology has been limited by the 20 amino acids of the canonical genetic code, which use all 64 possible triplet codons, limiting efforts to expand the chemical properties of proteins by means of nonstandard amino acids (NSAAs) (7, 8).

Changing the genetic code could solve these challenges and reveal new principles that explain how genetic information is conserved, encoded, and exchanged (fig. S1). We propose that genomically recoded organisms (GROs, whose codons have been reassigned to create an alternate genetic code) would be genetically isolated from natural organisms and viruses, as horizontally transferred genes would be mistranslated, producing nonfunctional proteins. Furthermore, GROs could provide dedicated codons to improve the purity and yield of NSAA-containing proteins, enabling robust and sustained incorporation of more than 20 amino acids as part of the genetic code.

We constructed a GRO in which all instances of the UAG codon have been removed, permitting the deletion of release factor 1 (RF1; terminates translation at UAG and UAA) and, hence, eliminating translational termination at UAG codons. This GRO allows us to reintroduce UAG codons, along with orthogonal translation machinery [i.e., aminoacyl–tRNA synthetases (aaRSs) and tRNAs] (7, 9), to permit efficient and site-specific incorporation of NSAAs into proteins (Fig. 1). That is, UAG has been transformed from a nonsense codon (terminates translation) to a sense codon (incorporates amino acid of choice), provided the appropriate translation machinery is present. We selected UAG as our first target for genome-wide codon reassignment because UAG is the rarest codon in *Escherichia coli* MG1655 (321 known instances), prior studies (7, 10) demonstrated the feasibility of amino acid incorporation at UAG, and a rich collection of translation machinery capable of incorporating NSAAs has been developed for UAG (7).

We used an in vivo genome-editing approach (11), which is more efficient than de novo genome synthesis at exploring new genotypic landscapes and overcoming genome design flaws. Although a single lethal mutation can prevent transplantation of a synthetic genome (12), our approach allowed us to harness genetic diversity and evolution to overcome any potential deleterious mutations at a cost considerably less than de novo genome synthesis (supplementary text section B, "Time and cost"). In prior work, we used multiplex automated genome engineering [MAGE (13)] to remove all known UAG codons in groups of 10 across 32 *E. coli* strains (11), and conjugative assembly genome engineering [CAGE (11)] to consolidate these codon changes in groups of ~80 across four strains. In this work, we overcome technical hurdles (supplementary text) to complete the assembly of the GRO and describe the biological properties derived from its altered genetic code.

The GRO [C321.ΔA, named for 321 UAG→UAA conversions and deletion of *prfA* (encodes RF1, Table 1)] and its $RF1^+$ precursor (C321) exhibit normal prototrophy and morphology (fig. S2), with 60% increased doubling time compared with *E. coli* MG1655 (table S1). Genome sequencing [GenBank accession CP006698] confirmed that all 321 known UAGs were removed from its genome and that 355 additional mutations were acquired during construction (10^{-8} mutations per base pair per doubling over ~7340 doublings; fig. S3 and tables S2 to S4). Although maintaining the *E. coli* MG1655 genotype was not a primary goal of this work, future applications requiring increased genome stability could exploit reversible switching of *mutS* function (14) to reduce off-target mutagenesis. CAGE improved the fitness of several strains in the C321 lineage (fig. S3), implicating off-target mutations in the reduced fitness.

C321.ΔA exhibited improved performance compared with previous strategies for UAG codon reassignment (15, 16), permitting the complete reassignment of UAG from a stop codon to a sense codon capable of incorporating NSAAs into proteins. One previous strategy used a variant of release factor 2 (RF2) that exhibits enhanced UAA termination (16) and weak UAG termination (17). The second strategy substituted a UAA stop codon in each of the seven essential genes naturally terminating with UAG (table S5) and reduced ribosome toxicity by efficiently incorporating amino acids at the remaining 314 UAGs (15). For comparative purposes, we used MAGE to create strains C0.B*.ΔA::S [expresses enhanced RF2 variant (16)], C7.ΔA::S (UAG changed to UAA in seven essential genes), and C13.ΔA::S [UAG changed to UAA in seven essential genes plus six nonessential genes (table S5)] (Table 1). C refers to the number of codon changes, while A and B refer to *prfA* (RF1) and *prfB* (RF2) manipulations, respectively. In contrast to previous work (15), we deleted RF1 in these strains without introducing a UAG suppressor, perhaps because near-cognate suppression is increased in *E. coli* MG1655 (18). Nevertheless, these strains exhibited a strong selective pressure to acquire UAG suppressor mutations (see below).

To assess the fitness effects of RF1 removal and UAG reassignment, we measured the doubling time and maximum cell density of each strain (table S1 and fig. S4). We found that C321 was the only strain for which RF1 removal and UAG reassignment was not deleterious (Fig. 2). Because we did not modify RF2 to enhance UAA termination (16), this confirms that RF1 is essential only for UAG translational termination and not for UAA termination or other essential cellular functions. By contrast, RF1 removal significantly impaired fitness for C0.B*.ΔA::S, and codon reassignment exacerbated this effect (Fig. 2 and fig. S5), probably because NSAA incorporation outcompeted the weak UAG termination activity (17) exerted by the RF2 variant (16). C7.ΔA::S and C13.ΔA::S also exhibited strongly impaired fitness, likely due to more than 300 nonessential UAG codons stalling translation in the absence of RF1-mediated translation at UAG codons (15); accordingly, *p*-acetylphenylalanine (pAcF) incorporation partially alleviated this effect (Fig. 2). However, not all NSAAs improved fitness in partially recoded strains; phosphoserine (Sep) impairs fitness in similar strains (19), perhaps by causing proteome-scale misfolding. Together, these results indicate that only the complete removal of all instances of the UAG codon overcomes these deleterious effects; therefore, it may be the only scalable strategy for sustained NSAA translation and for complete reassignment of additional codons.

We tested the capacity of our recoded strains to efficiently incorporate NSAAs [pAcF, *p*-azidophenylalanine (pAzF), or 2-naphthalalanine (NapA)] into green fluorescent protein (GFP) variants containing zero, one, or three UAG codons (Fig. 3 and fig. S6). In the presence of NSAAs, the $RF1^+$ strains efficiently read through variants containing three UAGs, demonstrating that the episomal pEVOL translation system, which expresses an aaRS and tRNA that incorporate a NSAA at UAG codons (9), is extremely active and strongly outcompetes RF1. In the absence of NSAAs, the $RF1^-$ strains exhibited detectable amounts of near-cognate suppression (18) of a single UAG. C321.ΔA::S exhibited strong expression of UAG-containing GFP variants only in the presence of the correct NSAA, whereas C7.ΔA::S and C13.ΔA::S displayed read-through of all three UAG codons even in the absence of NSAAs, suggesting efficient incorporation of natural amino acids at native UAGs (17). Mass spectrometry indicated that C13.ΔA::S incorporated Gln, Lys, and Tyr at UAG codons. DNA sequencing in C7.ΔA::S and C13.ΔA::S revealed UAG suppressor mutations in *glnV*, providing direct genetic evidence of Gln suppression observed by Western blot (Fig. 3A) and mass spectrometry (table S13). C0.B*.ΔA::S displayed truncated GFP variants corresponding with UAG termination in the absence of RF1 (17) (Fig. 3A).

We directly investigated the impact of pAcF and Sep incorporation on the proteomes (Fig. 3B) (20) of our panel of strains (Table 1) using mass spectrometry (tables S6 to S12). No Sep-containing peptides were observed for EcNR2, illustrating that RF1 removal is necessary for NSAA incorporation by the episomal phosphoserine system (21), which is an inefficient orthogonal translation machinery (19) (Fig. 3C and table S10). By contrast, we observed NSAA-containing peptides in unrecoded (C0.B*.ΔA::S) and partially recoded (C13.ΔA::S) strains, and not the GRO (C321.ΔA::S), which lacks UAGs in its genome (Fig. 3, B and C, fig. S7, and tables S6 to S12). Such undesired incorporation of NSAAs (or natural amino acids) likely underlies the fitness impairments observed for C0.B*.ΔA::S, C7.ΔA::S, and C13.ΔA::S. In contrast to the other $RF1^-$strains, C321.ΔA::S demonstrated equivalent fitness to its $RF1^+$ precursor (Fig. 2) and efficiently expressed all GFP variants without incorporating NSAAs at unintended sites (Figs. 2 and 3 and fig. S6). Therefore, complete UAG removal is the only strategy that provides a devoted codon for plug-and-play NSAA incorporation without impairing fitness (Figs. 2 and 3).

To determine whether this GRO can obstruct viral infection, we challenged $RF1^-$ strains with bacteriophages T4 and T7. Viruses rely on their host to express proteins necessary for propagation. Because hosts with altered genetic codes would mistranslate viral proteins (3), recoding may provide a general mechanism for resistance to all natural viruses. Given that UAG codons occur rarely and only at the end of genes, we did not expect UAG reassignment to result in broad phage resistance. Although the absence of RF1 did not appear to affect T4 (19 of 277 stop codons are UAG), it significantly enhanced resistance to T7 (6 of 60 stop codons are UAG) (Fig. 4).

$RF1^-$ hosts produced significantly smaller T7 plaques independent of host doubling time (Fig. 4A and fig. S8). The only exception was C0.B*.ΔA::S, which produced statistically equivalent plaque sizes regardless of whether RF1 was present (Fig. 4A and table S14). Consistent with the observation that the modified RF2 variant could weakly terminate UAG

[(17) and herein], our results suggest that C0.B*.ΔA::S terminates UAG codons well enough to support normal T7 infection.

Given that plaque area and phage fitness (doublings per hour) do not always correlate, we confirmed that T7 infection is inhibited in RF1^{-} hosts by comparing T7 fitness and lysis time in C321 versus C321.ΔA (Fig. 4B). Phage fitness (doublings per hour) is perhaps the most relevant measure for assessing phage resistance because it indicates how quickly a log-phase phage infection expands (22). We found that T7 fitness was significantly impaired in strains lacking RF1 ($P = 0.002$), and kinetic lysis curves (fig. S9) confirmed that lysis was significantly delayed in the absence of RF1 ($P < 0.0001$, Fig. 4B). Meanwhile, one-step growth curves (fig. S10) indicated that burst size (average number of phages produced per lysed cell) in RF1^{-} hosts was also reduced by 59% (±9%), and phage packaging was delayed by 30% (±2%) (table S15). We hypothesize that ribosome stalling at the gene 6 (T7 exonuclease) UAG explains the T7 fitness defect in RF1^{-} hosts, whereas T4 may not possess a UAG-terminating essential gene with a similar sensitivity (supplementary text). Abolishing the function of additional codons could block the translation of viral proteins and prevent infections entirely.

Using multiplex genome editing, we removed all instances of the UAG codon and reassigned its function in the genome of a living cell. The resulting GRO possesses a devoted UAG sense codon for robust NSAA incorporation that is suitable for industrial protein production. GROs also establish the basis for genetic isolation and virus resistance, and additional recoding will help fully realize these goals—additional triplets could be reassigned, unnatural nucleotides could be used to produce new codons (23), and individual triplet codons could be split into several unique quadruplets (8, 24) that each encode their own NSAA. In an accompanying study (25), we show that 12 additional triplet codons may be amenable to removal and eventual reassignment in *E. coli*. However, codon usage rules are not fully understood, and recoded genome designs are likely to contain unknown lethal elements. Thus, it will be necessary to sample vast genetic landscapes, efficiently assess phenotypes arising from individual changes and their combinations, and rapidly iterate designs to change the genetic code at the genome level.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We dedicate this paper to the memory of our friend, colleague, and gifted scientist, Tara Gianoulis. We thank R. Kolter for JC411, D. Reich for help with sequencing libraries, and C. and J. Seidman for Covaris E210; J. Aach, S. Kosuri, and U. Laserson for bioinformatics; T. Young, F. Peters, and W. Barnes for NSAA incorporation advice; I. Molineux and S. Kosuri for phage advice; S. Vassallo and P. Mali for experimental support; and D. Söll, A. Forster, T. Wu, K. Oye, C. Gregg, M. Napolitano, U. Laserson, A. Briggs, D. Mandell, and R. Chari for helpful comments. Funding was from the U.S. Department of Energy (DE-FG02-02ER63445), NSF (SA5283-11210), NIH (NIDDK-K01DK089006 to J.R.), Defense Advanced Research Projects Agency (N66001-12-C-4040, N66001-12-C-4020, N66001-12-C-4211), Arnold and Mabel Beckman Foundation (F.J.I.), U.S. Department of Defense National Defense Science and Engineering Graduate Fellowship (M.J.L.), NIH-MSTP-TG-T32GM07205 (A.D.H.), NSF graduate fellowships (H.H.W. and D.B.G.), NIH Director's Early Independence Award (1DP5OD009172-01 to H.H.W.), and the Assistant Secretary of Defense for Research and Engineering (Air Force Contract no. FA8721-05-

C-0002 to P.A.C.). Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the U.S. government.

References and Notes

1. Vetsigian K, Woese C, Goldenfeld N. Proc Natl Acad Sci USA. 2006; 103:10696–10701. [PubMed: 16818880]
2. Goeddel DV, et al. Proc Natl Acad Sci USA. 1979; 76:106–110. [PubMed: 85300]
3. Krakauer DC, Jansen VAA. J Theor Biol. 2002; 218:97–109. [PubMed: 12297073]
4. Schafer MG, et al. PLOS ONE. 2011; 6:e25736. [PubMed: 21998689]
5. Sturino JM, Klaenhammer TR. Nat Rev Microbiol. 2006; 4:395–404. [PubMed: 16715051]
6. Schmidt M, de Lorenzo V. FEBS Lett. 2012; 586:2199–2206. [PubMed: 22710182]
7. Liu CC, Schultz PG. Annu Rev Biochem. 2010; 79:413–444. [PubMed: 20307192]
8. Neumann H, Wang K, Davis L, Garcia-Alai M, Chin JW. Nature. 2010; 464:441–444. [PubMed: 20154731]
9. Young TS, Ahmad I, Yin JA, Schultz PG. J Mol Biol. 2010; 395:361–374. [PubMed: 19852970]
10. Eggertsson G, Söll D. Microbiol Rev. 1988; 52:354–374. [PubMed: 3054467]
11. Isaacs FJ, et al. Science. 2011; 333:348–353. [PubMed: 21764749]
12. Gibson DG, et al. Science. 2010; 329:52–56. [PubMed: 20488990]
13. Wang HH, et al. Nature. 2009; 460:894–898. [PubMed: 19633652]
14. Carr PA, et al. Nucleic Acids Res. 2012; 40:e132. [PubMed: 22638574]
15. Mukai T, et al. Nucleic Acids Res. 2010; 38:8188–8195. [PubMed: 20702426]
16. Johnson DBF, et al. Nat Chem Biol. 2011; 7:779–786. [PubMed: 21926996]
17. Ohtake K, et al. J Bacteriol. 2012; 194:2606–2613. [PubMed: 22427623]
18. O'Donoghue P, et al. FEBS Lett. 2012; 586:3931–3937. [PubMed: 23036644]
19. Heinemann IU, et al. FEBS Lett. 2012; 586:3716–3722. [PubMed: 22982858]
20. Ngo JT, Tirrell DA. Acc Chem Res. 2011; 44:677–685. [PubMed: 21815659]
21. Park H-S, et al. Science. 2011; 333:1151–1154. [PubMed: 21868676]
22. Heineman RH, Molineux IJ, Bull JJ. J Mol Evol. 2005; 61:181–191. [PubMed: 16096681]
23. Bain JD, Switzer C, Chamberlin R, Benner SA. Nature. 1992; 356:537–539. [PubMed: 1560827]
24. Anderson JC, et al. Proc Natl Acad Sci USA. 2004; 101:7566–7571. [PubMed: 15138302]
25. Lajoie MJ, et al. Science. 2013; 342:361–363. [PubMed: 24136967]

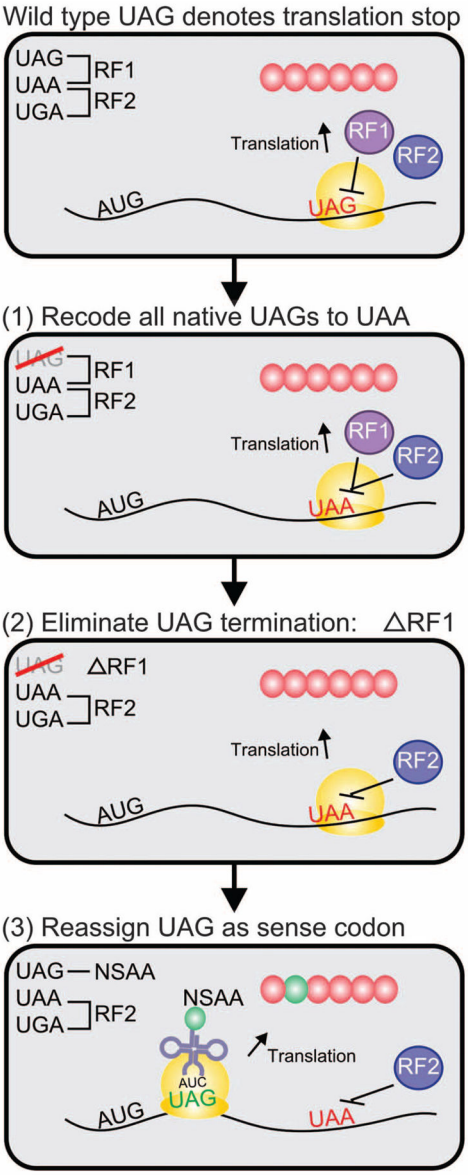


Fig. 1. Engineering a GRO with a reassigned UAG codon

Wild-type *E. coli* MG1655 has 321 known UAG codons that are decoded as translation stops by RF1 (for UAG and UAA). (1) Remove codons: converted all known UAG codons to UAA, relieving dependence on RF1 for termination. (2) Eliminate natural codon function: abolished UAG translational termination by deleting RF1, creating a blank codon. (3) Expand the genetic code: introduced an orthogonal aminoacyl–tRNA synthetase (aaRS) and tRNA to reassign UAG as a dedicated sense codon capable of incorporating nonstandard amino acids (NSAAs) with new chemical properties.

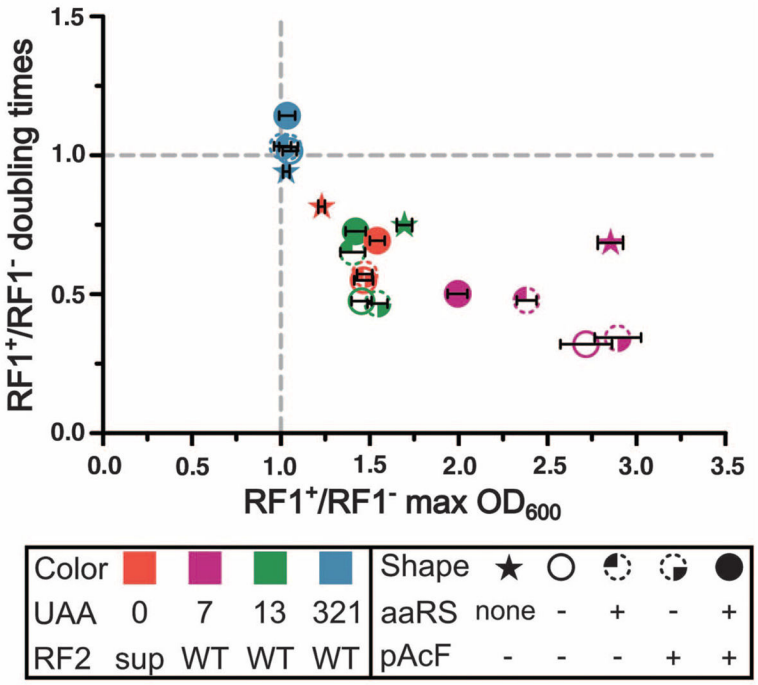


Fig. 2. Effects of UAG reassignment at natural UAG codons

Ratios of maximum cell densities (horizontal axis) and doubling times (vertical axis) were determined for $RF1^+$ strains versus their corresponding $RF1^-$ strains ($n = 3$) in the presence or absence of UAG suppression. Symbol color specifies genotype: UAA is the number of UAG→UAA mutations, and RF2 is “WT” (wild type) or “sup” [RF2 variant that can compensate for RF1 deletion (16)]. Symbol shape specifies NSAA expression: aaRS (aminoacyl–tRNA synthetase) is “none” (genes for UAG reassignment were absent), “−” [pEVOL-pAcF (9) is present but not induced, so only the constitutive aaRS and tRNA are expressed], or “+” (pEVOL-pAcF is fully induced using L-arabinose), and pAcF is “−” (excluded) or “+” (supplemented). Strains that do not rely on RF1 are expected to have a $RF1^+/RF1^-$ ratio at (1,1). $RF1^-$ strains exhibiting slower growth are below the horizontal gray line, and $RF1^-$ strains exhibiting lower maximum cell density are to the right of the vertical gray line. The doubling-time error bars are too small to visualize.

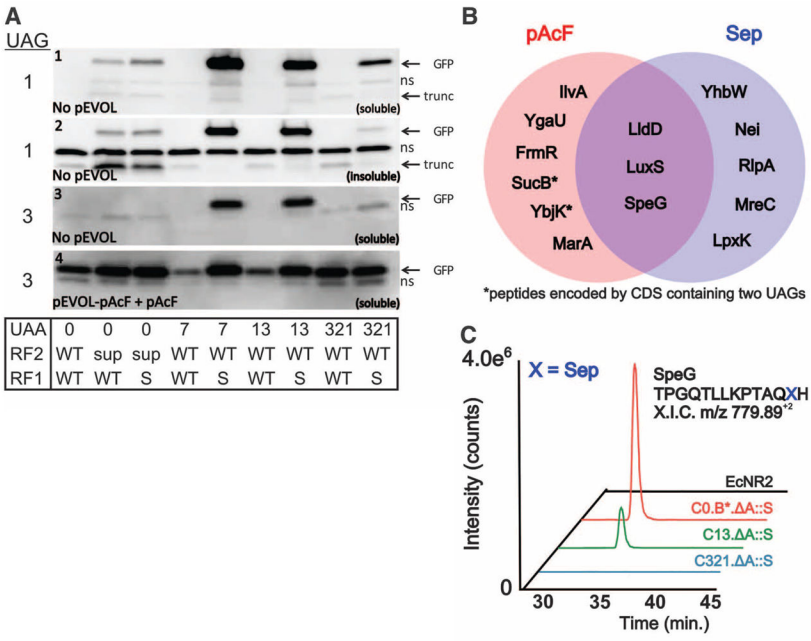


Fig. 3. NSAA incorporation in GROs

(**A**) Western blots demonstrate that C0.B*.ΔA::S terminates at UAG in the absence of RF1 and that C7.ΔA::S and C13.ΔA::S have acquired natural suppressors that allow strong NSAA-independent read-through of three UAG codons. When pAcF was omitted, one UAG reduced the production of full-length GFP, and three UAGs reduced production to undetectable levels for all strains except C7.ΔA::S and C13.ΔA::S, demonstrating that undesired near-cognate suppression (18) is weak for most strains even when RF1 is inactivated. However, all strains show efficient translation through three UAG codons when pAcF is incorporated. Western blots were probed with an antibody to GFP that recognizes an N-terminal epitope. UAA is the number of UAG→UAA mutations; RF2 is “WT” (wild type) or “sup” [RF2 variant that can compensate for RF1 deletion (16)]; RF1 is “WT” (wild type) or “S” (Δ*prfA::spec*^{R}). “GFP” is full-length GFP; “trunc” is truncated GFP from UAG termination and is enriched in the insoluble fraction; “ns” indicates a nonspecific band. (**B**) Venn diagram representing NSAA-containing peptides detected by mass spectrometry in C0.B*.ΔA::S when UAG was reassigned to incorporate *p*-acetylphenylalanine (pAcF, red) or phosphoserine (Sep, blue). No NSAA-containing peptides were identified in C321.ΔA::S. Asterisk (*) indicates coding DNA sequence possessing two tandem UAG codons. (**C**) Extracted ion chromatograms are shown for UAG suppression of the SpeG peptide to investigate Sep incorporation in natural proteins. Peptides containing Sep were only observed in C0.B*.ΔA::S, C7.ΔA::S, and C13.ΔA::S, as Sep incorporation was below the detection limit in EcNR2 (RF1^{+}), and *speG* was recoded in C321.ΔA::S.

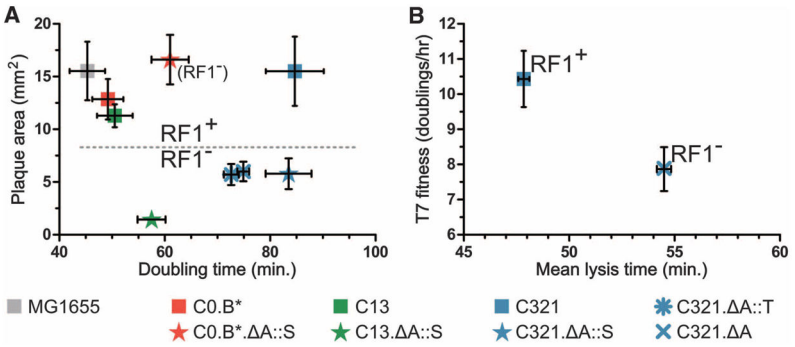


Fig. 4. Bacteriophage T7 infection is attenuated in GROs lacking RF1

RF1 (*prfA*) status is denoted by symbol shape: (■) wt *prfA* (WT); (★) Δ*prfA::spec*^{R} (ΔA::S); (✳) Δ*prfA::tolC* (ΔA::T); and (×) a clean deletion of *prfA* (ΔA). (**A**) RF1 status affects plaque area (Kruskal-Wallis one-way analysis of variance, $P < 0.001$), but strain doubling time does not (Pearson correlation, $P = 0.49$). Plaque areas (mm^2) were calculated with ImageJ, and means ± 95% confidence intervals are reported ($n > 12$ for each strain). In the absence of RF1, all strains except C0.B*.ΔA::S yielded significantly smaller plaques, indicating that the RF2 variant (16) can terminate UAG adequately to maintain T7 fitness. A statistical summary can be found in table S14. (**B**) T7 fitness (doublings/hour) (22) is impaired ($P = 0.002$) and mean lysis time (min) is increased ($P < 0.0001$) in C321.ΔA compared to C321. Significance was assessed for each metric by using an unpaired *t* test with Welch's correction.

Table 1

Recoded strains and their genotypes.

Strain*	Essential codons changed†	Total codons changed‡	Previously essential codon functions manipulated§	Expected (obs.) UAG translation function‖
EcNR2	0	0	None	Stop
C0.B¶	0	0	*prfB*#	Stop
C0.B¶.ΔA::S	0	0	*prfB*,# Δ*prfA::spec*^{R}	None (stop¶)
C7	7	7	None	Stop
C7.ΔA::S	7	7	Δ*prfA::spec*^{R}	None (sup)
C13	7	13	None	Stop
C13.ΔA::S	7	13	Δ*prfA::spec*^{R}	None (sup)
C321	7	321	None	Stop
C321.ΔA::S	7	321	Δ*prfA::spec*^{R}	None (nc)
C321.ΔA::T	7	321	Δ*prfA::tolC*	None (nc)
C321.ΔA	7	321	Δ*prfA*	None (nc)

* All strains are based on EcNR2 {*E. coli* MG1655 Δ(*ybhB-bioAB*)::[λcI857 N(*cro-ea59*)::*tetR-bla*] Δ*mutS::cat*}, which is mismatch repair deficient (Δ*mutS*) to achieve high-frequency allelic replacement; C0 and C321 strains are Δ*mutS::zeo*^{R}; C7 and C13 strains are Δ*mutS::tolC*; C7, C13, and C321 strains have the endogenous *tolC* deleted, making it available for use as a selectable marker. Spectinomycin resistance (S) or tolC (T) were used to delete *prfA* (A). Bacterial genetic nomenclature describing these strains includes:: (insertion) and Δ (deletion).

† Out of a total of 7.

‡ Out of a total of 321.

§ *prfA* encodes RF1, terminating UAG and UAA; *prfB* encodes RF2, terminating UGA and UAA; *prfB*# is an RF2 variant (T246A, A293E, and removed frameshift) exhibiting enhanced UAA termination (16) and weak UAG termination (17).

‖ Observed translation function: Stop, expected UAG termination; stop¶, weak UAG termination from RF2 variant; sup, strong selection for UAG suppressor mutations; nc, weak near-cognate suppression (i.e., reduced expression compared to C7.ΔA::S and C13.ΔA::S) in the absence of all other UAG translation function.

nature COMMUNICATIONS

ARTICLE

Received 27 Jan 2015 | Accepted 22 Jul 2015 | Published 9 Sep 2015

DOI: 10.1038/ncomms9130 OPEN

A flexible codon in genomically recoded *Escherichia coli* permits programmable protein phosphorylation

Natasha L. Pirman[1,3,*], Karl W. Barber[1,3,*], Hans R. Aerni[1,3], Natalie J. Ma[2,3], Adrian D. Haimovich[2,3], Svetlana Rogulina[1,3], Farren J. Isaacs[2,3] & Jesse Rinehart[1,3]

Biochemical investigation of protein phosphorylation events is limited by inefficient production of the phosphorylated and non-phosphorylated forms of full-length proteins. Here using a genomically recoded strain of *E. coli* with a flexible UAG codon we produce site-specific serine- or phosphoserine-containing proteins, with purities approaching 90%, from a single recombinant DNA. Specifically, we synthesize human MEK1 kinase with two serines or two phosphoserines, from one DNA template, and demonstrate programmable kinase activity. Programmable protein phosphorylation is poised to help reveal the structural and functional information encoded in the phosphoproteome.

[1] Department of Cellular & Molecular Physiology, Yale University, New Haven, Connecticut 06520-8114, USA. [2] Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520-8114, USA. [3] Systems Biology Institute, Yale University, New Haven, Connecticut 06520-8114, USA. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.R. (email: jesse.rinehart@yale.edu).

Protein phosphorylation modulates most cellular functions and is controlled by networks of kinases and phosphatases that add or remove phosphate groups at precisely defined positions. Large-scale phosphoproteomics efforts have mapped phosphorylation sites on the majority of human proteins[1,2]. Our knowledge of the phosphoproteome has outpaced our understanding of which protein kinases phosphorylate these important sites. Nature has evolved elaborate mechanisms to ensure proper kinase/substrate pairings[3]. Recombinant phosphoprotein synthesis is impeded by the fact that most of these pairs are unknown. We recently described a solution to this problem using *Escherichia coli* strains engineered to incorporate genetically encoded phosphoserine (Sep) into recombinant proteins by employing an orthogonal translation system (SepOTS)[4]. The SepOTS uses a Sep aminoacyl-tRNA synthetase (SepRS) to charge Sep onto a UAG-decoding $tRNA^{Sep}$ and an engineered elongation factor Tu (EF-Sep) that delivers Sep–$tRNA^{Sep}$ to the ribosome[4]. The SepOTS can direct phosphoserine incorporation into physiologically relevant positions within proteins without any knowledge of upstream kinases and thus provides a simplified platform for phosphoprotein synthesis.

The first described iteration of the SepOTS consisted of two plasmids housing genes for SepRS, EF-Sep and one gene copy of $tRNA^{Sep}$ (here referred to as SepOTSα)[4]. SepOTSα was functional in a standard BL21 *E. coli* strain, but inefficiently encoded Sep at UAG amber codons due to competition with release factor one (RF1). This barrier has long been recognized in the field of non-standard amino-acid (NSAA) incorporation and has compromised the yield and purity of NSAA-containing proteins. We recently demonstrated that recoding[5] the native TAG sites terminating seven essential genes in the *E. coli* genome to TAA enabled RF1 deletion and enhanced Sep incorporation[6]. This partially recoded cell (EcAR7.ΔA) enhanced UAG read-through, but exhibited severe growth impairment and suppression with natural amino acids. The deletion of RF1 causes ribosome stalling on the non-recoded, UAG-containing mRNAs, and protein instability from proteome-wide UAG read-through is elevated in the presence of the SepOTS[6,7] (Fig. 1a). We found that Sep incorporation at UAG sites was enhanced in EcAR7.ΔA, but we quantified natural amino-acid incorporation accounting for as much as 60% of UAG decoding[7], indicating that further advances in strain and SepOTS engineering were required. We recently introduced the first genomically recoded organism (C321.ΔA) where all 321 TAG sites in the genome were reassigned to TAA[8]. This strain not only tolerated the deletion of RF1 but also eliminated all RF1-knockout-associated growth defects observed in partially recoded strains. Furthermore, C321.ΔA exhibited muted UAG read-through by natural aminoacyl-tRNAs that compete with NSAA incorporation via OTSs[8] (Fig. 1a). The properties of C321.ΔA suggested that this strain would provide the ideal setting for Sep incorporation, offering a completely open and assignable UAG codon in a background that does not suffer from the low purity and compromised fitness observed in other RF1-knockout strains[6,7].

Here we demonstrate the utility of genomically recoded *E. coli* to produce phosphoproteins via site-specific incorporation of Sep with an enhanced version of the SepOTS. We take advantage of the flexible UAG codon to genetically program recombinant human MEK1 kinase activity, and we characterize our platform based on phosphoprotein yield, purity and positional bias. Overall, we conclude that our expression system enables robust expression of diverse phosphoproteins for potential biochemical and phosphoproteomic applications.

Results

Recoding *E. coli* permits increased phosphoprotein purity. We aimed to leverage the enhanced properties of C321.ΔA to further improve the SepOTS and to generate recombinant phosphoproteins with higher yield and purity. To evaluate SepOTS variants, we utilized an established green fluorescent protein (GFP) reporter in which position 17 is permissive to a glutamate-to-Sep substitution[6]. We recently used this reporter and quantitative mass spectrometry to show that using five $tRNA^{Sep}$ gene copies in our SepOTS plasmid (here within referred to as SepOTSμ) increases the yield and purity of phosphoproteins produced in EcAR7.ΔA[7]. Our analysis showed that Sep incorporation reached only 40% and UAG read-through by native aminoacyl-tRNAs accounted for 60% of the UAG decoding[7]. To efficiently compare and track this phenomenon across different strains, we adopted a Phos-tag gel shift assay[9] in which phosphoserine-containing proteins are separated from non-phosphorylated proteins and can be quantified on the same western blot[10–12]. The gel shift assays showed that increasing $tRNA^{Sep}$ gene copy number from one to five raised Sep incorporation at one UAG in EcAR7.ΔA from 40 to 60% and was generally in good agreement with our previous studies[7] (Fig. 1b, Supplementary Fig. 1). As a first step, we moved these two SepOTS variants into C321.ΔA. SepOTSα showed increased Sep incorporation to over 80% in C321.ΔA compared with only 40% in EcAR7.ΔA (Fig. 1b). This was the first clear indication that the C321.ΔA background significantly enhanced the ability of SepOTSα to encode Sep at higher purity. Interestingly, there was little difference in phospho-GFP purity using the SepOTSμ with 5 $tRNA^{Sep}$ gene copies and suggested that simple tRNA gene copy number manipulation would not be sufficient to further enhance SepOTS performance in C321.ΔA.

Modifying SepOTS to enhance phosphoserine incorporation. The properties of C321.ΔA, in particular the increased fitness and reduction of strong UAG read-through observed in partially recoded strains[7,8] that directly competes with Sep insertion, inspired us to perform a more extensive assessment of SepOTS variants in the C321.ΔA background. To accomplish this, we assembled combinations of known and novel SepOTS components and compared the total phospho-GFP synthesis and purity in C321.ΔA. We used the enhanced SepRS9 and EF-Sep21 variants and multiple gene copies of $tRNA^{Sep}$ known separately to increase SepOTS performance[7,13]. We also introduced a new $tRNA^{Sep}$ variant containing a G37A substitution ($tRNA^{Sep\text{-}A37}$) known to improve UAG read-through in other natural suppressor tRNAs[14]. We hypothesized that this $tRNA^{Sep}$ mutation would enhance Sep decoding by stabilizing anticodon base stacking in the ribosome[15]. We compared relative phospho-GFP expression for 12 different SepOTS combinations (renamed SepOTSβ–SepOTSν for simplicity) to the original two-plasmid system SepOTSα (ref. 4) (Fig. 1c). The most dramatic increase was observed using SepOTSλ, which contained enhanced SepRS9, EF-Sep21 and four gene copies of $tRNA^{Sep\text{-}A37}$ and yielded approximately ninefold elevation in GFP expression compared with SepOTSα. However, the variants containing SepRS9 and EF-Sep21 did not always outperform the original SepOTSα in the combinations tested. We did not observe a consistent correlation between increased tRNA gene copy number and increased phospho-GFP production, but observed that four gene copies of $tRNA^{Sep\text{-}A37}$ performed better than five gene copies of $tRNA^{Sep}$. This suggests that the G37A mutation increased the performance of the $tRNA^{Sep}$ synergistically with a tRNA gene copy number effect. Northern blot analysis of these cells confirmed that $tRNA^{Sep}$ levels generally correlate with gene copy number for all SepOTS

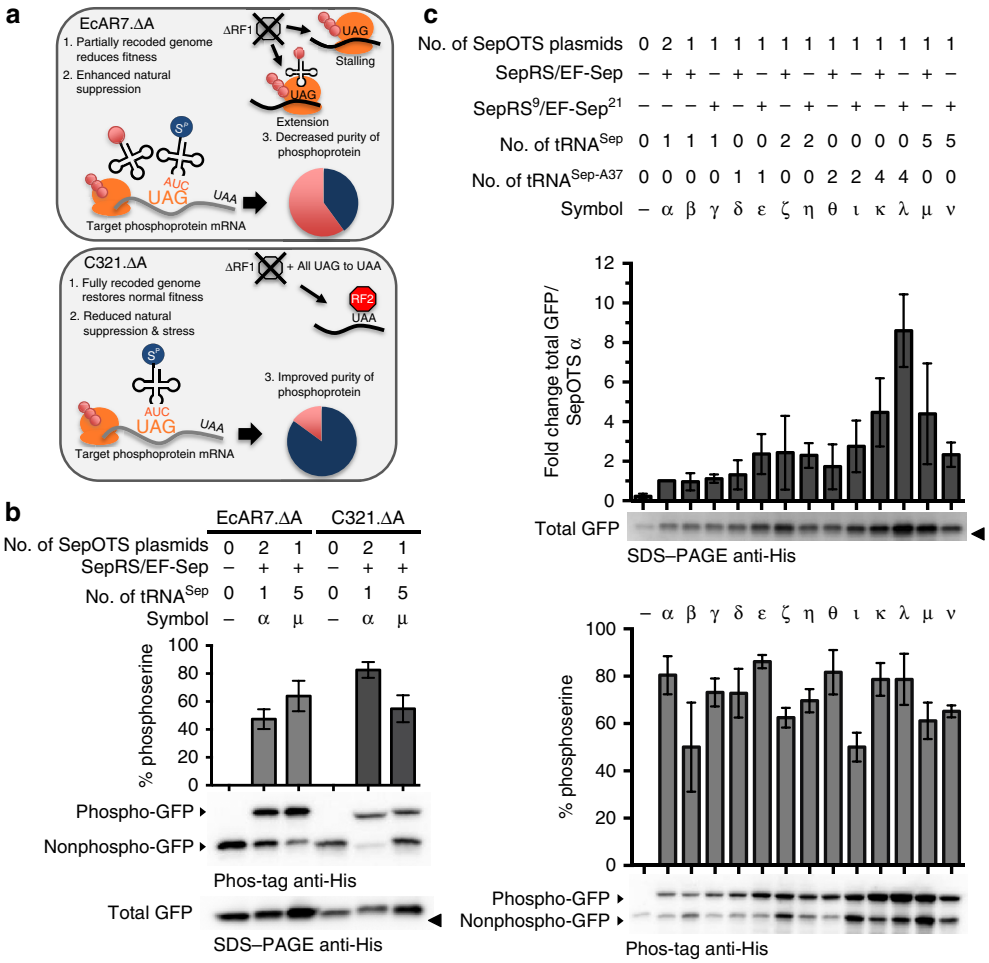


Figure 1 | Development of SepOTS variants for improved phosphoprotein yields and purity in C321.ΔA. (**a**) Schematic of Sep incorporation into a target recombinant phosphoprotein in the setting of the partially recoded EcAR7.ΔA and the fully recoded C321.ΔA strains. The 314 TAG-containing non-recoded loci in EcAR7.ΔA lead to ribosome stalling and increased near-cognate suppression. Sep incorporation at UAG codons is reduced due to competition for the target UAG by near-cognate suppression. The fully recoded C321.ΔA cell restores natural protein synthesis by restoring release factor function at all 321 recoded loci. Improved cellular fitness and reduced near-cognate suppression significantly improve the purity of recombinant phosphoproteins. (**b**) Phospho-GFP expression using two SepOTS variants (two-plasmid system including SepRS, EF-Sep and 1× $tRNA^{Sep}$ denoted SepOTSα and one-plasmid system including SepRS, EF-Sep and 5× $tRNA^{Sep}$ denoted SepOTSμ) in the EcAR7.ΔA versus C321.ΔA strains. A UAG codon at position 17 directs Sep incorporation into GFP. Purity was determined by Phos-tag gel shift and western blot and comparing intensities of the upper band (phospho-GFP) and the lower band (non-phospho-GFP). Error bars report s.d. from six biological replicates. Normalized loading was approximated by regular SDS–PAGE analysis of total GFP expression. Extended data is shown in Supplementary Fig. 1. (**c**) Phospho-GFP (Sep at position 17) expression using 13 SepOTS variants in the C321.ΔA strain. Fold change reflects GFP expression for each SepOTS variant compared with SepOTSα. Loading was normalized by OD_{600} measurement. Error bars report s.d. of three biological replicates. Purity of phospho-GFP was determined by Phos-tag gel shift and western blot. Error bars report s.d. from three biological replicates. Blots in (**b**) and (**c**) were performed using antibody against C-terminal 6xHis tag in GFP. The solid triangles denote the 25-kDa molecular weight marker. Phos-tag gels are run without molecular weight standards.

variants (Supplementary Fig. 2). Although we cannot fully explain the fluctuations in the expression patterns, we speculate that perhaps tRNA gene copy number may alter tRNA base modifications and contribute to the patterns observed. When we examined the purity of phospho-GFP with gel shift we observed that most of the SepOTS variants produce consistently higher phosphoprotein purity in C321.ΔA than the variants tested in the partially recoded EcAR7.ΔA (Fig. 1b,c). We noted that the increased protein yield from twofold with SepOTSε to ninefold with SepOTSλ did not compromise the purity of Sep insertion at UAG since both OTS systems achieved purities between 80 and 90% in the C321.ΔA strain (Fig. 1c).

Flexible codon for incorporation of serine or phosphoserine. The flexibility of amber codon amino-acid assignment in C321.ΔA presents the possibility of synthesizing proteins encoding either Ser or Sep at desired positions within the recombinant protein (Supplementary Fig. 3a). Taking advantage of this synthetic flexibility, we employed a known amber suppressor tRNA SupD to incorporate Ser at UAG codons[16]. We envision this modularity having important practical applications for the incorporation of either Sep or Ser at amber codons in large synthetic gene libraries, obviating the need for separate DNA pools encoding Sep and Ser. To test amber codon flexibility in an enzyme, we used beta-lactamase (Bla) since its catalytic site requires Ser at position 68 to produce ampicillin-resistant cells[17]. A Bla variant replacing S68 with a TAG codon (S68TAG Bla) was then created. C321.ΔA cells transformed with SupD and S68TAG Bla plasmids were resistant to ampicillin and confirmed that the amber codon could be used to encode serine

(Supplementary Fig. 3b). C321.ΔA transformed with SepOTSλ and S68TAG Bla plasmids to encode Sep at the amber codon yielded expression of inactive Bla and no cell growth in the presence of ampicillin (Supplementary Fig. 3b). This experiment also validates the stability of Sep incorporation at position 68 in the Bla protein, since robust *in vivo* Sep dephosphorylation to Ser would have produced an active Bla enzyme. These results show that the amber codon in C321.ΔA can be used to create an active and inactive enzyme from the same gene. Furthermore, this demonstrates that the phosphorylation state and function of a protein can be programmed by simply employing different translation machinery for the amber codon.

Programmable active kinase production. We next wanted to use our enhanced SepOTS systems and the flexible UAG to control phosphoserine-related function in a physiologically relevant mammalian protein containing more than one phosphoserine. Mitogen-activated protein kinase kinase 1 (MEK1) directly controls a broad range of cell cycle functions, is implicated in oncogenesis[18], and like many kinases[19–22] is widely believed to require phosphorylation at two sites for activity. To measure the robustness of our experimental systems we synthesized MEK1 with genetically encoded Sep at positions 218 and 222 (MEK1-S^PS^P) in its activation segment. We first used SepOTSμ (refs 7,23) to compare MEK1-S^PS^P expression levels in BL21 $RF1^+$, EcAR7.ΔA and C321.ΔA. MEK1-S^PS^P expression in BL21 $RF1^+$ was undetectable due to RF1 activity, whereas both EcAR7.ΔA and C321.ΔA RF1-deficient backgrounds easily produced detectible levels of MEK1-S^PS^P (Supplementary Fig. 4). C321.ΔA has a growth curve with a steep log phase and final OD_{600} closer to wild-type cells[8] and consistently outperforms the poor growth characteristics of EcAR7.ΔA[6] (Supplementary Fig. 5). This substantial boost in fitness, and the enhanced SepOTS performance in the C321.ΔA strain, rendered it the optimal strain for MEK1-S^PS^P expression.

To further investigate MEK1 synthesis we used the highest performing SepOTS variant, SepOTSλ, and found expression of MEK1-S^PS^P was twofold higher (~2 mg per l culture) than with SepOTSμ (~1 mg per l culture, Fig. 2a). In parallel, we used SupD in C321.ΔA with the same *MEK1* gene (TAGs at 218 and 222) to produce inactive MEK1 containing Ser at positions 218 and 222 (Fig. 2a). MEK1-S^PS^P production was confirmed with a commercially available phosphospecific antibody for positions 218 and 222 (Fig. 2a) and further characterized by examining singly phosphorylated forms of MEK1 (MEK1-S^PS and MEK1-SS^P) by Phos-tag gel shift (Supplementary Fig. 6a). These studies confirmed the synthesis of singly and doubly phosphorylated MEK1 but revealed that the MEK1 phospho-specific antibody could readily detect both MEK1-S^PS^P and MEK1-SS^P, but not MEK1-S^PS. We also analysed the MEK1-SS, SS^P and S^PS^P variants by mass spectrometry and verified Sep incorporation (Supplementary Fig. 6b and Supplementary Table 1). Despite our best efforts, we were unable to directly observe the doubly phosphorylated peptide likely due to poor ionization. To overcome this problem, we treated MEK1-S^PS^P with calf intestinal alkaline phosphatase to dephosphorylate the protein. Calf intestinal alkaline phosphatase-treated MEK1-S^PS^P produced a peptide containing two serine residues at the S218 and S222 positions (absent in untreated MEK1-S^PS^P sample), confirming that Sep was originally inserted at both positions within the protein. We next used our established data analysis techniques to identify natural amino acids incorporated at the two different UAG sites in the MEK1 activation loop[7] (Supplementary Table 1). These non-phosphorylated MEK1 peptides explain a portion of the unshifted bands detected in our Phos-tag analysis and, consistent with our previous studies[7,8], demonstrate that near-cognate suppression of the amber codon can lead to natural amino-acid incorporations that interfere with SepOTS activity and phosphoprotein purity.

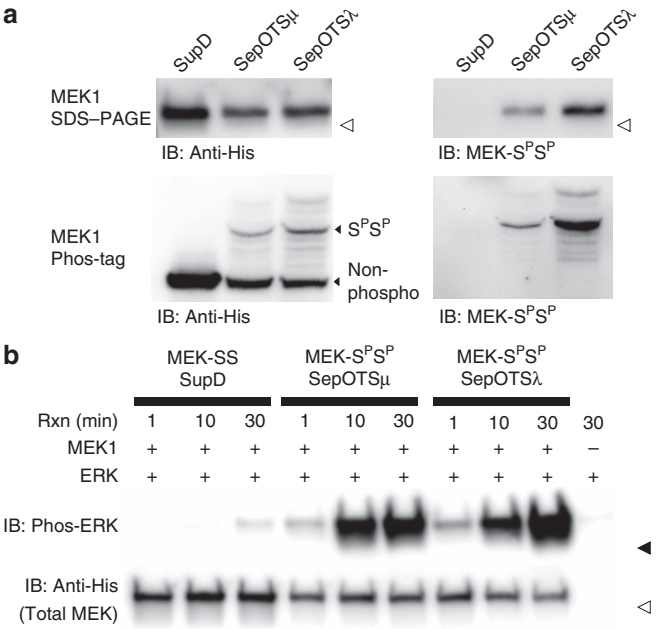


Figure 2 | A single-gene system for robust production of phosphorylated and non-phosphorylated human kinases. (**a**) Analysis of MBP-MEK1 S218TAG/S222TAG expression in the fully recoded C321.ΔA strain using plasmids that enable incorporation of Ser (one-plasmid system encoding SepRS, EF-Sep and 2 × $tRNA^{SupD}$ denoted SupD) or Sep (one-plasmid system encoding SepRS, EF-Sep and either 5 × $tRNA^{Sep}$ denoted SepOTSμ or 4 × $tRNA^{Sep\text{-}A37}$ denoted SepOTSλ) at amber codons. Total expression was evaluated by western blot using an antibody against the C-terminal 6xHis tag on MBP-MEK1. Loading was normalized by OD_{600} measurement. Presence of phosphorylated MBP-MEK1 was verified using a phosphospecific antibody (extended data shown in Supplementary Fig. 6a). Phosphoprotein purity was determined by Phos-tag gel shift and western blot by comparing the relative intensity of the upper bands compared with the lower band, corresponding to the phosphorylated and non-phosphorylated forms of MBP-MEK1, respectively. (**b**) Assay of MBP-MEK1 kinase activity monitored by phosphorylation of a kinase-dead ERK2 substrate *in vitro*. Representative blots showing phosphorylated (anti-Phos-ERK) and total MEK (anti-His) are shown. Assays and westerns were performed in triplicate and representative results are shown. The open and solid triangles denote 75- and 37-kDa molecular weight markers, respectively. Phos-tag gels are run without molecular weight standards.

A side-by-side comparison of MEK1 production showed that SepOTSμ and SepOTSλ yielded ~12% and ~30% phosphorylated MEK1-S^PS^P, respectively (Fig. 2a). While SepOTSλ increased the yield of MEK1-S^PS^P as predicted by the GFP reporter studies, the yield of phosphorylated MEK1 was lower than phospho-GFP. This could be due to instability of phosphorylated MEK or context-specific effects of UAG suppression in the MEK activation segment. To examine protein stability we conducted time course expression studies that showed higher MEK1-S^PS^P phosphoprotein yield with shorter expression times (Supplementary Fig. 6c). We assessed positional effects of Sep incorporation or stability using several singly and doubly phosphorylated forms of GFP (Supplementary Fig. 6d). The differences in phosphoserine incorporation between the GFP variants led us to conclude that UAG context and

phosphoprotein stability are both contributing factors to the overall purity of recombinant phosphoproteins.

To demonstrate the utility of the flexible UAG codon in synthesizing active mammalian kinases, we carried out MEK1 *in vitro* phosphorylation assays. MEK1-S^PS^P produced robust *in vitro* ERK2 phosphorylation while non-phosphorylated MEK1 (produced with SupD in C321.ΔA) was completely inactive (Fig. 2b). This demonstrates that the activity of a protein kinase can be programmed by simply using different translational machinery for the UAG codon in C321.ΔA cells. We envision that this flexibility could be leveraged in experiments with large panels or arrays of recombinant enzymes designed to explore the phosphoproteome at greater depth.

Discussion

These results clearly demonstrate that the physiological properties of protein phosphorylation can be controlled by a programmable genetic code. The flexible UAG codon in C321.ΔA can be manipulated with different OTS systems to produce proteins encoding either Sep or Ser at amber codons. We show that C321.ΔA is an optimal host for phosphoprotein expression with SepOTSε and SepOTSλ spanning a ninefold expression range without compromising purity. These systems provide a platform to reveal the functional consequences of serine phosphorylation. As a proof of principle, we reproduced the phosphorylation events underlying the MEK1 activity switch that controls a human signalling cascade. Our programmable UAG can be used to model other important kinase regulatory switches to enable applications such as kinase substrate discovery and screens for novel kinase inhibitors. More broadly, different pairs of natural suppressor tRNAs and OTSs could be introduced into C321.ΔA to explore more types of post-translational modification[24]. Certainly, exploring the function of protein phosphorylation in more enzymatic classes and across different species should now be possible. We recently showed that our C321.ΔA-based phosphoserine systems can be used for *in vitro* protein synthesis platforms, which may further extend the experimental landscape[25]. Guided by the large number of physiological protein phosphorylation sites[2], systematic investigations into the structure and function of any phosphoproteome could be within reach.

Methods

Transformations and strain storage. All *E. coli* strains used in this study were made chemically competent using a standard $RbCl_2$ method. EcAR7.ΔA was co-transformed with E17TAG GFP in the modified pCR-Blunt II-TOPO vector with the signified SepOTS version. C321.ΔA was co-transformed with PCRT7 NT Topo tetR pLtetO plasmid containing S2TAG, E17TAG, Q157TAG, S2TAG/E17TAG, E17TAG/Q157TAG GFP, maltose binding protein-MEK1 (MBP-MEK1) (S218TAG/S222), MBP-MEK1 (S218/S222TAG) or MBP-MEK1 (S218TAG/S222TAG) cassettes with the designated SepOTS. BL21 RF1+ was co-transformed with MBP-MEK1 and the signified SepOTS. All strains were stored as frozen glycerol stocks and restreaked on selective Luria–Bertani (LB) agar plates before protein expression. Strains harbouring the following plasmids were grown with the indicated antibiotic concentrations: E17TAG GFP in the modified pCR-Blunt II-TOPO plasmids, 25 μg per ml Zeocin (Zeo); all pCRT7 NT Topo tetR pLtetO plasmids, 100 μg per ml ampicillin (Amp); SupD and all SepOTS plasmids, 25 μg per ml kanamycin (Kan) with the exception of the two plasmid SepOTSα which requires 6 μg per ml tetracycline (Tet) and 25 μg per ml Kan; β-lactamase S68TAG plasmid, 10 μg per ml chloramphenicol (Cam) or 100 μg per ml Amp, as indicated. All cultures were started from a freshly streaked glycerol stock on LB agar plates with the appropriate combination of antibiotics and 0.08% glucose. Detailed information about strain and plasmid generation can be found in the Supplementary Methods. All primer sequences can be found in Supplementary Table 2. Sequences of tRNA-related synthesized genes can be found in Supplementary Table 3.

GFP variant protein expression. *E. coli* strains were streaked from −80 °C frozen glycerol stocks on LB agar selective plates and grown ~20 (BL21-based strains), ~24 (C321.ΔA-based strains), or ~48 h (EcAR7.ΔA-based strains) at 30 °C. Five colonies were inoculated in 5 ml of LB media supplemented with appropriate antibiotic and 0.08% glucose. Precultures were grown ~16 h at 30 °C shaking at 230 r.p.m. Cells were diluted to OD_{600} of 0.15 a.u. in 20 ml of LB with antibiotics, 0.08% glucose, 2 mM *O*-phospho-L-serine (Sep), and protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside and 100 ng per ml anhydrotetracycline. Protein was expressed for 20 h at 30 °C shaking at 230 r.p.m. unless otherwise noted. An equivalent number of cells as 1 ml of OD_{600} 2.5 a.u. was collected and spun down at 4,000*g* for 5 min at 4 °C. Media was aspirated and dry cell pellets were stored at −80 °C. Transformation and retention of the correct SepOTS variant (including tRNA gene copy number) were confirmed by PCR from frozen cell pellets. Cells were resuspended in 40 μl bacterial lysis buffer (50 mM Tris/HCl (pH 7.4), 150 mM NaCl, 1 mM dithiothreitol (DTT), 50 mM NaF, 1 mM $NaVO_4$ and 5% glycerol) supplemented with cOmplete mini-EDTA-Free protease inhibitor cocktail tablets (Roche) and 1× BugBuster detergent (Novagen), and kept on ice for 10 min. Samples were then spun down at 21,000*g* for 7 min at 4 °C and the supernatant (lysate) was transferred to a new tube and stored at −80 °C.

For western analysis, equal volumes of each sample were diluted in 1× Laemmli buffer (Bio-Rad). Samples were run on a 15-well 4–15% acrylamide gels (Bio-Rad) or on a handcast 15-well 12% acrylamide gel containing 100 μM Phos-tag Acrylamide (Waco Pure Chemical Industries, Inc., AAL-107). Transferred polyvinylidene difluoride (PVDF) membranes were blotted with 1:2,500 Anti-His (6xHis Epitope TAG, PA1-983B, Thermo Fisher Scientific), followed by 1:10,000 DAR–HRP (Peroxidase-conjugated AffiniPure Donkey Anti-Rabbit IgG, 711-035-152, Jackson ImmunoResearch). Signal was detected by enhanced chemiluminescence (Bio-Rad) imaged on a ChemiDoc XRS + CCD camera (Bio-Rad). Densitometry was performed using the Bio-Rad Image Lab software. Three biological replicates (starting from glycerol stocks) were performed for both the total expression and phospho-shifted westerns. Uncropped western blot images corresponding to text and Supplementary Figures are included in Supplementary Figs 7–12.

MBP-MEK1 recombinant protein expression and purification. The 5-ml precultures were inoculated with 5–20 colonies and grown overnight to confluency in LB media containing 0.08% glucose and antibiotics. The precultures were diluted to OD_{600} 0.15 a.u. in 100 ml LB media containing 0.08% glucose, antibiotics and 2 mM Sep, and were incubated at 30 °C, 230 r.p.m. for ~3 h to an OD_{600} of 0.8 a.u. Protein expression was then induced with 100 ng per ml anhydrotetracycline and 1 mM isopropyl β-D-1-thiogalactopyranoside, and cultures were grown at 20 °C, 230 r.p.m. for ~20–22 h unless otherwise noted. An equivalent number of cells as 1 ml of OD_{600} 2.5 a.u. was collected and spun down at 4,000*g* for 5 min at 4 °C. The remaining cells were collected at 4,000*g*, 20 min, 4 °C. Pellets were resuspended in ~30 ml of the used LB media and transferred to 50-ml centrifuge tubes and centrifuged under the same conditions again. All media was decanted, and pellets were stored at −80 °C.

The 1-ml OD_{600} 2.5 a.u. cell pellets were resuspended in 40 μl bacterial lysis buffer (with protease inhibitors and detergent, as for E17TAG GFP expression above) and kept on ice for 10 min. Samples were then spun down at 21,000*g* for 7 min at 4 °C and the supernatant (lysate) was transferred to a new tube and stored at −80 °C.

For western analysis, equal volumes of OD_{600} normalized lysates were diluted in 1× Laemmli buffer (Bio-Rad). Samples were run on a 15-well 4–15% acrylamide gels (Bio-Rad) or on a handcast 10-well 7.5% acrylamide gel containing 25 μM Phos-tag Acrylamide. Transferred PVDF membranes were blotted with either 1:2,500 anti-His (6xHis Epitope TAG, PA1-983B, Thermo Fisher Scientific) for total MEK1 or 1:1,000 Anti-MEK-S^PS^P (Phospho-MEK1/2 (Ser217/221), 9154, Cell Signaling Technology), followed by 1:10,000 DAR–HRP. Signal was detected by enhanced chemiluminescence (Bio-Rad) imaged on a ChemiDoc XRS + CCD camera. Densitometry was performed using the Bio-Rad Image Lab software.

Samples for MBP-MEK1 (S218TAG/S222TAG) in C321.ΔA harbouring either the SepOTSλ or SupD plasmids expression time course were expressed as described above, with exception that after induction 1-ml aliquots were removed at denoted time points and spun down at 4,000*g* for 5 min at 4 °C. All media was decanted, and pellets were stored at −80 °C. For western analysis samples were lysed in 80 μl of 2× Laemmli buffer and ran on a handcast 15-well 7.5% acrylamide gel containing 25 μM Phos-tag Acrylamide. Membranes were blotted with Anti-His as described above.

Cell pellets for purification were thawed in a 37 °C water bath for ~20 s and then put on ice. The pellets were resuspended in 5 ml of bacterial lysis buffer for sonication (50 mM Tris/HCl (pH 7.4), 500 mM NaCl, 0.5 mM EDTA, 0.5 mM EGTA, 1 mM DTT, 1 mg per ml lysozyme, 50 mM NaF, 1 mM NaVO4, 10% glycerol and complete mini-EDTA-Free protease inhibitor cocktail tablets) and incubated on ice for 30 min, followed by sonication. The lysates were centrifuged at 22,000*g*, 15 min, 4 °C and the clarified lysate was transferred to a 15-ml tube, and centrifuged under the same conditions again to remove all remaining insoluble material. The 200 μl bed volumes of Ni-NTA agarose resin (Qiagen Valencia, CA) were transferred to Pierce spin columns (Thermo Scientific Waltham, MA) and resin was equilibrated with 5 ml of Ni-NTA equilibration buffer (50 mM Tris/HCl (pH 7.4), 500 mM NaCl, 0.5 mM EDTA, 0.5 mM EGTA, 1 mM DTT, 50 mM NaF, 1 mM $NaVO_4$ and 10% glycerol). The clarified lysate was loaded onto the column via syringe then the column was washed with 10 ml of Ni-NTA wash buffer

(50 mM Tris/HCl (pH 7.4), 500 mM NaCl, 0.5 mM EDTA, 0.5 mM EGTA, 1 mM DTT, 50 mM NaF, 1 mM $NaVO_4$, 10% glycerol and 20 mM imidazole). Protein was eluted using Ni-NTA elution buffer (50 mM Tris/HCl (pH 7.4), 500 mM NaCl, 0.5 mM EDTA, 0.5 mM EGTA, 1 mM DTT, 50 mM NaF, 1 mM $NaVO_4$, 10% glycerol and 250 mM imidazole) and 400 µl elutions were collected in a clean Eppendorf tubes. Each fraction was assessed by SDS–polyacrylamide gel electrophoresis (SDS–PAGE).

The eluates were pooled, concentrated and buffer exchanged into the protein storage buffer (50 mM Tris/HCl (pH 7.4), 150 mM NaCl, 1 mM DTT and 20% glycerol) using a 0.5-ml Amicon Ultra centrifugal filter (Millipore Billerica, MA) and the protein was stored at −20 °C. The protein concentration was estimated by UV280 and by comparing known quantities of BSA standards on an SDS–PAGE gel.

K54R ERK2 was expressed and purified as described for MEK1 above.

MEK1-S^PS^P kinase activity assay. Kinase activity of purified MBP-MEK1 (S218/S222 or S^P218/S^P222) was evaluated by measuring ERK2 phosphorylation. In all, 1.0-µM MBP-MEK1 variants were pre-incubated in kinase activity buffer (50 mM Tris/HCl (pH 7.4), 150 mM NaCl, 1 mM DTT, 20% glycerol, 10 mM $MgCl_2$ and 1 mM ATP) at 30 °C for 5 min, then 2.5 µM ERK2 substrate was added to the reaction and further incubated at 30 °C. A 7.5 µl volume of the reaction was removed at 1, 10 and 30 min time points and quenched with 7.5 µl of 2 × Laemmli sample buffer (Bio-Rad), then heated to 55 °C for 5 min. A negative control was run with only ERK2 substrate for 30 min and quenched in the same manner as kinase-containing samples. The quenched reactions were run on 15-well 4–15% acrylamide SDS–PAGE gels and transferred to a PVDF membrane. Each membrane was cut between the 50–75-kDa protein markers and the bottom portion of the membrane was blotted with 1:1,000 anti-Phos-Erk antibody (Phospho p44/42 MAPK (Erk1/2) (Thr 202/Tyr204), 9101, Cell Signaling Technology) and the top portion of the membrane was blotted with 1:2,500 Anti-His antibody (6xHis Epitope TAG, PA1-983B, Thermo Fisher Scientific) for total MBP-MEK1, followed by 1:10,000 DAR–HRP. Signal was detected by enhanced chemiluminescence (Bio-rad) imaged on a ChemiDoc XRS + CCD camera. The kinase activity assay was run in triplicate using the same purified preparation of MBP-MEK1 kinase and K54R ERK2.

References

1. Olsen, J. V. & Mann, M. Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol. Cell. Proteomics.* **12,** 3444–3452 (2013).
2. Hornbeck, P. V. *et al.* PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* **40,** D261–D270 (2012).
3. Mok, J. *et al.* Deciphering protein kinase specificity through large-scale analysis of yeast phosphorylation site motifs. *Sci. Signal* **3,** ra12 (2010).
4. Park, H. S. *et al.* Expanding the genetic code of *Escherichia coli* with phosphoserine. *Science* **333,** 1151–1154 (2011).
5. Isaacs, F. J. *et al.* Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science* **333,** 348–353 (2011).
6. Heinemann, I. U. *et al.* Enhanced phosphoserine insertion during Escherichia coli protein synthesis via partial UAG codon reassignment and release factor 1 deletion. *FEBS Lett.* **586,** 3716–3722 (2012).
7. Aerni, H. R., Shifman, M. A., Rogulina, S., O'Donoghue, P. & Rinehart, J. Revealing the amino acid composition of proteins within an expanded genetic code. *Nucleic Acids Res.* **43,** e8 (2014).
8. Lajoie, M. J. *et al.* Genomically recoded organisms expand biological functions. *Science* **342,** 357–360 (2013).
9. Kinoshita, E., Kinoshita-Kikuta, E., Takiyama, K. & Koike, T. Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Mol. Cell. Proteomics* **5,** 749–757 (2006).
10. Takeya, K., Loutzenhiser, K., Shiraishi, M., Loutzenhiser, R. & Walsh, M. P. A highly sensitive technique to measure myosin regulatory light chain phosphorylation: the first quantification in renal arterioles. *Am. J. Physiol. Renal Physiol.* **294,** F1487–F1492 (2008).
11. Yamada, S. *et al.* Separation of a phosphorylated histidine protein using phosphate affinity polyacrylamide gel electrophoresis. *Anal. Biochem.* **360,** 160–162 (2007).
12. Kinoshita-Kikuta, E., Aoki, Y., Kinoshita, E. & Koike, T. Label-free kinase profiling using phosphate affinity polyacrylamide gel electrophoresis. *Mol. Cell. Proteomics* **6,** 356–366 (2007).
13. Lee, S. *et al.* A facile strategy for selective incorporation of phosphoserine into histones. *Angew. Chem. Int. Ed.* **52,** 5771–5775 (2013).
14. Eggertsson, G. & Soll, D. Transfer ribonucleic acid-mediated suppression of termination codons in Escherichia coli. *Microbiol. Rev.* **52,** 354–374 (1988).
15. Demeshkina, N., Jenner, L., Westhof, E., Yusupov, M. & Yusupova, G. A new understanding of the decoding principle on the ribosome. *Nature* **484,** 256–259 (2012).
16. Steege, D. A. A nucleotide change in the anticodon of an Escherichia coli serine transfer RNA results in supD-amber suppression. *Nucleic Acids Res.* **11,** 3823–3832 (1983).
17. Normanly, J., Ogden, R. C., Horvath, S. J. & Abelson, J. Changing the identity of a transfer RNA. *Nature* **321,** 213–219 (1986).
18. Scholl, F. A., Dumesic, P. A. & Khavari, P. A. Effects of active MEK1 expression in vivo. *Cancer Lett.* **230,** 1–5 (2005).
19. Zheng, C. F. & Guan, K. L. Activation of MEK family kinases requires phosphorylation of two conserved Ser/Thr residues. *EMBO J.* **13,** 1123–1131 (1994).
20. Alessi, D. R. *et al.* Mechanism of activation of protein kinase B by insulin and IGF-1. *EMBO J.* **15,** 6541–6551 (1996).
21. Kyriakis, J. M., Brautigan, D. L., Ingebritsen, T. S. & Avruch, J. pp54 microtubule-associated protein-2 kinase requires both tyrosine and serine/threonine phosphorylation for activity. *J. Biol. Chem.* **266,** 10043–10046 (1991).
22. Raingeaud, J. *et al.* Pro-inflammatory cytokines and environmental stress cause p38 mitogen-activated protein kinase activation by dual phosphorylation on tyrosine and threonine. *J. Biol. Chem.* **270,** 7420–7426 (1995).
23. Sawyer, N. *et al.* Designed phosphoprotein recognition in Escherichia coli. *ACS Chem. Biol.* **9,** 2502–2507 (2014).
24. Davis, L. & Chin, J. W. Designer proteins: applications of genetic code expansion in cell biology. *Nat. Rev. Mol. Cell Biol.* **13,** 168–182 (2012).
25. Oza, J. *et al.* Robust production of recombinant phosphoproteins using cell-free protein synthesis. *Nat. Commun* **6,** 8168 (2015).

Acknowledgements

This research was supported by National Institutes of Health (NIDDK-K01DK089006 and P01DK01743341to J.R. and NIH-MSTP-TG-T32GM07205 to A.D.H.); Defense Advanced Research Projects Agency (N66001-12-C-4211 to J.R. and F.J.I.); DuPont, Inc. (F.J.I.); Gen9, Inc. (F.J.I.); The Arnold and Mabel Beckman Foundation (F.J.I.); K.W.B. is supported by NIH-T32GM100884 and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1122492. We would also like to thank Noah Reynolds, Hubert Salvail and Eduardo Groisman for assistance with northern blot experiments.

Author contributions

N.L.P. and K.W.B. assisted in cloning, expression, performing experiments, analysing data, preparing figures and writing the manuscript; H.R.A performed the mass spectrometry experiments; N.J.M. and A.D.H. designed and constructed the modified C321.ΔA strain; S.R. assisted in cloning and expression; F.J.I. supervised the construction of the recoded strains, and provided advice on experimental design and manuscript content and J.R. supervised the research, assisted in analysing data, preparing the figures and writing the manuscript.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

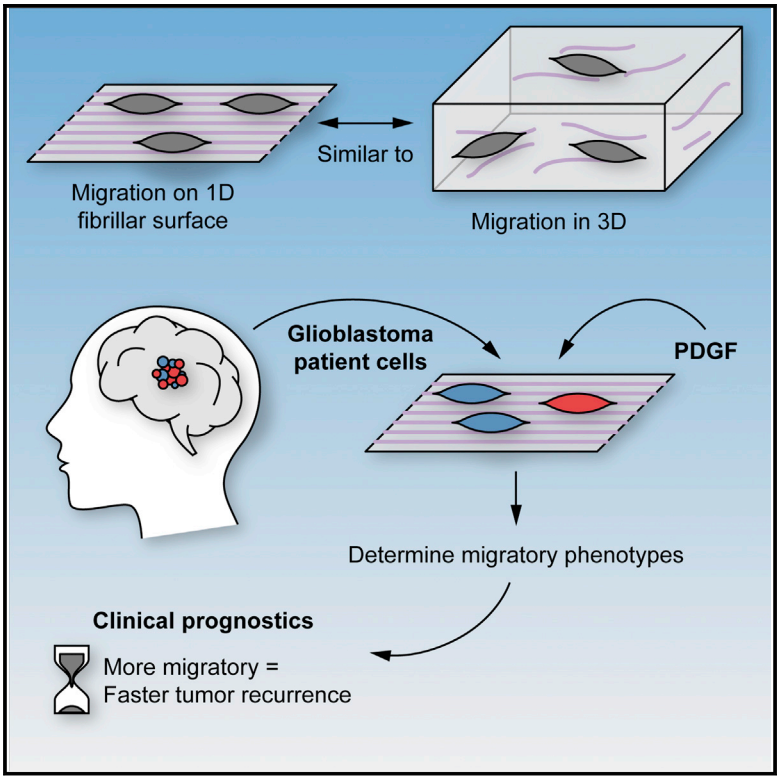
Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Pirman, N. L. *et al.* A flexible codon in genomically recoded *Escherichia coli* permits programmable protein phosphorylation. *Nat. Commun.* 6:8130 doi: 10.1038/ncomms9130 (2015).

Cell Reports

Migration Phenotype of Brain-Cancer Cells Predicts Patient Outcomes

Graphical Abstract



Authors

Chris L. Smith, Onur Kilic, Paula Schiapparelli, ..., JinSeok Park, Alfredo Quiñones-Hinojosa, Andre Levchenko

Correspondence

aquinon2@jhmi.edu (A.Q.-H.), andre.levchenko@yale.edu (A.L.)

In Brief

Smith et al. describe a method to quantitatively screen cell migration responses of glioma cells to the growth factor PDGF. They demonstrate that this technique can differentiate glioma cells into two groups—strong and weak responders—which in turn strongly correlate with patient relapse and tumor location.

Highlights

- High-throughput analysis for tumor single-cell migration
- More sensitive and physiologically relevant than classical screening assays
- Glioma cells showing inter- and intra-patient differential sensitivity to PDGF
- Glioma cell sensitivity to PDGF correlating with tumor recurrence and tumor location

Smith et al., 2016, Cell Reports *15*, 2616–2624
June 21, 2016
http://dx.doi.org/10.1016/j.celrep.2016.05.042

CellPress

OPEN ACCESS
CellPress

Cell Reports
Report

Migration Phenotype of Brain-Cancer Cells Predicts Patient Outcomes

Chris L. Smith,[1,2] Onur Kilic,[1] Paula Schiapparelli,[2] Hugo Guerrero-Cazares,[2] Deok-Ho Kim,[5] Neda I. Sedora-Roman,[6] Saksham Gupta,[2,3] Thomas O'Donnell,[2,3] Kaisorn L. Chaichana,[2] Fausto J. Rodriguez,[4] Sara Abbadi,[2] JinSeok Park,[7] Alfredo Quiñones-Hinojosa,[2,*] and Andre Levchenko[7,*]

[1]Department of Biomedical Engineering
[2]Department of Neurosurgery
[3]Department of Neuroscience
[4]Department of Pathology
Johns Hopkins University School of Medicine, Baltimore, MD 21231, USA
[5]Department of Bioengineering, University of Washington, Seattle, WA 98195, USA
[6]Department of Radiology, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA
[7]Department of Biomedical Engineering and Yale Systems Biology Institute, Yale University, New Haven, CT 06516, USA
*Correspondence: aquinon2@jhmi.edu (A.Q.-H.), andre.levchenko@yale.edu (A.L.)
http://dx.doi.org/10.1016/j.celrep.2016.05.042

SUMMARY

Glioblastoma multiforme is a heterogeneous and infiltrative cancer with dismal prognosis. Studying the migratory behavior of tumor-derived cell populations can be informative, but it places a high premium on the precision of in vitro methods and the relevance of in vivo conditions. In particular, the analysis of 2D cell migration may not reflect invasion into 3D extracellular matrices in vivo. Here, we describe a method that allows time-resolved studies of primary cell migration with single-cell resolution on a fibrillar surface that closely mimics in vivo 3D migration. We used this platform to screen 14 patient-derived glioblastoma samples. We observed that the migratory phenotype of a subset of cells in response to platelet-derived growth factor was highly predictive of tumor location and recurrence in the clinic. Therefore, migratory phenotypic classifiers analyzed at the single-cell level in a patient-specific way can provide high diagnostic and prognostic value for invasive cancers.

INTRODUCTION

Aggressive cancers, such as glioblastoma multiforme (GBM), are of particular interest due to the heterogeneous nature of individual tumors (Snuderl et al., 2011; Szerlip et al., 2012) and high recurrence following surgical resection (Filippini et al., 2008; McGirt et al., 2009; Chaichana et al., 2013, 2014). Genomic and proteomic profiling can provide a wealth of information about tumor samples, including cancer-specific mutations and clinically relevant subclasses (Verhaak et al., 2010). However, these cell population-based analyses ignore the diversity of individual cells that can predetermine the aggressiveness of a given tumor. High-throughput genomic and proteomic single-cell analyses of multiple samples are not within reach of clinical applications (Kalisky et al., 2011; Meier et al., 2013). Furthermore, aggressive cell migration can be a product of multiple and distinct combinations of genetic alterations and would benefit from a complementary analysis of phenotypic properties at the individual cell level. As with many other cancers, individual GBM cells can spread from the primary tumor bulk, avoid detection, and form secondary tumor foci (Sahai, 2007). Growth factors, e.g., platelet-derived growth factor (PDGF), have emerged as enhancers of malignant potential in GBM, because they affect cell migration and proliferation (Fomchenko and Holland, 2007; Shih and Holland, 2006). Various components of the extracellular matrix (ECM), such as laminin, have also been implicated in modulation of cell migration (Friedl and Wolf, 2010; Petrie et al., 2009; Wirtz et al., 2011; Anton et al., 1999; Porcionatto, 2006). Directional cell migration can be guided by a variety of mechanical cues presented by ECM structures ranging in size from nanometers to microns (Kim et al., 2009a, 2012; Park et al., 2016). Therefore, a challenge is to develop an experimental platform that will model the mechano-chemical cellular milieu yet remain simple and accessible to allow practical, high-throughput use. In this study, we demonstrate that single-cell-resolution phenotypic screening holds great promise in prognostic analysis of glioblastoma samples. We found that the clinical outcome of GBM tumors strongly correlated with the response to two environmental inputs: the nanotopography and the growth factor PDGF. The ability to observe this responsiveness at the single-cell level and thus examine different cell subpopulations was critical for the success of this phenotypic analysis, revealing correlations with such critical prognostic tumor characteristics as time of recurrence after resection.

RESULTS

Construction and Application of a Phenotypic-Screening Platform

To create 1D fibrillar surfaces that mimic nanometer-scale features of the 3D ECM microenvironment, we fabricated

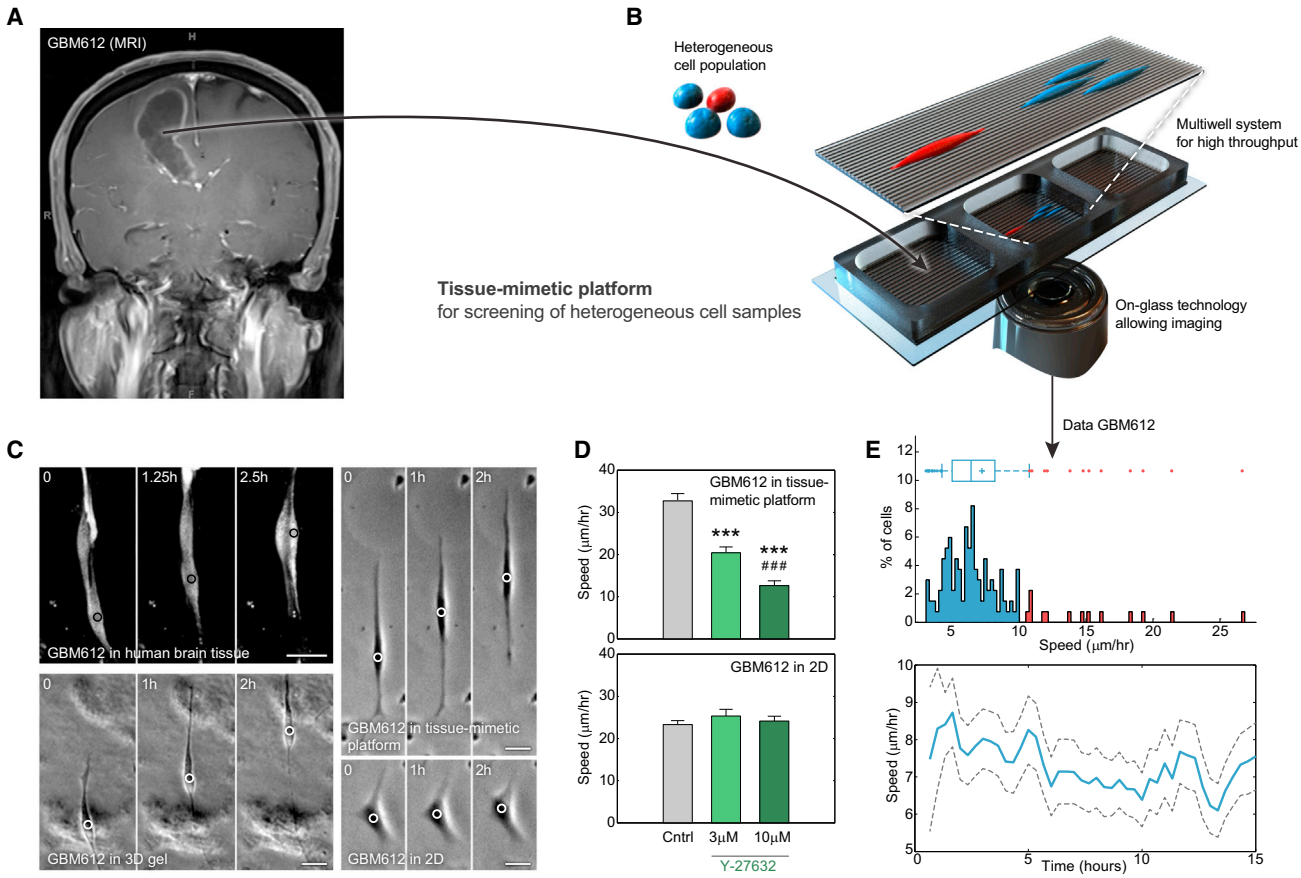


Figure 1. Phenotypic Screening of Heterogeneous Cell Populations Recapitulates the Microenvironment of Migrating Cells

(A) Cells with heterogeneous phenotypes are isolated from a patient's tumor (MRI of tumor for sample GBM 612).

(B) The cells are seeded on a platform that has a multi-well structure, allowing testing of multiple conditions, and is an on-glass technology, allowing direct imaging of migration and morphology with single-cell resolution.

(C) Images show that GBM 612 cells migrating on the platform have similar morphology and migration speed compared to GBM 612 cells migrating in ex vivo human brain tissue and 3D Matrigel. In comparison, cells migrating on flat surfaces are not polarized and have reduced migration speeds compared to the other cases (see Supplemental Information) (scale bars, 25 μm; duration between each frame, 1 hr).

(D) ROCK inhibitor (Y-27632) affects migration of GBM 612 cells on tissue-mimetic substrates and flat surfaces (n = 30 cells, mean + SEM, ***p < 0.0001, *paired against control group, #paired against 3 μM group, Kruskal-Wallis one-way ANOVA on ranks, Dunn's method).

(E) The platform provides important information on the migration response of heterogeneous cell populations. GBM 612 samples show a subpopulation of cells whose migration is fast and stable over time.

topographic patterns consisting of regular, parallel ridges (Figures 1A, 1B, and S1A–S1C) similar in size to those found in the brain tissue ECM (Kim et al., 2009b; Bellail et al., 2004; Ottani et al., 2001). We explored how well cell migration on the quasi-3D, fibrillar topography approximated cell migration in a true 3D ECM environment. We compared the morphology, directionality, and speed of GBM 612 cells on our platform and cells migrating in distinct 3D settings: Matrigel matrix and organotypic human brain slice cultures (Figure 1C). We found that cell morphology on fibrillar, but not flat, surfaces mirrored the characteristic cell shapes found in 3D environments. Cell morphology in collagen 3D matrices was similar to that observed in our tissue mimetic platform (Huang et al., 2016). Furthermore, in contrast to 2D surfaces, migration speed was enhanced to levels similar to those observed in 3D settings (Figure 1C), displaying essentially 1D migration patterns. We replicated the same observations using other patient-derived cell lines (GBM 318 and GBM 276) (Figures S1B, S1D, and S1J). Consistent with a tendency of mobile cells to align parallel to oriented topographic structures (Friedl and Wolf, 2010), the direction of cell migration was strongly biased along the axis of the ridge pattern (Figures S1D and S1J; Movies S1, S2, and S3). Compared to smooth surfaces, cells cultured on the platform showed increased cell area and spindle shape factor (Figures S1E, S1H, and S1I), and their migration was enhanced based on the three metrics scored: average speed, alignment, and persistence (Figures S1F, S1G, S1K, and S1L). These results suggested a high degree of similarity between our platform and the true 3D microenvironments, possibly due to similar molecular mechanisms observed in cells migrating on 1D fibrillar surfaces and 3D matrices but distinct from those observed in cells cultured on 2D surfaces (Doyle et al., 2009). To test the similarity between 1D fibrillar surfaces

and 3D microenvironments, we analyzed the effect of the Rho-associated protein kinase (ROCK) inhibitor Y-27632 on the migration of GBM 612 cells (Figure 1D). We found that Y-27632 inhibited migration on fibrillar surfaces but not on flat surfaces, suggesting that myosin II plays a role in migration on the fibrillar surfaces, similar to findings in vivo (Beadle et al., 2008). We replicated these experiments with GBM 965 (Figures S1M and S1N). Using the multi-well setup of the device, we then explored the influence of two critical environmental cues: ECM density and growth factor (Figure S2). On the fibrillar surfaces, but not on the flat surfaces, we observed a strong dependence of cell velocity values on the surface density of laminin (Figures S2A, S2B, S2E, and S2F). The density at which cell migration speed was maximized was used in all subsequent experiments. Next, we examined the distributions of cell velocities. We found that cell migration was highly heterogeneous, displaying a substantial number of outliers, some exceeding the average cell migration 3- to 4-fold (Figure 1E). Our results confirm that the heterogeneity of cell behavior can be assessed in this platform, enabling the identification and characterization of rare cells with extreme properties.

Migratory Behavior of GBM Cells Can Be Altered in Response to a Combination of PDGF and Nanotopographic Cues

We explored whether 3D-like cell migration might be differentially sensitive to the effects of growth factors implicated in the onset and progression of glioma, e.g., PDGF-AA (PDGF). PDGF can control glioma cell proliferation, but its effect on GBM migration and invasion is less clear (Feng et al., 2012; Laurent et al., 2003). We tested the effects of PDGF on cell speed and persistence at different doses. On our platform, we found a dose-dependent response, with maximal motility achieved at intermediate PDGF concentrations (Figure S2D, S2G, and S2H). Similar experiments on flat substrata showed more limited response to PDGF (Figure S2C). We then explored whether the effect of PDGF could be ascribed to the activation of PDGF receptor alpha (PDGFRα). PDGFRα is thought to be the exclusive receptor for the PDGF-AA isoform employed in this study (Fomchenko and Holland, 2007). First, we examined the expression levels of PDGFRα in patient-derived cell lines by RT-PCR and immunoblotting (Figure 2A, 2B, and S3A–S3C). Then, we analyzed the migratory behavior of cells with low and high PDGFRα expression levels (GBM 253 and GBM 276, respectively). Following exposure to PDGF, we found an increase in cell speed and directionality of GBM 276 cells (high PDGFRα expression) but no response in GBM 253 (low PDGFRα expression) (Figures 2C–2G). To further ascertain that PDGFRα was functionally involved, we used the tyrosine kinase inhibitor imatinib and found that migration was attenuated to similar levels in both cell lines (Figures 2C, 2D, S3D, and S3E). However, a detailed analysis of the GBM 276 subpopulations showed a highly heterogeneous single-cell response to PDGF. In particular, only the fastest 25% quartile of cells responded to PDGF, with the response abrogated by the inhibitor (Figures 2F and 2G). Furthermore, the response of the slowest 25% quartile of GBM 276 cells was analogous to that of the PDGF-unresponsive cell line GBM 253 (Figures 2E and 2F). These results suggest that migration of only a subset of cells is responsive to PDGF stimulation and that this subset represents the fastest cell subpopulation.

PDGF Enhances Invasiveness of Patient-Derived Cells In Vivo

Using orthotopic human GBM tumor models in mice (Figures S3F–S3J) (Garzon-Muvdi et al., 2012; Gonzalez-Perez et al., 2010), we explored the role of PDGF in tumor growth and survival. First, we examined the behavior of the xenografts using GBM 276 cells, cultured in the presence or absence of PDGF for 3 weeks before injection. This PDGF preculture selects for PDGF-responsive cells by stimulating their growth and enriching this cell subpopulation. In vitro analysis of GBM 276 cells showed that increased proliferation correlated with the PDGF-induced migratory response (Figures S3M–S3R). We observed significantly reduced survival of mice injected with PDGF-preconditioned cells (n = 4 each group) (Figure 2H). Although this result suggested the importance of highly PDGF-responsive cells for tumor aggressiveness, we could exclude other effects of prolonged PDGF exposure, e.g., transdifferentiation (Figure S3Q). We thus tested the putative role of PDGF in tumor spreading by supplying this factor exogenously in vivo to existing tumor xenografts via infusion pump (Figure S3F). Quantification of tumor size and qualitative analysis by a blinded neuropathologist suggested that continuous exposure of tumor xenografts to PDGF generated larger, more invasive tumors with more eccentric shapes (Figure S3K). These samples displayed features indicative of migration along fiber tracts. We also observed increased dispersion of GBM cells beyond the tumor margins (Figure S3L). These findings confirm the role of this growth factor in tumor induction and progression that was observed in other animal models (Fomchenko and Holland, 2007; Jackson et al., 2006; Lokker et al., 2002), and they are consistent with clinical data from The Cancer Genome Atlas (TCGA) (Goswami and Nakshatri, 2013) suggesting significant correlation between PDGF expression and survival (Figure 2I). TCGA data did not show a similar correlation for PDGFRα expression (Figure 2J). Because the genomic data did not support the correlation between average PDGFR expression and patient outcomes, we examined the correlation between the tumor characteristics and the cell migration data obtained in the tissue-mimetic platform.

Screening Heterogeneity within and between Patient-Specific Tumor Samples

Heterogeneity of cell properties within the same tumor reflects subpopulations promoting tumor growth, progression, and therapeutic resistance (Snuderl et al., 2011). GBM is also known to have populations with distinct expression profiles of receptor tyrosine kinases, particularly PDGFRα (Snuderl et al., 2011; Szerlip et al., 2012). This heterogeneity requires analysis on the single-cell level, which is yielded in our platform with less than 1,000 cells (particularly beneficial for screening precious intraoperative human tissue specimens). We took advantage of the single-cell resolution to quantify the distribution of cell speed in control versus PDGF-exposed conditions and investigate the difference in migratory behaviors among 14 glioblastoma patients (Figures 3A–3C). We found both intra- and inter-patient

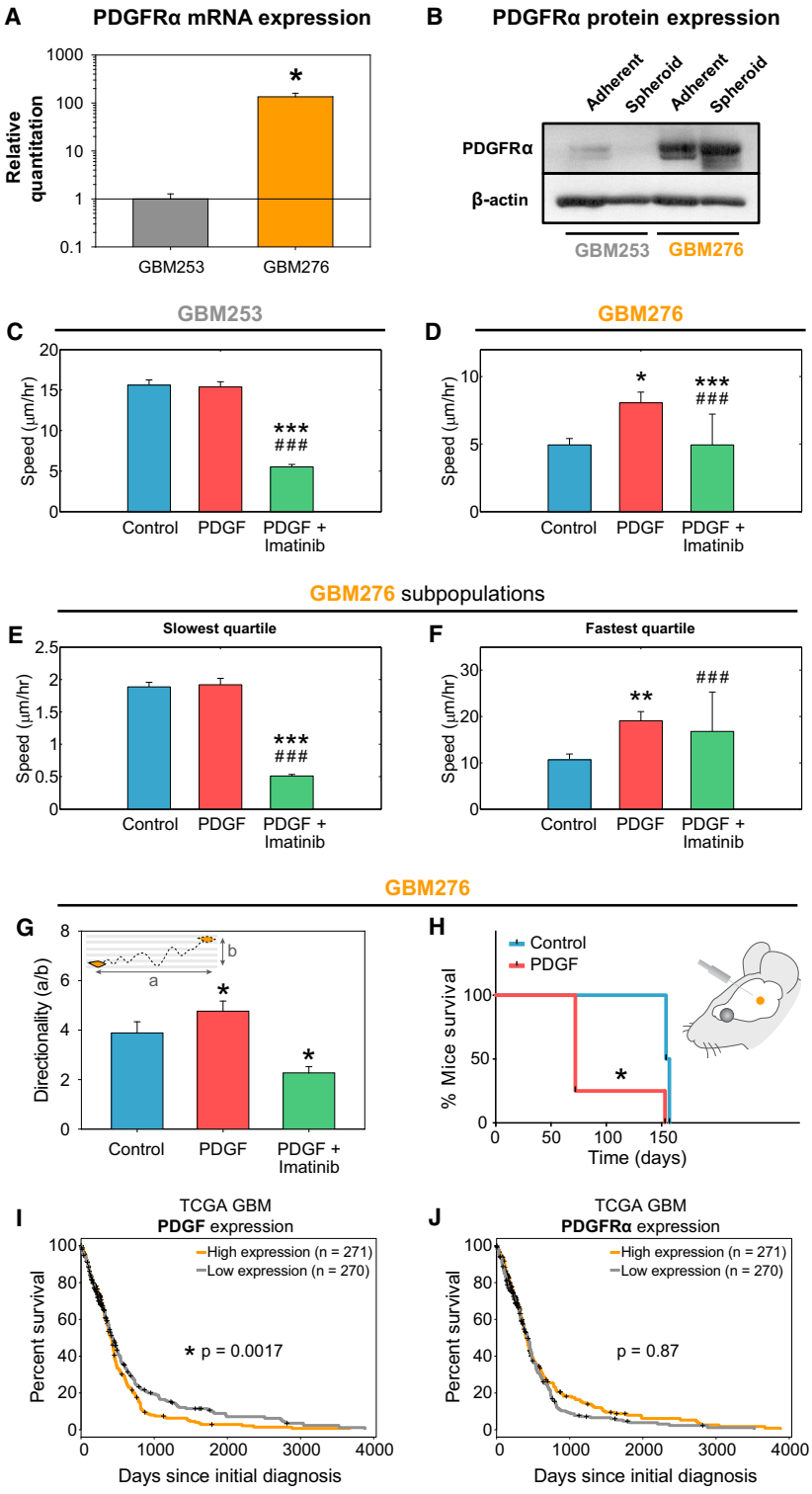


Figure 2. Migratory Response to PDGF Correlates with Tumor Characteristics Both In Vitro and In Vivo

(A) RT-PCR analysis of PDGFRα mRNA levels in responding sample GBM 276 and non-responding sample GBM 253 (n = 3, *p = 0.006, Student's t test).

(B) Western blot for PDGFRα protein expression in GBM samples grown as adherent or spheroid cultures.

(C and D) Quantification of migration speed of cell lines GBM 253 (C) and GBM 276 (D) in the presence of PDGF-AA (50 ng/ml) and imatinib (30 μM) (n ≈ 80 cells, mean + SEM, *p < 0.05, **p < 0.01, ***p < 0.001, *paired against control group, #paired against PDGF group, Kruskal-Wallis one-way ANOVA on ranks, Dunn's method) (asterisks indicate pairing against the control group, hash marks indicated pairing against the PDGF group).

(E and F) Migration speed of GBM 276 for the slowest (E) and fastest (F) quartile of the cells (i.e., 25% of the slowest- and fastest-moving cells, respectively), showing that only a subpopulation responds to PDGF.

(G) Quantification of migration measured by alignment of GBM 276 cells (*p < 0.05, Kruskal-Wallis one-way ANOVA on ranks, Dunn's method) (asterisks indicate comparisons to all other conditions).

(H) Survival curves of mice injected with GBM 276 cells cultured in control spheroid conditions or in the presence of PDGF-AA (n = 4 mice per group, *p = 0.0097, Gehan-Breslow-Wilcoxon test).

(I and J) Kaplan-Meier plots based on clinical TCGA data of GBM patients, comparing survival between high and low expression of PDGF-AA (I) and PDGFRα (J), respectively. The cohorts were divided at the median of the expression level of the respective gene.

differences in the cell population behavior. When analyzing GBM 499 cells based on their speed, the total population average showed no significant response to PDGF, while analysis of the 25% fastest quartile subpopulation revealed significant differences (Figure 3A). Such masked responses and heterogeneities were also present in the time-domain data. For instance, two patient samples, GBM 501 and GBM 609, responded significantly to PDGF. However, for GBM 501, this response was not persistent throughout the experiment duration, in contrast to the response of GBM 609 (Figure 3B). Finally, the platform allowed us to investigate high-speed outliers. For example, for both the GBM 630 and the GBM 544 samples, cells experienced a significant increase in migration speed in response to PDGF. However, a detailed analysis of the speed distribution revealed that only for GBM 630 were fast-moving outliers clearly identifiable; GBM 544 showed a substantially more uniform response (Figure 3C).

We observed considerable differences in the cell speed across the spectrum of 14 patient-derived samples, both in the presence and in the absence of PDGF (Figure 3D). Recognizing these patient-specific differences, we analyzed the distinct features of single-cell response based on the criteria presented in Figures 3A–3C, namely, based on the average population response (group I), persistence of response over time

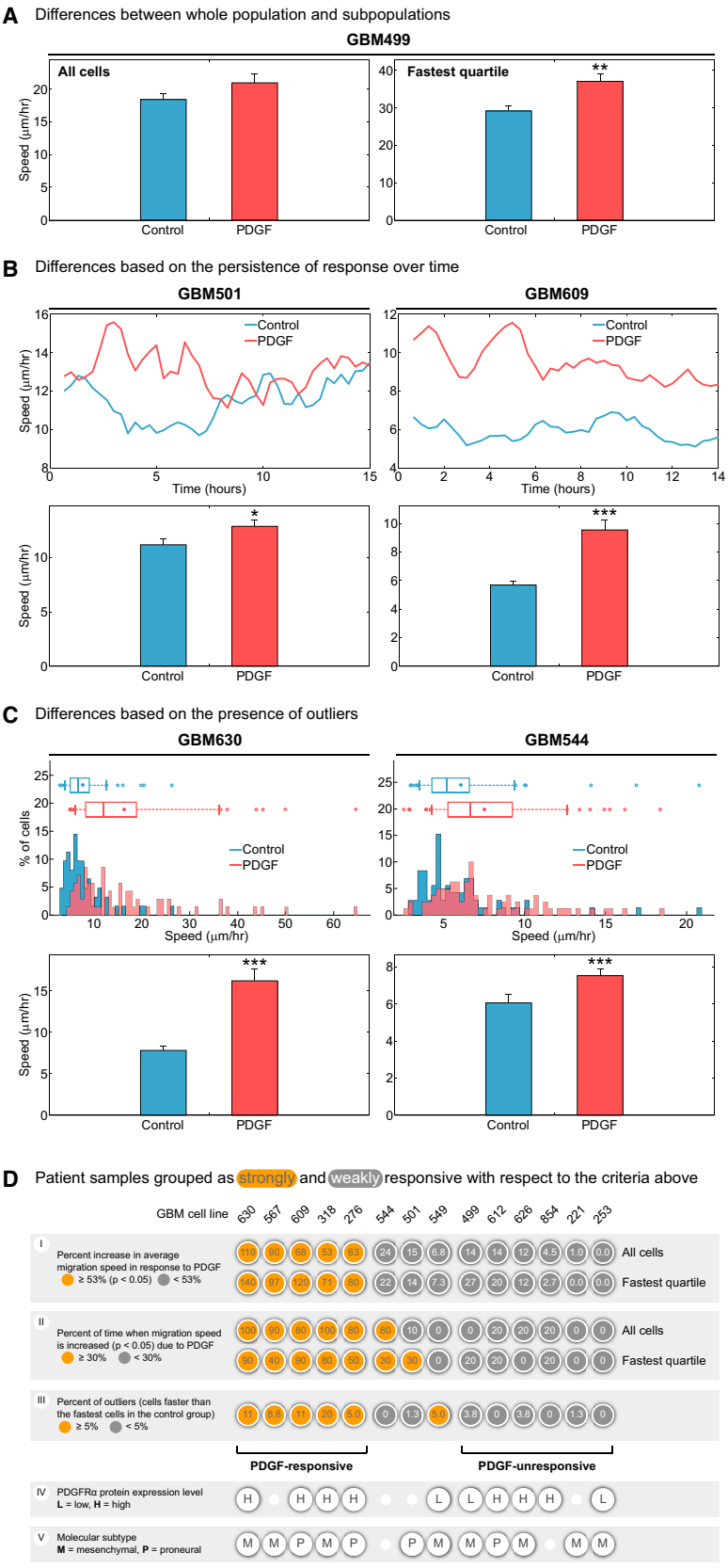


Figure 3. Information on Migration Speed Reveals Important Differences among Patient Samples in Response to PDGF

(A) Analyzing the fastest quartile (GBM 499) reveals that the subpopulations display a significant response to PDGF. In contrast, for the whole population, there is no significant response.

(B) Migration speed time lapse demonstrates that sample GBM 501 does not respond to PDGF at all times (compared to GBM 609); on average, however, both samples respond significantly to PDGF.

(C) Both GBM 630 and GBM 544 samples display a significant increase in average migration speed (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, Wilcoxon rank-sum test). However, GBM 630 has a significantly larger number of fast outliers, while GBM 544 displays a uniform increase in speed.

(D) The platform allows patient sample classification based on multiple characteristics, permitting a better description of the heterogeneity of the samples. We compare 14 patients' GBM cell lines. The samples were grouped based on whether there is a significant increase ($p < 0.05$, Wilcoxon rank-sum test) in average migration speed in response to PDGF (group I) and whether this significant increase is persistent over time (group II). Furthermore, samples are grouped based on the number of outliers (cells faster than the fastest cells in the control group) with a threshold of 4 cells (5%, for a total of 80 cells) (group III). For reference, the PDGFRα protein expression level (group IV) and the subclass of the GBM cells (group V) are also provided.

(group II), and presence of outliers (group III) (Figure 3D). For each of these criteria, we identified PDGF-responsive and PDGF-unresponsive groups. There was overlap among strongly PDGF-responsive groups in all groups, which we thus treated as the consensus PDGF-responsive samples (marked in Figure 3D). However, several patient samples (GBM 544, GBM 549, and GBM 501) were PDGF responsive by some criteria. To determine whether PDGFRα expression level could serve as a molecular marker predictive of enhanced PDGF responsiveness, we assayed the tumor samples for expression of this receptor. Although all consensus PDGF-responsive samples had high expression of PDGFRα (Figures 2A, 2B, S3B, and S3C), similarly high levels of the receptor were found in several samples in the consensus PDGF-unresponsive group (GBM 626, GBM 612, and GBM 854). The failure of these samples to respond in migration experiments may be due to differences in downstream effectors of PDGFRα, stressing the difficulty of using molecular markers alone in the classification of the aggressive migration phenotype, as is evident from clinical data showing no correlation between PDGFRα expression and survival (Figure 2J). Differential expression of PDGFRα across the groups might also relate to the subtype of the patient tumors. A subclassification study of our GBM samples only yielded two subtypes: mesenchymal and proneural. Previous studies have observed higher amplification and mutation rates of the PDGFRA gene in the proneural subtype of GBM (Verhaak et al., 2010). We did not observe a clear correlation between expression of this gene and tumor subclass. High levels of PDGF-triggered cell migration can be achieved in various ways, and they may not be revealed by a simple molecular signature. However, the aggressive migration phenotype can be translated into enhanced invasiveness. These results highlight potential advantages of our single-cell phenotype analysis.

Migratory Behavior Correlates with Clinical Tumor Characteristics

The degree of cell migration may reflect the propensity for invasive tumor spread. We examined more than 35 factors related to each patient's tumor, general health, and demographics (Table S1). We found that the migratory response of GBM samples to PDGF correlated with time to tumor recurrence after surgical resection (Figures 4A and 4B). This correlation was particularly significant when the analysis was focused on the consensus-responsive and consensus-unresponsive groups (Figure 4A). In comparisons to the whole-cell populations, correlations were more significant for the aggressively moving cells: either the fastest 25% of the cells or the outlier population (Figure 4B). We contrasted several characteristic tumor features visualized in the MRI of the patients with the responsiveness to PDGF (Figures 4C and 4D). Tumors from the consensus PDGF-responsive subset (Figure 4D) were larger and more spread out than those from the PDGF-unresponsive subset. We observed statistically significant differences in the anatomical location of the tumors; all consensus PDGF-responsive samples were in the frontal lobe (Figure 4E), but temporal lobe tumor samples commonly fell into the PDGF-unresponsive subgroup. Migration analysis of GBM cells revealed that higher directionality (i.e., alignment of the migration to the patterns) correlated with longer recurrence times (univariate Cox analysis, $p = 0.002$) (Figures 4F and S4A). Because alignment of the migration is associated with the strength of cell-substrate adhesion (Kim et al., 2009a; Garzon-Muvdi et al., 2012), this result may reflect a higher propensity in more aligned cells to adhere to ECM, leading to retarded migratory response and delayed tumor spread. Blind, qualitative analysis of patient tumor samples by a neuropathologist revealed that the cells were small and that the tumors yielding PDGF-responsive samples had marked microvascular proliferation. The latter feature is commonly associated with advanced progression, receptor tyrosine kinase amplification, and worse prognosis (Louis, 2006; Fomchenko and Holland, 2007). PDGF-responsive samples emerged from tumors that resulted in shorter recurrence times, after controlling for factors known to be associated with recurrence (age, Karnofsky Performance Scale score, extent of resection, and adjuvant therapy) ($p = 0.0009$) (Table S1). However, survival times showed a less significant correlation to PDGF responsiveness with the limited number of patient samples tested (Figure S4B).

We supplemented these findings with comparisons to more traditional protein expression analyses. First, we segregated the tumor sample into groups based on PDGFRα expression, which revealed trends in patient tumor characteristics that were supportive of our migration analysis (Figure 3D). However, grouping the samples according to PDGFRα expression or molecular subclassification did not yield significant differences in predicting time to recurrence (Figures S4C and S4D). Moreover, the difference in tumor location between high- and low-expression groups was not statistically significant. Second, we mined the public Rembrandt database of glioblastoma patients to investigate the relationship between tumor location and PDGFRα expression. Examination of 47 patients revealed no significant relationship between the two characteristics (Figures S4E and S4F). Taken together, these findings suggest that RNA expression is a weaker predictor of patient outcomes than is our phenotypic migration analysis.

DISCUSSION

The heterogeneity and invasive nature of glioblastoma and other aggressive cancers highlights the importance of assaying cell migration as a phenotypic feature predictive of clinical outcomes. Here we describe a simple but information-rich experimental platform aimed at the analysis of primary patient samples on a single-cell level. This platform allows high-throughput screening of the effects of variable extracellular milieu. Using this method on a range of patient-derived samples and contrasting the results of the analysis with respective clinical information revealed substantial predictive power, particularly when cell migration was examined in conjunction with the effects of PDGF. This result strongly suggests that cell migration, as examined in structured, mechanically defined culture conditions, can be predictive of more complex in vivo invasion processes and can be a powerful phenotypic analysis tool with strong clinical implications. Prior attempts to examine glioma cell migration and its relationship to tumor progression (Friedlander et al., 1996) have used 2D surfaces and have not achieved such direct predictions of patient-specific tumor features as we have in the

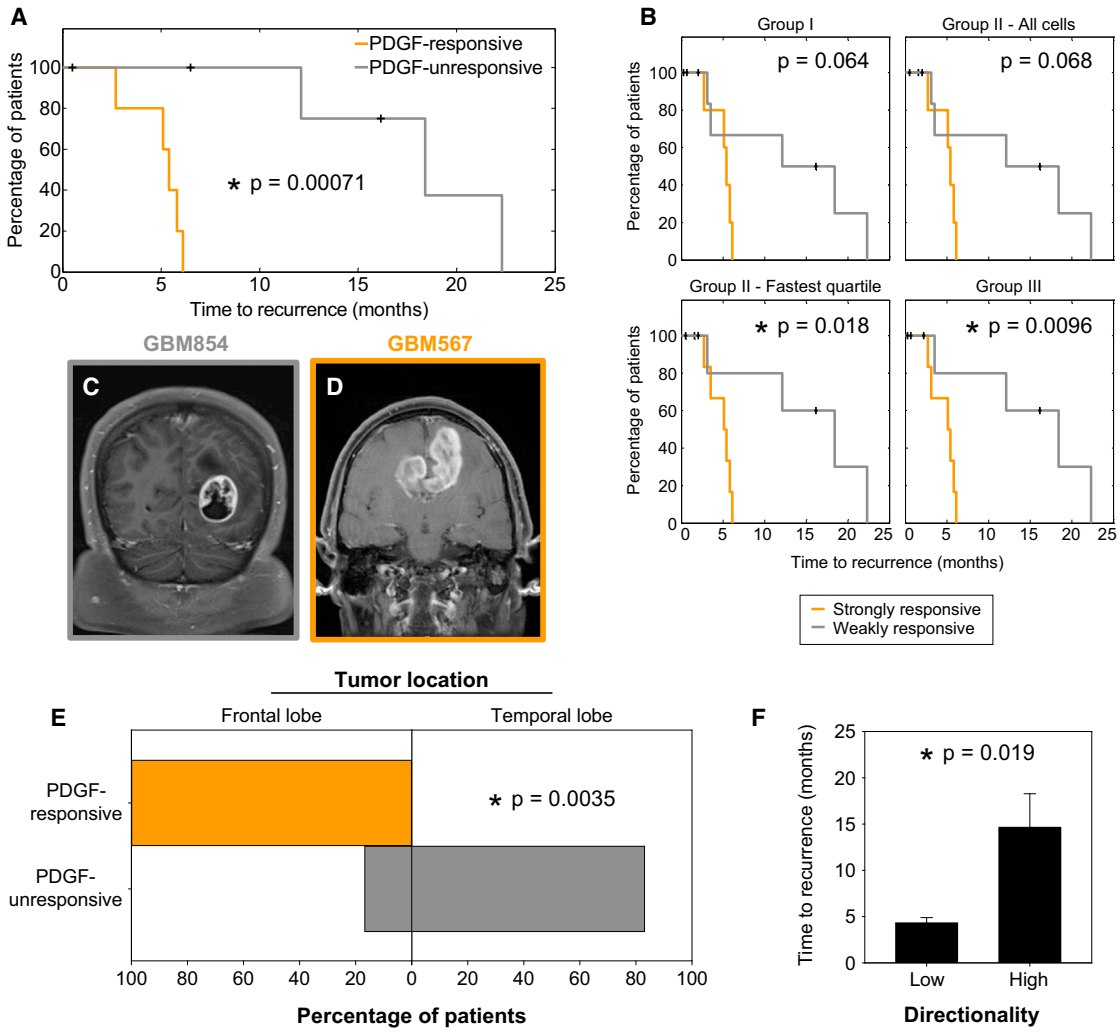


Figure 4. GBM Migratory Response to PDGF Correlates with Patient Tumor Characteristics

(A and B) Kaplan-Meier plots comparing recurrence between PDGF-responsive and PDGF-unresponsive groups (n = 11 patient tumors, consensus group) (A) and based on the criteria in Figure 3D (n = 14 patient tumors, weak and strong responders) (B). The p values were calculated using a two-tailed log-rank (Mantel-Cox) test.

(C and D) MRI scans of patients with tumors from the unresponsive group (C) and the responsive group (D).

(E) Distribution of PDGF-responsive and PDGF-unresponsive tumors that formed in specific locations in the brain (n = 11 patient tumors, Barnard's exact test).

(F) Time to recurrence for GBM samples separated into low-directionality (<3.25) and high-directionality (≥3.25) groups (n ≥ 4, mean + SEM, Wilcoxon rank-sum test) (the threshold of 3.25 was determined using linear discriminant analysis).

present work. This emphasizes the benefits of analyzing heterogeneities within samples and using surfaces that better mimic in vivo conditions.

The significant correlation of migratory behavior with time to recurrence and tumor location provides crucial insight into this disease. Recurrence of glioblastoma after tumor resection is the primary cause of death in patients and is one of the most important predictors of future patient outcomes (Chaichana et al., 2013, 2014). This study provides a simple method to glean information about these phenomena. Our finding that most PDGF-unresponsive tumors are derived from the temporal lobe could suggest that PDGF signaling is less critical to tumor progression in this region of the brain. Direct access to individual cell migration analysis can be important for future treatment modalities. Our experimental platform has important advantages over 2D migration assays, because it provides a cellular environment similar to in vivo conditions (as evidenced in the similarity of several aspects of migration in ex vivo human brain tissue and a 3D hydrogel, e.g., increased cell polarity and migration speed). Previous studies have highlighted the importance of these factors in migration (Friedl and Wolf, 2010; Petrie et al., 2009; Kim et al., 2012; Beadle et al., 2008; Louis, 2006). Another advantage is the reduced number of cells required when compared to commonly used transwell migration assays. Furthermore, transwell assays fail to yield the information on migration and morphology of individual cells and only originate endpoint

information (Kim et al., 2009a). We found a substantial degree of heterogeneity in the glioblastoma samples analyzed. The increased average migration speed of a cell population in the presence of PDGF was ascribed to a small subpopulation of aggressive cells (approximately 25%). Knowledge of the degree of population heterogeneity can be critical to the decision-making in the clinic (Snuderl et al., 2011; Szerlip et al., 2012). In addition, our tissue mimetic platform can distinguish the effects of cell proliferation and migration phenotypes, which can be a confounding factor in both transwell and in vivo migration studies.

Our results also highlight advantages of the proposed method over traditional protein expression assays. We observed an incomplete correlation between receptor expression and response to PDGF signaling, possibly due to veiled differences in the signal transduction pathways. We demonstrated that protein expression analysis was less sensitive and less robust at predicting differences among patient tumor features. Genomic and proteomic approaches also suffer from limited supplies of primary tissues available for their cumbersome requirements of cellular material.

Overall, the results here support the proposed methodology as a simpler, more biomimetic, and informative method to gain critical information about patient tumors and cell populations. The analysis presented here reveals the importance of careful engineering of chemical and mechanical extracellular milieu in cell migration analysis. We believe that this methodology will provide an important prognostic tool, with benefits that include high-throughput, label-free analysis of single-cell resolution; low demand for precious primary cell samples; and better physiological relevance compared to other migration assays.

EXPERIMENTAL PROCEDURES

Cells

Human tissues were obtained at Johns Hopkins medical institutions and used with approval of the Institutional Review Board. Glioblastoma pathologically confirmed tumor samples (GBM 221, GBM 253, GBM 276, GBM 318, GBM 499, GBM 501, GBM 544, GBM 549, GBM 567, GBM 609, GBM 612, GBM 626, GBM 630, and GBM 854) were derived from primary intraoperative tissues of patients undergoing surgery. Tissue donors received no treatment before surgery.

Construction of a Multi-well Nanopatterned Device

The topographic nanopatterned substratum, consisting of parallel ridges 350 nm wide and 500 nm high that are spaced 1.5 μm apart, was fabricated onto glass coverslips as previously described (Kim et al., 2009a, 2009b) using UV-assisted capillary molding techniques.

Time-Lapse Microscopy of Live Cells and Quantitative Analysis of Cell Morphology and Migration

Cell migration was observed using time-lapse microscopy (Movies S1, S2, and S3). To enable long-term observation, the multi-well, nanopatterned device was mounted on the stage of a motorized inverted microscope (Olympus IX81) equipped with a Photometrics hCascade 512B II charge-coupled device camera and temperature- and gas-controlling environmental chamber. Phase-contrast cell images were automatically recorded under a 10× objective (numerical aperture = 0.30) using SlideBook 4.1 software (Intelligent Imaging Innovations) for 10–15 hr at 10 or 20 min intervals. Because cell-cell contact is known to affect the extent of cell spreading and migration, cells were plated at low density (~4×10^4 cells ml^{-1}) to allow isolated movements. A custom-made MATLAB script was used to allow manual tracking and measurement of cells frame by frame.

Tumor Xenografts

Animal protocols were approved by the Johns Hopkins School of Medicine Animal Care and Use Committee. For intracranial xenografts, severe combined immunodeficiency mice received 100,000 viable cells in 1 μl of DMEM/F12 serum media without growth factors by stereotactic injection into the right striatum. Cells were cultured in DMEM/F12 serum media with epidermal growth factor, fibroblast growth factor, and PDGF ligand for 3 weeks before injections were performed. Cell viability was determined by trypan blue dye exclusion. Mice were perfused with 4% paraformaldehyde at the indicated times, and the brains were removed for histological analysis.

Statistical Analysis

Results are presented as mean + SEM. The Mann-Whitney rank-sum test was for pairwise comparisons; Dunn's test (rank-based ANOVA) was used in multiple group comparisons. When noted, Student's t test or standard ANOVA (the Holm-Sidak method) was used. Univariate Cox analysis was used to identify correlations among tumor characteristics. To group data, thresholds were determined using linear discriminant analysis as previously described (Lin et al., 2012). Statistics were analyzed using Sigmaplot, GraphPad Prism, and MATLAB software.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, one table, and three movies and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.05.042.

AUTHOR CONTRIBUTIONS

Conceptualization, C.L.S., O.K., P.S., A.Q.-H., and A.L.; Analysis, C.L.S., O.K., P.S., F.J.R., and K.L.C.; Investigation, C.L.S., O.K., P.S., N.I.S.-R., H.G.-C., S.G., T.O., S.A., and J.P.; Writing, C.L.S., O.K., P.S., H.G.-C., A.Q.-H., and A.L.; Supervision, A.Q.-H. and A.L.

CONFLICTS OF INTEREST

Some authors are listed as inventors in pending patents (C.L.S., D.-H.K., O.K., P.S., A.Q.-H., and A.L.).

ACKNOWLEDGMENTS

We thank H. Basdag, L. Johnson, and L. Noiman for establishing cell cultures; H. Nam Kim and K.Y. Suh for help in building nanopatterned surfaces; and M. Delannoy for assistance with SEM. This work was funded by NIH RO1 NS070024 (to A.Q.-H.), a Ford Foundation fellowship (to C.L.S.), AHA fellowship 13POST17140090 (to O.K.), and NIH grants U01CA15578 and CA16359 (Yale Cancer Center) (to A.L.).

Received: December 9, 2015
Revised: February 24, 2016
Accepted: May 9, 2016
Published: June 9, 2016

REFERENCES

Anton, E.S., Kreidberg, J.A., and Rakic, P. (1999). Distinct functions of α3 and α(v) integrin receptors in neuronal migration and laminar organization of the cerebral cortex. Neuron *22*, 277–289.

Beadle, C., Assanah, M.C., Monzo, P., Vallee, R., Rosenfeld, S.S., and Canoll, P. (2008). The role of myosin II in glioma invasion of the brain. Mol. Biol. Cell *19*, 3357–3368.

Bellail, A.C., Hunter, S.B., Brat, D.J., Tan, C., and Van Meir, E.G. (2004). Microregional extracellular matrix heterogeneity in brain modulates glioma cell invasion. Int. J. Biochem. Cell Biol. *36*, 1046–1069.

Chaichana, K.L., Zadnik, P., Weingart, J.D., Olivi, A., Gallia, G.L., Blakeley, J., Lim, M., Brem, H., and Quiñones-Hinojosa, A. (2013). Multiple resections for patients with glioblastoma: prolonging survival. J. Neurosurg. *118*, 812–820.

Chaichana, K.L., Jusue-Torres, I., Navarro-Ramirez, R., Raza, S.M., Pascual-Gallego, M., Ibrahim, A., Hernandez-Hermann, M., Gomez, L., Ye, X., Weingart, J.D., et al. (2014). Establishing percent resection and residual volume thresholds affecting survival and recurrence for patients with newly diagnosed intracranial glioblastoma. Neuro Oncol. *16*, 113–122.

Doyle, A.D., Wang, F.W., Matsumoto, K., and Yamada, K.M. (2009). One-dimensional topography underlies three-dimensional fibrillar cell migration. J. Cell Biol. *184*, 481–490.

Feng, H., Liu, K.W., Guo, P., Zhang, P., Cheng, T., McNiven, M.A., Johnson, G.R., Hu, B., and Cheng, S.Y. (2012). Dynamin 2 mediates PDGFRα-SHP-2-promoted glioblastoma growth and invasion. Oncogene *31*, 2691–2702.

Filippini, G., Falcone, C., Boiardi, A., Broggi, G., Bruzzone, M.G., Caldiroli, D., Farina, R., Farinotti, M., Fariselli, L., Finocchiaro, G., et al.; Brain Cancer Register of the Fondazione IRCCS (Istituto Ricovero e Cura a Carattere Scientifico) Istituto Neurologico Carlo Besta (2008). Prognostic factors for survival in 676 consecutive patients with newly diagnosed primary glioblastoma. Neuro-oncol. *10*, 79–87.

Fomchenko, E.I., and Holland, E.C. (2007). Platelet-derived growth factor-mediated gliomagenesis and brain tumor recruitment. Neurosurg. Clin. N. Am. *18*, 39–58, viii.

Friedl, P., and Wolf, K. (2010). Plasticity of cell migration: a multiscale tuning model. J. Cell Biol. *188*, 11–19.

Friedlander, D.R., Zagzag, D., Shiff, B., Cohen, H., Allen, J.C., Kelly, P.J., and Grumet, M. (1996). Migration of brain tumor cells on extracellular matrix proteins in vitro correlates with tumor type and grade and involves alphaV and β1 integrins. Cancer Res. *56*, 1939–1947.

Garzon-Muvdi, T., Schiapparelli, P., ap Rhys, C., Guerrero-Cazares, H., Smith, C., Kim, D.-H., Kone, L., Farber, H., Lee, D.Y., An, S.S., et al. (2012). Regulation of brain tumor dispersal by NKCC1 through a novel role in focal adhesion regulation. PLoS Biol. *10*, e1001320.

Gonzalez-Perez, O., Guerrero-Cazares, H., and Quiñones-Hinojosa, A. (2010). Targeting of deep brain structures with microinjections for delivery of drugs, viral vectors, or cell transplants. J. Vis. Exp. (*46*), 2082.

Goswami, C.P., and Nakshatri, H. (2013). PROGgene: gene expression based survival analysis web application for multiple cancers. J. Clin. Bioinforma. *3*, 22.

Huang, Y.-J., Hoffmann, G., Wheeler, B., Schiapparelli, P., Quinones-Hinojosa, A., and Searson, P. (2016). Cellular microenvironment modulates the galvanotaxis of brain tumor initiating cells. Sci. Rep. *6*, 21583.

Jackson, E.L., Garcia-Verdugo, J.M., Gil-Perotin, S., Roy, M., Quinones-Hinojosa, A., VandenBerg, S., and Alvarez-Buylla, A. (2006). PDGFR α-positive B cells are neural stem cells in the adult SVZ that form glioma-like growths in response to increased PDGF signaling. Neuron *51*, 187–199.

Kalisky, T., Blainey, P., and Quake, S.R. (2011). Genomic analysis at the single-cell level. Annu. Rev. Genet. *45*, 431–445.

Kim, D.-H., Han, K., Gupta, K., Kwon, K.W., Suh, K.-Y., and Levchenko, A. (2009a). Mechanosensitivity of fibroblast cell shape and movement to anisotropic substratum topography gradients. Biomaterials *30*, 5433–5444.

Kim, D.-H., Seo, C.-H., Han, K., Kwon, K.W., Levchenko, A., and Suh, K.-Y. (2009b). Guided cell migration on microtextured substrates with variable local density and anisotropy. Adv. Funct. Mater. *19*, 1579–1586.

Kim, D.H., Provenzano, P.P., Smith, C.L., and Levchenko, A. (2012). Matrix nanotopography as a regulator of cell function. J. Cell Biol. *197*, 351–360.

Laurent, M., Martinerie, C., Thibout, H., Hoffman, M.P., Verrecchia, F., Le Bouc, Y., Mauviel, A., and Kleinman, H.K. (2003). NOVH increases MMP3 expression and cell migration in glioblastoma cells via a PDGFR-alpha-dependent mechanism. FASEB J. *17*, 1919–1921.

Lin, B., Holmes, W.R., Wang, C.J., Ueno, T., Harwell, A., Edelstein-Keshet, L., Inoue, T., and Levchenko, A. (2012). Synthetic spatially graded Rac activation drives cell polarization and movement. Proc. Natl. Acad. Sci. USA *109*, E3668–E3677.

Lokker, N.A., Sullivan, C.M., Hollenbach, S.J., Israel, M.A., and Giese, N.A. (2002). Platelet-derived growth factor (PDGF) autocrine signaling regulates survival and mitogenic pathways in glioblastoma cells: evidence that the novel PDGF-C and PDGF-D ligands may play a role in the development of brain tumors. Cancer Res. *62*, 3729–3735.

Louis, D.N. (2006). Molecular pathology of malignant gliomas. Annu. Rev. Pathol. *1*, 97–117.

McGirt, M.J., Chaichana, K.L., Gathinji, M., Attenello, F.J., Than, K., Olivi, A., Weingart, J.D., Brem, H., and Quiñones-Hinojosa, A.R. (2009). Independent association of extent of resection with survival in patients with malignant brain astrocytoma. J. Neurosurg. *110*, 156–162.

Meier, M., Sit, R.V., and Quake, S.R. (2013). Proteome-wide protein interaction measurements of bacterial proteins of unknown function. Proc. Natl. Acad. Sci. USA *110*, 477–482.

Ottani, V., Raspanti, M., and Ruggeri, A. (2001). Collagen structure and functional implications. Micron *32*, 251–260.

Park, J., Kim, D.H., Kim, H.N., Wang, C.J., Kwak, M.K., Hur, E., Suh, K.Y., An, S.S., and Levchenko, A. (2016). Directed migration of cancer cells guided by the graded texture of the underlying matrix. Nat. Mater. Published online March 14, 2016. http://dx.doi.org/10.1038/nmat4586.

Petrie, R.J., Doyle, A.D., and Yamada, K.M. (2009). Random versus directionally persistent cell migration. Nat. Rev. Mol. Cell Biol. *10*, 538–549.

Porcionatto, M.A. (2006). The extracellular matrix provides directional cues for neuronal migration during cerebellar development. Braz. J. Med. Biol. Res. *39*, 313–320.

Sahai, E. (2007). Illuminating the metastatic process. Nat. Rev. Cancer *7*, 737–749.

Shih, A.H., and Holland, E.C. (2006). Platelet-derived growth factor (PDGF) and glial tumorigenesis. Cancer Lett. *232*, 139–147.

Snuderl, M., Fazlollahi, L., Le, L.P., Nitta, M., Zhelyazkova, B.H., Davidson, C.J., Akhavanfard, S., Cahill, D.P., Aldape, K.D., Betensky, R.A., et al. (2011). Mosaic amplification of multiple receptor tyrosine kinase genes in glioblastoma. Cancer Cell *20*, 810–817.

Szerlip, N.J., Pedraza, A., Chakravarty, D., Azim, M., McGuire, J., Fang, Y., Ozawa, T., Holland, E.C., Huse, J.T., Jhanwar, S., et al. (2012). Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response. Proc. Natl. Acad. Sci. USA *109*, 3041–3046.

Verhaak, R.G.W., Hoadley, K.A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M.D., Miller, C.R., Ding, L., Golub, T., Mesirov, J.P., et al.; Cancer Genome Atlas Research Network (2010). Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell *17*, 98–110.

Wirtz, D., Konstantopoulos, K., and Searson, P.C. (2011). The physics of cancer: the role of physical interactions and mechanical forces in metastasis. Nat. Rev. Cancer *11*, 512–522.