NEB EXPRESSIONS

A scientific update

Issue I • 2020

IN THIS ISSUE

Breeding a better tomato with the NEBNext Direct® Genotyping Solution

By Andrew Barry, MS, New England Biolabs

The global population is anticipated to eclipse 8 billion by 2025, and approach 10 billion people by 2050 (1). Along with this rate of growth comes global challenges for feeding this population, pressing the need for more efficient farming practices. This efficiency is perturbed by reduced land availability for farming, emergence of novel pathogens, diminishing table water, and observed changes in climate, generating the need for sustainable crops with improved resilience to these stresses.

Traditional crop breeding approaches rely on interbreeding, or "crossing" of plant varieties for allelic transfer to generate genetic diversity within new crop varieties. This is followed by phenotypic assessment, and subsequent backcrossing with the parental lines for selection of desired traits. These traits can be quantitative and practical, such as pathogen resistance, drought tolerance and improvements in crop yield, or they can be more subjective and aesthetic traits including flavor and color. While these methods are effective, they rely on plant growth, exposure to stress, and observation of the desired phenotype in order to assess the presence of the desired trait; therefore, the breeding process is greatly lengthened.

Quantitative trait locus, or QTL mapping, generates linkage information between a desired phenotype and the associated genotypic information. There are several approaches available to perform QTL mapping, but the goal is to identify a set of genetic markers that can be used in place of phenotypic information to assess whether plants are harboring the specific markers that are positive indicators for the presence of traits being selected for. The development of crop-specific databases to guide breeding programs has created a need for novel methods for genotyping plants that result from genetic crossing or backcrossing in order to guide future breeding efforts. To generate necessary genetic diversity, thousands to tens of thousands of plants are used for breeding efforts, so ideally these approaches are fast and scalable to address the throughput demand.

Traditional genotyping approaches include endpoint PCR assays, whereby only a limited number of markers for any given plant can be assayed, and input samples must be split across multiple PCR reactions, requiring high amounts of PCR consumables and specialized equipment for high-throughput sample processing. Another common option for genotyping is the use of SNP-based microarrays, which can assay hundreds of thousands of markers in parallel, yet the challenge lies in scaling the experiment for high sample numbers, because a single DNA sample per chip is required. In parallel, the creation of full genome reference sequences for many crops has increased our knowledge, and therefore, additional marker types including genomic insertions and deletions, and combinations

Figure 1A:

The NEBNext Direct Genotyping Solution Workflow

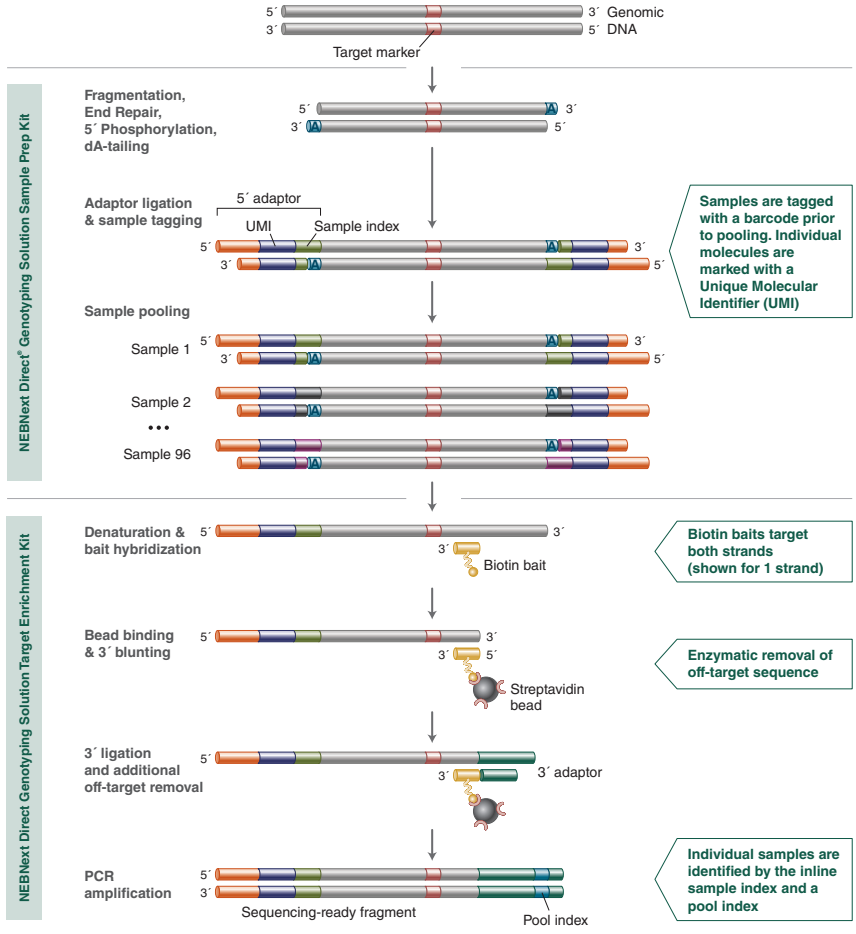
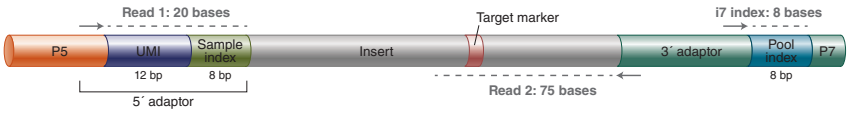


Figure 1B:

Final library and sequencing details



Using the NEBNext Direct Genotyping Solution, 25 ng of 96 individual tomato DNA samples were enzymatically fragmented and 5´ tagged with an Illumina®-compatible P5 adaptor that incorporates both an inline sample index to tag each sample prior to pooling and an inline UMI to mark each unique DNA fragment within the samples (Figure 1A). The 96 samples were pooled and enriched using SolCap panel in a single hybridization reaction, followed by library preparation and 16 cycles of PCR amplification. After purification and quantification, the 96-plex library was sequenced in a single MiSeq® run as indicated (Figure 1B). Following sequencing, reads were demultiplexed with a Picard-based workflow(3). Sequencing reads were aligned to the SL2.40 reference genome(4) using BWA-MEM(5) and PCR duplicates were identified using the unique molecular identifiers(6).

of markers, or haplotypes, are increasingly used as phenotypic indicators. Microarrays and endpoint PCR assays are limited in their ability to solely provide information on the presence or absence of a known marker and cannot be used for discovery of novel genomic information.

The advent of next-generation sequencing has provided scientists in many research areas with a tool to understand genomic information in a cost-effective manner. The continual decreases in the cost of sequencing have made this an attractive readout that provides not only genotype information, but contextual information of the areas surrounding target genomic loci, that can increase the types of available genomic markers, and also lead to the discovery of new markers. The efficiency of next-generation sequencing has shifted the throughput challenge further upstream, necessitating improved methods for preparing samples for sequencing analysis, in the most efficient way possible.

The NEBNext Direct Genotyping Solution was developed to address the specific needs of high-throughput, targeted genotyping required for breeding applications. Developed in collaboration with industrial crop breeders, the method employs high levels of sample multiplexing along with highly efficient, capture-based enrichment of hundreds to thousands of genomic loci, to enable next-generation sequencing to be fully leveraged as a means for high-throughput, cost-effective genotyping.

In order to demonstrate the efficacy of the approach, a panel of 2309 genomic markers was developed to target SNPs from the Solanaceae Coordinated Agricultural Project (SolCAP) database (2). This panel was assayed against extracted DNA from 96 samples of the tomato crop, *Solanum lycopersicum*, and processed using the NEBNext Genotyping Solution before Illumina® sequencing.

Key features aiding the efficiency of the NEBNext Direct Genotyping Solution include the consolidation of DNA fragmentation with end repair, 5′ phosphorylation, and dA-tailing into a single enzymatic step. This is immediately followed by ligation of the Sample Index, which contains a 5′ adaptor to barcode the samples. These two steps represent the only workflow steps where samples are processed individually, as sample pooling immediately follows. By pooling samples (up to 96) prior to capture-based enrichment, the processing steps are significantly reduced, and there is a vast reduction in laboratory consumables required. The 5′ adaptor also contains a 12 base pair, random sequence known as a Unique Molecular Identifier, or UMI. The UMI is used to individually index each library molecule and is used in data analysis to identify duplicate molecules that are generated during the downstream amplification processes, as well as aiding more accurate genotype calls.

The bait hybridization step hybridizes both DNA strands using synthetic, biotinylated oligonucleotides designed against all 2,309 genomic regions harboring the loci of interest for all 96 samples, capturing a total of over 220,000 data points in a single enrichment reaction. These captured molecules are subsequently fully converted into a next-generation sequencing library, during which specificity enhancing enzymatic treatments are performed, and a second, pool-specific barcode is added to the 3′ end of molecules. This dual-indexing strategy enables further pooling prior to sequencing, maximizing the output of Illumina® sequencing.

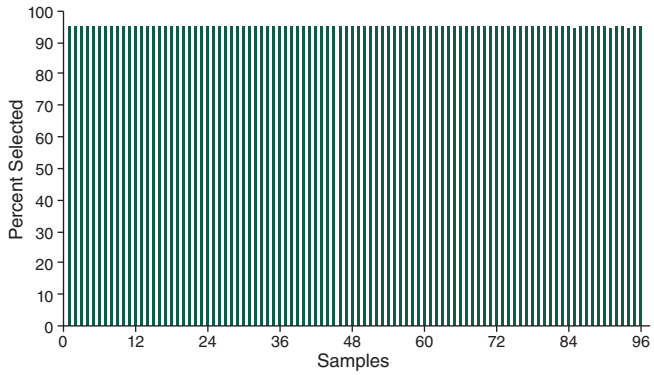
In order to assess performance across 96 samples, sequencing data was processed and aligned to the reference genome sequence, and key metrics were obtained. Analysis of the percent selected across each of the 96 samples demonstrates the method's ability to efficiently enrich molecules containing the target SNP markers, with consistent values across samples showing > 95% of sequencing data mapping to the defined targets (Figure 2).

Further analysis into the ability to confidently determine the genotype of the 2,309 SNPs, demonstrates the ability of the method to produce consistent sequencing coverage at depths sufficient to assess the presence or absence of SNP markers across the 96 pooled samples (Figure 3).

A closer examination of the specific performance across targets included in the panel within a single sample shows coverage across the highest and lowest performing targeted regions within.

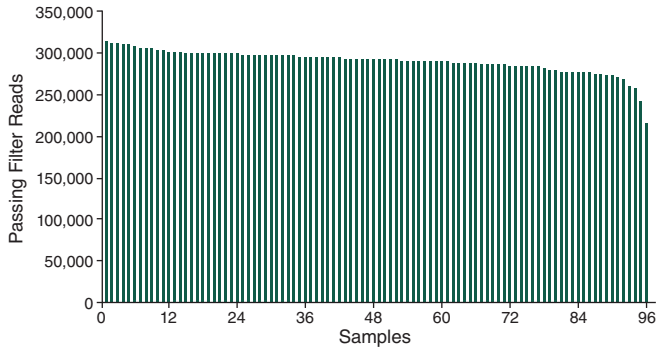
These data suggest that by using the NEBNext Direct Genotyping Solution, high-throughput, massively parallel enrichment of genomic loci can be achieved in an efficient manner upfront of next generation sequencing. The combined efficiency of a novel sample-preparation strategy and continued advances in next generation sequencing present a tractable solution for genotyping hundreds to thousands of genomic loci that can be employed to accelerate plant breeding programs aimed at production of new crop variants that can overcome the challenges our global population growth will continue to present.

Figure 2:
Percent selected across 96 pooled samples



The percent of passing filter reads mapping to targeted regions demonstrates high specificity across 96 multiplexed samples using the NEBNext Direct Genotyping Solution. 25 ng of purified tomato DNA was used as input for each sample. Samples were index-tagged and pooled prior to hybridization and libraries were sequenced on an Illumina MiSeq with 20 cycles of Read 1 to sequence the 12 base UMI and 8 base sample index, and 75 cycles of Read 2 to sequence the targets.

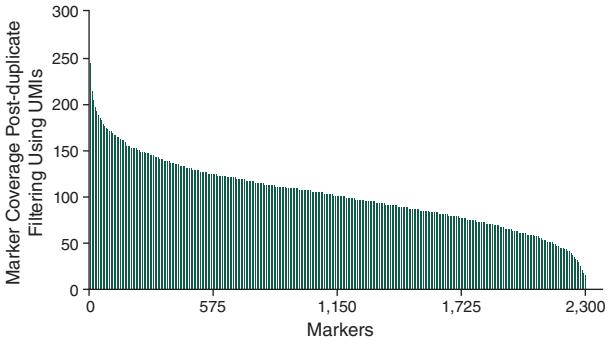
Figure 3:
Mean SNP coverage across 96 pooled samples



Mean SNP coverage of 2,309 SolCAP markers across 96 samples. 25 ng of purified tomato DNA was used for each sample. Samples were index-tagged and pooled prior to hybridization and Libraries were sequenced on an Illumina MiSeq with 20 cycles of Read 1 to sequence the 12 base UMI and 8 base sample index, and 75 cycles of Read 2 to sequence the targets.

continued on page 4...

Figure 4:
Sequencing coverage depth observed across 2,309 marker loci within a single sample



Histogram of coverage across each of the 2,309 SolCAP markers demonstrates evenness of enrichment across targets and coverage levels sufficient for genotyping calls. These data represent enrichment of a single tomato sample pooled with 95 others prior to hybridization. 25 ng of purified tomato DNA was used for each sample. Samples were index-tagged and pooled prior to hybridization and libraries were sequenced on an Illumina MiSeq with 20 cycles of Read 1 to sequence the 12 base UMI and 8 base sample index, and 75 cycles of Read 2 to sequence the targets.

References

1. United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population 2019: Wall Chart (ST/ESA/ SER.A/434).
2. Solanaceae Coordinated Agricultural Project (SolCAP). 2011. Tomato intervarietal TA496 vs. Heinz1706 SNPs. Michigan State University, Dept. of Plant and Soil Science, East Lansing, MI. http://solcap.msu.edu
3. http://broadinstitute.github.io/picard
4. https://solgenomics.net/help/index.pl
5. Li H. (2013) Aligning sequence reads, clone sequences and assem-bly contigs with BWA-MEM. arXiv:1303.3997v2 [q-bio.GN]
6. Fulcrum Genomics, https://github.com/fulcrumgenomics/fgbio

In the NEB Expressions feature article "Over 40 years in protein expression and purification – a historical perspective" (Issue III, 2019) Har Gobind Khorana was accidentally omitted from the timeline "Advances in DNA Understanding were Foundational for Protein Overexpression" (Figure 1, page 2). Khorana published seminal work on cracking the genetic code. In 1968, he shared the Nobel Prize with Marshall Nirenberg and Robert Holley for their interpretation of the genetic code and its function in protein synthesis. This figure has been corrected in both web and PDF versions of the article available on www.neb.com. We apologize for this oversight.

NEBNext Direct® Genotyping Solution – High throughput targeted genotyping for Illumina sequencing

The NEBNext Direct Genotyping Solution combines highly multiplexed, capture-based enrichment with maximum efficiency next generation sequencing to deliver cost-effective, high throughput genotyping for a wide variety of applications. Applicable for ranges spanning 100-5,000 markers, pre-capture multiplexing of up to 96 samples combined with dual-indexed sequencing allows over 3.8 million genotypes in a single Illumina sequencing run.

Sample Indexing and Multiplexing

With 96 pre-capture sample indexes and 8 post-capture pool indexes available, up to 96 samples can be combined for a single capture, and 768 samples can be pooled into a single Illumina sequencing run. Additionally, a 12 bp Unique Molecular Identifier (UMI) is added prior to sample pooling and enrichment, allowing for accurate assessment of input coverage and improving the accuracy of genotyping calls. Finally, sequencing cycle numbers are optimized to sequence only the necessary target region, indexes and UMI required for marker genotyping. The NEBNext Direct Genotyping Solution is compatible with the full range of Illumina sequencers.

Optimized Bait Design

The NEBNext Direct Genotyping Solution employs a purpose-built bait designer that has been optimized to provide both highly specific capture of target loci and maximized sequencer efficiency. By designing baits independently to each target DNA strand with proximity to the target loci, shorter sequence reads can be utilized for genotyping calls. Further, by removing upstream off-target sequence, individual baits can be unambiguously linked to their corresponding sequencing read, presenting opportunities for bait optimization on a per target level and resulting in extremely uniform coverage levels across markers.

For ordering information and to download the full brochure, visit
www.neb.com/E9500

NGS-based targeted genotyping for a wide range of applications

PLANT

Marker Assisted Selection / Breeding

Quantitative Trait Locus (QTL) Screening

ANIMAL

Mouse Genotyping

Livestock Breeding

HUMAN

Biobanking

NGS Sample Tracking

Meet our new ClimaCell® Shipper

NEB has always placed environmental stewardship as one of our highest priorities – one of our goals has been to continuously improve our business processes in order to minimize our impact on the environment.

Cold chain shipping has long been a challenge from an environmental standpoint. Many of our products require shipment on ice or dry ice, and maintaining proper shipping temperature conditions is critical, particularly when shipping long distances or to warmer climates. Expanded polystyrene (EPS), commonly referred to as Styrofoam®, has always been the gold standard – it is light, durable, and well-known for its insulative properties. Unfortunately, EPS is difficult to recycle and often makes its way to landfills. To address this, we have maintained a shipping box recycling program for over 40 years – customers simply use the free return label and send their shipping box back to NEB for re-use. In the meantime, we continued to look for alternative, more sustainable solutions.

We are excited to announce that beginning this Spring, NEB will be transitioning from EPS coolers to a more sustainable solution. The ClimaCell® cooler, developed by TemperPack®, maintains the cold shipping temperature requirements needed to ship NEB products, but is 100% recyclable and requires 94% less energy to manufacture.

Q. How is the ClimaCell cooler made?

A. *The ClimaCell insert is paper-based, and contains thousands of air pockets to prevent heat from moving into the package. It fits tightly inside the shipping box, which is made from corrugated cardboard. Both are 100% curbside recyclable, along with other corrugated cardboard.*

Q. Does the ClimaCell cooler keep product as cold as EPS coolers?

A. *Yes, the ClimaCell insert was specifically designed for thermal protection. It was tested under varying shipping conditions and maintains proper temperature requirements for the delivery of NEB products. Temperature stability data can be found in our Technical Note located at www.neb.com/ShippingBox.*

Q. When will customers begin to receive the ClimaCell cooler?

A. *Customers will start receiving the new shipper in Spring, 2020. There may be a period of time where you receive both ClimaCell and EPS coolers, until supplies are depleted.*

Q. Will NEB still be using EPS coolers?

A. *Until we identify additional solutions, NEB will still rely on EPS coolers for a portion of our products, including those that require shipment on dry ice (e.g. competent cells), as well as larger shipments.*

Q. Will NEB still maintain its shipping box recycling program?

A. *Yes, we will maintain the program for any EPS coolers in circulation. If you receive an EPS cooler, you are welcome to use the free return label to send your shipping box back to NEB, where we will arrange for its proper recycling.*

Q. Where can a customer go to learn more about the ClimaCell cooler?

A. *More information on the ClimaCell cooler can be found at* ***www.neb.com/ShippingBox***

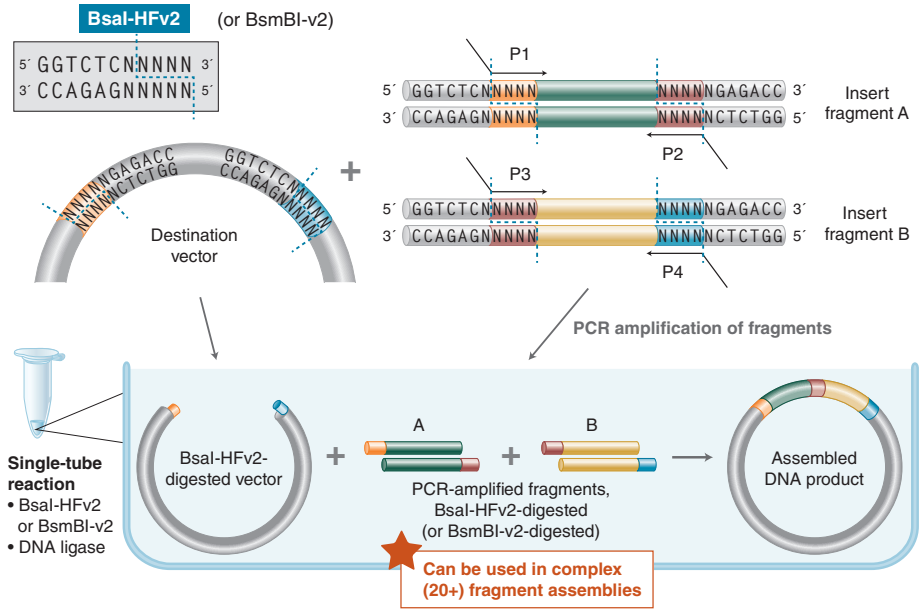
Push the Limits of Golden Gate Assembly

20+ fragment assembly now achievable with high efficiency and accuracy

The efficient and seamless assembly of DNA fragments, commonly referred to as Golden Gate Assembly (1,2), has its origins in 1996, when for the first time it was shown that multiple inserts could be assembled into a vector backbone using only the sequential (3) or simultaneous (4) activities of a single Type IIS restriction enzyme and T4 DNA Ligase. Golden Gate Assembly and its derivative methods exploit the ability of Type IIS restriction endonucleases (REases) to cleave DNA outside of the recognition sequence. The inserts and cloning vectors are designed to place the Type IIS recognition site distal to the cleavage site, such that the Type IIS REase can remove the recognition sequence from the assembly.

With constant advances in both the development of new enzymes and research on maximizing enzyme functionality (e.g., ligase fidelity), NEB has become the industry leader in pushing the limits of Golden Gate Assembly and related methods, such as MoClo, GoldenBraid, Mobius Assembly and Loop Assembly. We have all the products and information you need to perform complex assemblies (20+ fragments) with high efficiencies, >90% accuracy and low backgrounds.

Golden Gate Assembly workflow for both simple and complex assemblies



In its simplest form, Golden Gate Assembly requires a Type IIS recognition site, in this case, BsaI-HFv2 (GGTCTC), added to both ends of a dsDNA fragment. After digestion, these sites are left behind, with each fragment bearing the designed 4-base overhangs that direct the assembly.

Advantages:

- Clone seamlessly, with no scars remaining after assembly
- Perform single insert cloning in just 5 minutes using our fast protocols
- Generate libraries with high efficiencies
- Assemble multiple fragments (2-20+) in order, in a single reaction
- Experience high efficiency, even with regions of high GC content and areas of repeats
- Use with a broad range of fragment sizes (<100 bp to >15 kb)

Tools to help with designing your reactions

- For help designing primers, try the new NEB Golden Gate Assembly Tool at **GoldenGate.neb.com**
- Try our ligase fidelity tools for the design of high-fidelity Golden Gate assemblies at **neb.com/research/nebeta-tools**
  - **Ligase Fidelity Viewer™ (v2)** – Visualize overhang ligation preferences
  - **GetSet™** – Predict high-fidelity junction sets
  - **SplitSet™** – Split DNA sequence for scarless high-fidelity assembly

USAGE GUIDE:

Expanded "assembly standards" for MoClo, GoldenBraid 2.0 and other modular Golden Gate Assembly methods

What is Modular Cloning?

Modular Cloning, better known as MoClo, is a Type IIS assembly method commonly used by synthetic biologists, including those from the plant and AgBio community, for the assembly of multiple biological parts to engineer new biological systems. MoClo methods provide efficient and seamless assembly and allow users to generate and reuse "parts for assemblies" so that laboratories can share assembly-ready sequences.

MoClo methods depend on Type IIS restriction enzymes (such as BsaI-HFv2, BbsI/BbsI-HF and BsmBI-v2/Esp3I) that leave 4-base overhangs. NEB has extensively studied the fidelity of ligation for all possible 4-base overhangs. With this information, we are able to estimate the fidelity of any set of junctions (fusion sites).

Like Golden Gate Assembly, MoClo (and GoldenBraid 2.0) uses 3 levels of successive assembly. The community has agreed upon a set of common standard overhangs for each level as indicated below, along with predicted fidelity:

MoClo Standardized Assembly Overhangs

- Level 0 (Basic parts): ACAT, TTGT (94% fidelity)
- Level 1 (Transcriptional units): GGAG, TACT, CCAT, AATG, AGGT, TTCG, GCTT, GGTA, CGCT (93% fidelity)
- Level 2 (Multigene constructs): TGCC, GCAA, ACTA, TTAC, CAGA, TGTG, GAGC, GGGA (95% fidelity)

Utilizing the gathered ligase fidelity information, NEB has expanded each level of assembly overhangs without sacrificing fidelity. The expanded sets are described at right.

Expanded MoClo Standardized Assembly Overhangs

- Level 0 (Basic parts): ACAT, TTGT, ACTG, GCTA, CCCA, AATA, ATTC, GTGA, CGCC, AAGA, AAAC, AACG, CTGC. GACC, AAGA, AAAC, AACG, CTGC, GACC, CTAA, ACCC, TACA, GGAA, CAAG, AGAG (93% fidelity)
- Level 1 (Transcriptional units): GGAG, TACT, CCAT, AATG, AGGT, TTCG, GCTT, GGTA, CGCT, GAAA, TCAA, ATAA, GCGA, CGGC, GTCA, AACA, AAAT, GCAC, CTTA, TCCA (92% fidelity)
- Level 2 (Multigene constructs): TGCC, GCAA, ACTA, TTAC, CAGA, TGTG, GAGC, GGGA, CGTA, CTTC, ATCC, ATAG, CCAG, AATC, ACCG, AAAA, AGAC, AGGG, TGAA, ATGA (95% fidelity)

Advances in Ligase Fidelity

Research at NEB has led to an increased understanding of ligase fidelity, including the development of a comprehensive method for profiling end-joining ligation fidelity in order to predict which overhangs have improved fidelity (5). This research has enabled complex fragment assemblies with high efficiency and >90% accuracy. More information can be found in the NEB publication, Comprehensive Profiling of Four Base Overhang Ligation Fidelity by T4 DNA Ligase and Application to DNA Assembly (5) or in our webinar, "Fidelity and Bias in End-Joining Ligation: Enabling complex, multi-fragment Golden Gate DNA Assembly".

Both of these are accessible at **www.neb.com/GoldenGate**.

References

1. Engler, C., Kandzia, R., and Marillonnet, S. (2008) PLoS ONE 3, e3647.
2. Engler, C., et al. (2009) PLoS ONE 4, e5553.
3. Lee, J.H. et al, (1996) Genetic Analysis: Biomolecular Engineering, 13: 139-145.
4. Padgett, K.A. and Sorge, J.A. (1996) Gene, 168: 31-35.
5. Potapov, V. et. al. (2018) ACS Synth. Biol. 7,1, 2665-2674.

NEW GOLDEN GATE ASSEMBLY PRODUCTS:

NEB Golden Gate Assembly Kit (BsmBI-v2)

The absence of internal sites in a sequence determines the choice of which Type IIS restriction enzyme to drive the assembly. For your convenience, NEB now offers two kits for Golden Gate Assembly featuring BsaI-HF-v2 or BsmBI-v2. Both kits incorporate digestion followed by ligation with T4 DNA Ligase into a single reaction, and can be used to assemble 2-20+ fragments in a single step.

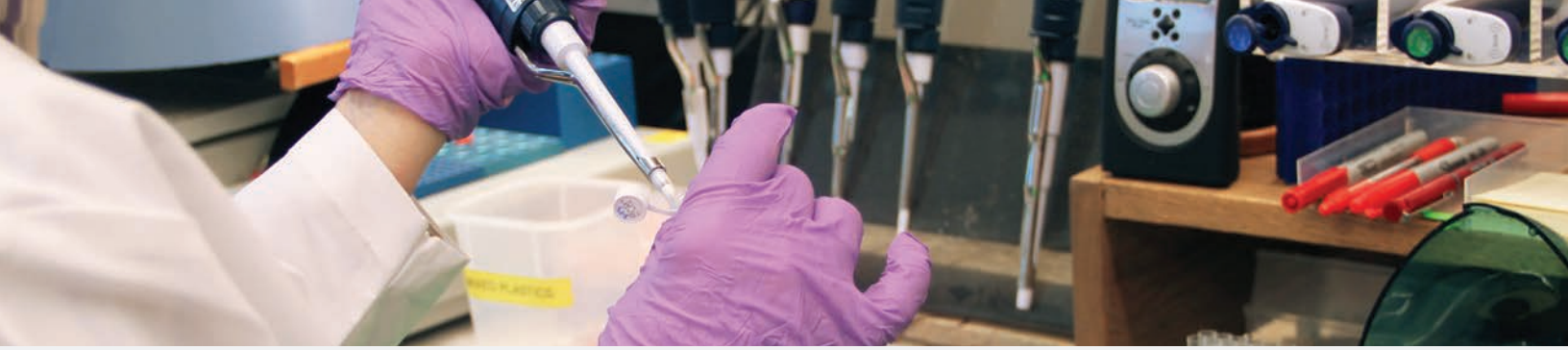
BsmBI-v2

NEB also offers more Type IIS restriction enzymes that any other supplier, many of which are used in Golden Gate assembly. These enzymes recognize asymmetric DNA sequences and cleave outside of their recognition sequence. BsmBI-v2 has been optimized for Golden Gate Assembly and replaces BsmBI.

For more information, visit **www.neb.com/GoldenGate**.

Ordering information:

Product	NEB #	Size
NEB Golden Gate Assembly Kit (BsmBI-v2)	E1602S/L	20/100 rxns
BsmBI-v2	R0739S/L	200/1,000 units



Eleven Tips For Optimizing Your Golden Gate Assembly Reactions

Looking to assemble multiple DNA fragments in a single reaction? Here are some tips to keep in mind when planning your Golden Gate Assembly experiment.

1 Check your sequences

Always check your assembly sequences for internal sites before choosing which Type IIS restriction endonuclease to use for your assembly. While single insert Golden Gate assembly has such high efficiencies of assembly that the desired product is obtainable regardless of the presence of an internal site, this is not true for assemblies with multiple inserts. Options include choosing a different Type IIS restriction enzyme to direct your assembly, or eliminating internal sites through mutagenesis. The Q5® Site-Directed Mutagenesis Kit (NEB #E0554S) and the NEB web tool NEBaseChanger work well for this purpose. Alternately, a junction point can be created at the internal site's recognition sequence.

2 Orient your primers

When designing PCR primers to introduce Type IIS restriction enzyme sites, either for amplicon insert assembly or as an intermediate for pre-cloning the insert, remember that the recognition sites should always face inwards towards your DNA to be assembled. Consult the Golden Gate Assembly Kit manuals for further information regarding the placement and orientation of the sites.

3 Choose the right plasmid

Consider switching to the pGGAselect Destination Plasmid for your Golden Gate assembly. This versatile new destination construct is included in all Golden Gate Assembly kits and can be used for Bsa-HFv2, BsmBI-v2 or BbsI directed assemblies. It also features T7 and SP6 promoter sequences flanking the assembly site, and has no internal BsaI, BsmBI or BbsI sites. The pGGAselect plasmid can also be transformed into any *E. coli* strain compatible with pUC19 for producing your own plasmid preparation if so desired.

4 Choose the right buffer

T4 DNA Ligase Buffer works well for Golden Gate Assembly with both BsaI-HFv2 and BsmBI-v2. However, alternate buffers would be NEBuffer 1.1 for Bsa-HFv2 and NEBuffer 2.1 for BsmBI-v2, as long as supplemented with 1 mM ATP and 5-10 mM DTT.

5 Increase your complex assembly efficiency by increasing the Golden Gate cycling levels

T4 DNA Ligase, BsaI-HFv2 and BsmBI-v2 are very stable and survive extended cycling protocols; an easy way to increase assembly efficiencies without sacrificing fidelity is to increase the total cycles from 30 to 45-65, even when using long (5-minute) segments for the temperature steps.

6 Make sure your plasmid prep is RNA-free

For pre-cloned inserts/modules, make sure your plasmid prep is free of RNA to avoid an over-estimation of your plasmid concentrations.

7 Avoid primer dimers

For amplicon inserts/modules, make sure your PCR amplicon is a specific product and contains no primer dimers. Primer dimers, with Type IIS restriction endonuclease sites (introduced in the primers used for the PCR reactions), would be active in the assembly reaction and result in mis-assemblies.

8 Avoid PCR-induced errors

Do not over-cycle and use a proofreading high fidelity DNA polymerase, such as Q5® DNA High-Fidelity Polymerase.

9 Decrease insert amount for complex assemblies

For complex assemblies involving >10 fragments, pre-cloned insert/modules levels can be decreased from 75 to 50 ng each without significantly decreasing the efficiency of assembly.

10 Carefully design EVERY insert's overhang

An assembly is only as good as its weakest junction. Research at NEB has led to an increased understanding of ligase fidelity, including the development of a comprehensive method for profiling end-joining ligation fidelity in order to predict which overhangs will result in improved accuracy. This ligase fidelity information can be used in conjunction with the NEB Golden Gate Assembly Kits (BsaI-HFv2 or BsmBI-v2) to achieve high efficiency and accurate complex assemblies. Please use the free NEB Golden Gate Assembly Tool to design primers for your Golden Gate Assembly reactions, predict overhang fidelity or find optimal Golden Gate junctions for long sequences. When working with complex assemblies (>20 fragments), refer to the ligase fidelity tools on the NEBeta Tools site.

11 Check for a sequence error if your assembly becomes non-functional

Be aware that occasionally a pre-cloned insert/module can become corrupted by an error during propagation in E. coli, usually a frameshift due to slippage in a run of a single base (e.g., AAAA) by the *E. coli* DNA Polymerase. This should be suspected if previously functional assembly components suddenly become nonfunctional.

A High Performance *E. coli* Cell Lysate-Based System for *in vitro* Protein Synthesis

Paula Magnelli, Ph.D., Haruichi Asahara, Julie Beaulieu, Jim Samuelson, Ph.D., Stephen Shi, Ph.D., New England Biolabs, Inc.

Introduction

Bacterial cell extract-based *in vitro* protein synthesis systems have been widely used for an array of applications. These systems offer a number of advantages; for example, the tight coupling of translation and transcription in prokaryotic cells still remains in the lysate, rendering the protein synthesis process exceptionally efficient. Additionally, genetic manipulation can be performed on the strain in which the lysates are made to enhance its ability to generate high yields of protein. Finally, lysates can be manufactured at a larger scale, as compared to reconstituted systems. Lysate-based protein synthesis systems are conceptually simple and relatively less expensive, and have therefore been home-brewed in many laboratories over the past decades, as well as supplied by several commercial sources. These systems, however, exhibit varying levels of performance, ease of use, and often do not produce consistent results across broad sizes and types of proteins.

The NEBExpress® Cell-free *E. coli* Protein Synthesis System (NEB #E5360) was developed using several strategies to enhance performance, ease of use, and ensure robustness. These include the use of an *E. coli* strain genetically engineered to maximize the stability of template DNA and RNA and the protein products, a highly optimized reaction buffer, and a stringent biomanufacturing process.

In this technical note, we examine how the NEBExpress Cell-free *E. coli* Protein Synthesis System performs in several frequently encountered applications, and demonstrate this system's high performance and versatility.

Results

Efficiency on different templates

Traditionally, cell-free protein synthesis systems use DNA plasmids as templates, because they are readily available and easily prepared. Recently, there has been an increasing need for linear DNA templates in applications such as high throughput screening, where linear DNA can be used directly from amplification or directed mutagenesis. Additionally, there are situations where protein synthesis from mRNA templates is desired.

For these reasons, a lysate system that can utilize plasmid DNA, linear DNA or mRNA as a template is desirable. When equimolar linear DNA templates were introduced in the NEBExpress Cell-free *E. coli* Protein Synthesis System, the target vGFP protein was produced, albeit at a reduced yield. However, yield was almost doubled by adding 1.5 µg of NEBExpress GamS Nuclease Inhibitor, which is known to stabilize linear DNA templates in *in vitro* protein synthesis reactions (data available in downloadable PDF). Protein yield was also enhanced by the addition of more linear DNA. With a combination of GamS and an increased amount of linear DNA, protein synthesis can be achieved at almost the same level as the plasmid DNA templates (data available in downloadable PDF).

Effects of target size and temperature

The ability to use one kit for synthesis of a wide range of proteins can have a big impact on experimental efficiency. The NEBExpress Cell-free *E. coli* Protein Synthesis System has been shown to generate high yields of proteins from a wide range of targets (16.7–230 kDa) (Figure 1). Further, the enzymatic activity of β-galactosidase S. (230 kDa) was confirmed via colorimetric assay using 2-nitrophenyl β-D-galactopyranoside as a substrate (data available in downloadable PDF). To our knowledge, this the largest active protein synthesized using a cell-free system.

Moreover, protein synthesis using the NEBExpress Cell-free *E. coli* Protein Synthesis System can be carried out at different temperatures. β-galactosidase S showed highest yield at the optimal synthesis temperature of 28°C (data available in downloadable PDF). Similar to *in vivo* protein expression, the relative rate of protein translation and folding can determine how much soluble form of the protein can be obtained. These variables can be easily tuned by incubation at different temperatures in *in vitro* reactions.

Disulfide bond formation

Disulfide bond formation is a challenging issue for protein expression in bacteria. Disulfide bond enhancers can be added to *in vitro* protein synthesis reactions to help correct disulfide bond formation in target proteins. In the absence of such enhancers, chitinase from *Plasmodium* that was synthesized using the NEBExpress Cell-free *E. coli* Protein Synthesis System displayed minimal activity. However, in the presence of PURExpress® Disulfide Bond Enhancer (NEB #E6820), there was a significant increase in the chitinase activity (data available in downloadable PDF).

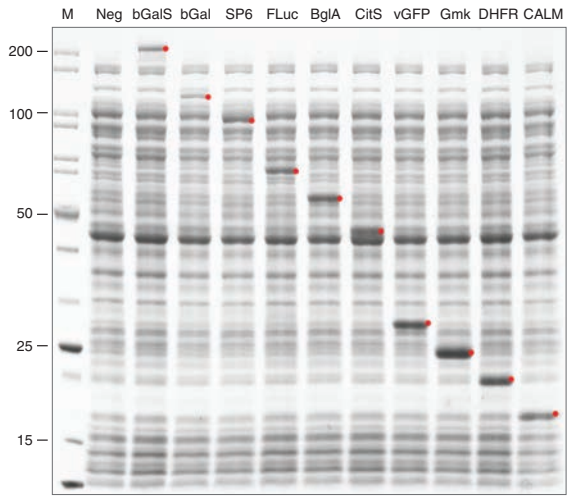
Sustained protein synthesis

The NEBExpress Cell-free *E. coli* Protein Synthesis System can produce high yields of target protein in approximately three hours. However, given adequate aeration and agitation, the reaction can continue for more than 10 hours, producing greater than 1 mg protein/ml (data available in downloadable PDF). The longer incubation time is particularly useful when it is necessary to carry out the protein synthesis reaction at milder temperatures.

Co-expression of proteins

It is possible to co-express multiple protein targets in a single reaction. Four DNA targets were added to a single reaction using the NEBExpress Cell-free *E. coli* Protein Synthesis System. All four targets were produced, in clear bands, although with a slightly reduced yield compared to experiments where only one plasmid is introduced (data available in downloadable PDF). This demonstrates the potential for the NEBExpress system to make multiple proteins for the purpose of assembling a protein complex, or engineer a metabolic pathway.

Figure 1: **The NEBExpress Cell-Free *E. coli* Protein Synthesis System can generate high yields of proteins from a wide range of targets**



CFPS of ten different targets, indicated by red dots, and a no DNA control (neg). Target proteins varied in size between 16.7 and 230 kDa.

Conclusion

This high-performing, versatile, and robust cell-free protein synthesis system offers the ability to rapidly produce a large number of proteins for further characterization. The NEBExpress Cell-free *E. coli* Protein Synthesis System was developed using genetically engineered *E. coli* strains, an optimized reaction buffer, and stringent manufacturing practices, and is capable of synthesizing proteins as large as 230 kDa. The yield, under optimal conditions, can reach milligrams per milliliter, with protein synthesis continuing for up to 24 hours. The versatility of this system makes it ideal for a variety of applications, including high throughput protein screening and engineering as well as synthetic biology.

The full technical note including performance data can be downloaded at ***www.neb.com/e5360****.*

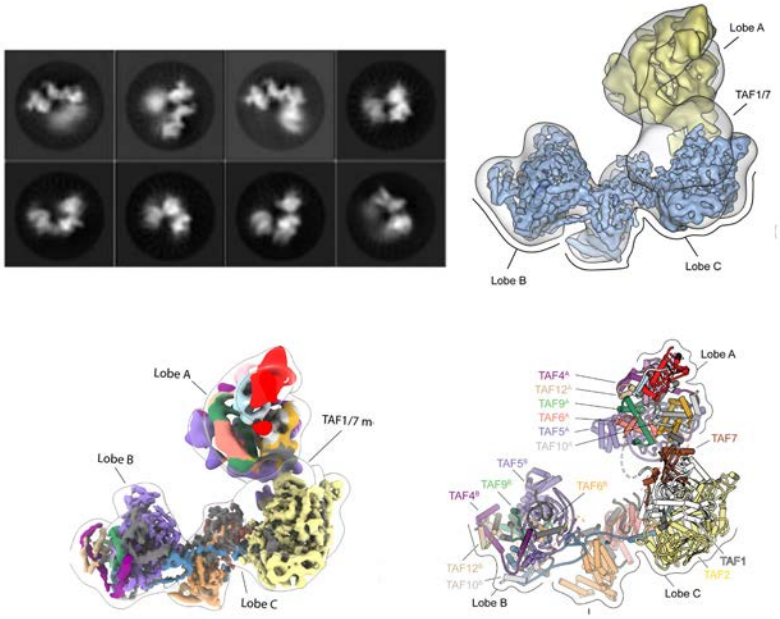
Honoring NEB's founder with lectures from pioneering researchers

By Greg Lohman, Ph.D., New England Biolabs

On Wednesday November 20th, New England Biolabs hosted the **Inaugural Donald G. Comb Honorary Lecture**, the first in an annual series named in honor of Don Comb, NEB's founder and first CEO. When Dr. Comb founded NEB he brought a passion for research and a vision for the company that would integrate production of reagents for molecular biology with basic research. As a result, NEB researchers have published over 1,200 publications in peer-reviewed journals, in fields ranging from biochemistry to genomics to parasitology. The annual Donald G. Comb Honorary Lecture series aims to feature speakers who have pioneered basic research with substantial impact on molecular biology and related fields, especially those whose work has significant impact on problems affecting the world at large. Further, selected speakers will embody the values of NEB, including a passion for science, innovative thinking, environmental stewardship, mentorship, genuineness and humility.

Our first lecturer was Dr. Eva Nogales, Howard Hughes Medical Institute Investigator and Professor of Biochemistry, Biophysics and Structural Biology at the University of California, Berkeley. Dr. Nogales is a pioneer in the application of cryogenic electron microscopy (Cryo-EM) to research in molecular biology. Cryo-EM is an imaging technique that bridges the gap between molecules and cells, allowing visualization of large protein complexes too big for crystallography but too small for light microscopy. Cryo-EM enables imaging of these large complexes by rapid freezing and evaporation of a thin layer of buffered solution, capturing the proteins still folded and in their functional arrangements with their protein and nucleic acid partners. Images of these macromolecular complexes are combined to generate a three-dimensional picture of the complex. Combining Cryo-EM structures with high-resolution structures obtained by other techniques and biochemical information allows the researchers to create a dynamic picture of these complexes in action with unprecedented levels of detail.

This technique allows the Nogales lab to study the macromolecular machines of the cell – such as cytoskeleton assembly and gene regulation – as whole units through direct visualization of the structure of not only individual proteins, but also dynamic multiprotein complexes in action. New insights provided by her lab have led to a deeper understanding of the large molecular machines that power the "central dogma" of DNA replication, transcription and translation. These studies have enabled an understanding of how these machines function and how they can fail — from how antibiotics block the cellular protein factory to how the DNA replication machinery can introduce mistakes that result in cancer or other diseases.



Structure of human transcription factor IID (TFIID), reconstructed from Cryo-EM images of the complex (top left). TFIID is a key component of the transcription pre-initiation complex, required for proper loading of RNA polymerase II onto core promoter DNA. CryoEM, in combination with crosslinking mass spectrometry and biochemical data, allowed the structures of this ~1.3 MDa transcription factor to be determined in biologically relevant conformations. Images adapted from Science. 2018 Dec 21; 362(6421): eaau8872 with permission of the authors.

Dr. Nogales, Howard Hughes Medical Institute Investigator and Professor of Biochemistry, Biophysics and Structural Biology at the University of California, Berkeley

Dr. Nogales, pictured here with NEB's founder, Don Comb, also received a piece of custom artwork from local artist, Michael Updike

In addition to the foundational work of her research group, Dr. Nogales has engaged in numerous collaborations with academic, government and industry scientists throughout her career, is well known for her energy and passion for her work, and has been a mentor to many pursuing scientific careers of their own. She was an ideal choice for our Inaugural lecture, which was held at NEB's Ipswich campus. A recording of her lecture is available at **www.neb.com/NogalesLecture**. Dr. Nogales spent the day at NEB meeting with scientists and touring the campus. Dr. Nogales also recorded a podcast interview discussing her work and career, which can be found at (**https://www.neb.com/NogalesPodcast**).

Can looking to the future help preserve a historical fishery against climate change?

Jake Kritzer, Ph.D., Environmental Defense Fund

Note: *This article is modified from a post in the Environmental Defense Fund's "Fisheries for the Future" series published October 10, 2019.*

The New England groundfish fishery

New England's storied groundfish fishery, which targets cod, haddock, a variety of flatfishes, and other bottom-dwelling predators, is among the oldest fisheries in the United States. It was once said that a fisherman could walk across the water on the backs of plentiful cod. The fishery fueled the regional economy following European settlement, and created a rich maritime heritage that continues today. That heritage first belonged to Native American peoples who long pre-dated colonization.

"Preserving this iconic fishery as well as the economy and culture it has subsequently built will require looking forward to an ocean evolving under a changing climate."

Fish on the move – forward-looking management of habitat and prey

The Gulf of Maine and Georges Bank, the complex basin and shallow underwater plateau that have been the foundation of our regional fisheries, sit at the southwestern edge of the range of many North Atlantic species. Historically, these have been coldwater ecosystems due to the Labrador Current delivering frigid polar waters southward from the Arctic. However, this corner of the Northwest Atlantic is warming faster than almost anywhere else on Earth. Warming waters are causing rapid shifts in the distribution of many species, generally to the north and offshore, seeking preferred water temperatures. Thus, species we normally associate with the Mid-Atlantic — black sea bass, summer flounder, striped bass, and others — are expected to become more abundant in New England as colder water species push northward. Fishermen will see a shifting mix of species as warming progresses, and governance must also change to manage the shifting portfolios.

Climate change is making Atlantic cod recovery especially difficult, but geographical shifts might have some benefits. Cod have become concentrated in a small pocket in the western Gulf of Maine bounded by Cape Ann and Cape Cod. Elsewhere, overfishing has caused near-complete localized extinction. Although warming waters are already decreasing the productivity of cod, spreading the stock more widely across the Gulf of Maine could increase resilience relative to today's much more restricted distribution by hedging bets against localized declines. Important efforts to restore coastal prey fishes that cod feed on, especially sea-run herring, are helping to give cod a chance where they have been lost.

If cod return to those areas, they will need time to re-establish. That process will be more complicated in a changing ecosystem, for the nature of seafloor habitats, water temperatures, surrounding fish and invertebrate species, and other factors will be different from what cod once knew. It was therefore with laudable foresight that the New England Fishery Management Council and National Marine Fisheries Service created a fishery closed area offshore from Penobscot Bay. That refuge is helping protect important habitats and can enable fledgling spawning groups to grow and possibly serve as a source of replenishment to areas elsewhere in the Gulf of Maine.

Climate change and uncertainty — responsive fishing policies for evolving conditions

Of the 20 stocks in the groundfish fishery, most live primarily away from shore. But one unique species, the winter or blackback flounder, historically moved inshore to spawning and nursery grounds in estuaries and salt ponds in the wintertime. Key habitats in those areas, including salt marshes, eelgrass beds and oyster reefs, are especially susceptible to climate change as sea levels rise, waters warm, and storms intensify. These habitat changes, among other impacts, mean that winter flounder are expected to suffer especially strong declines in productivity.

However, some winter flounder are known to spawn offshore as well. This means the stock might have the ability to counteract reduced inshore spawning success by capitalizing on deeper and colder waters. The effect of this life history diversity is but one of many scientific uncertainties we must confront, among other uncertainties related to climate change and incomplete accounting of how many fish are caught. Untangling these uncertainties and applying our findings to forward-looking management strategies will not be easy, but there are steps we can take in anticipation of changes that will come.

A central element of any fishery management strategy is a harvest control rule, or HCR, which determines how many fish can be caught based on how many fish are in the water. An HCR is arguably where science most directly confronts policy in fisheries management, as it reveals a great deal about objectives, scientific understanding and uncertainties, and risk tolerance. In many fisheries, the HCR is simply to fish at a fixed but precautionary rate of fishing mortality that strikes a balance between achieving high yields when the stock is large, but not overfishing when the stock is smaller. However, when facing climate-driven declines in productivity that are exacerbated by scientific uncertainty, the HCR must be more responsive.

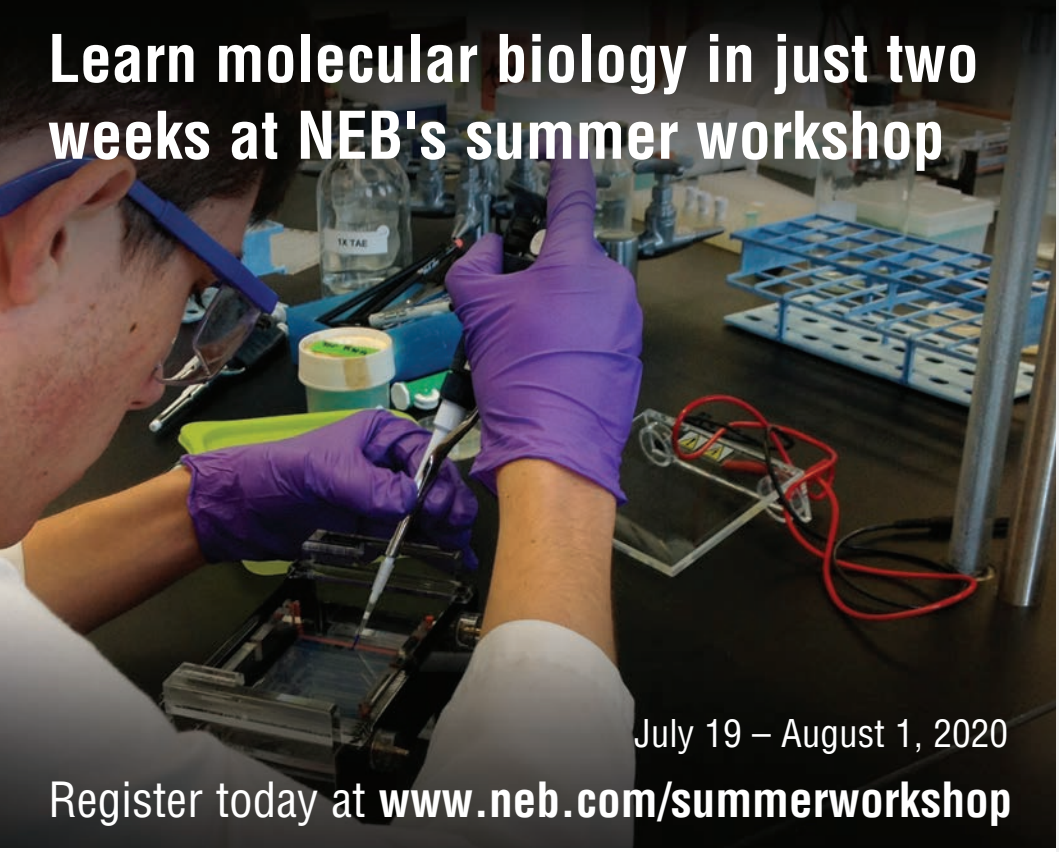
"Even if we do not understand all of the changes taking place, fishing mortality should decrease as we detect declines and can then rise again with evidence of recovery."

Such an approach is not yet used for New England groundfish, but could be adopted more readily than other measures.

New England can have abundant fisheries

For those who call New England home, the groundfish fishery has sculpted our waterfronts, history, folklore and cuisine. It can remain an indelible part of our region, as long as we look to the future while we embrace the past. The ecosystem will function differently as climate change continues to unfold and we must prepare for that future. Strategic use of protected areas and responsive harvest policies, alongside other actions like recovery of prey fish and improved monitoring to track changes, can help us keep pace with a changing ocean and retain vibrant fisheries as part of our regional economy and culture.

Visit **http://blogs.edf.org/edfish/2019/11/21/building-fisheries-for-the-future/** for more blog posts from the multi-part series: Fisheries for the Future.



NEW ENGLAND BioLabs® Inc.
240 County Road, Ipswich, MA 01938-2723

LOCATIONS

USA
New England Biolabs, Inc.
Telephone: (978) 927-5054
Toll Free: (U.S. Orders) 1-800-632-5227
Toll Free: (U.S. Tech) 1-800-632-7799
info@neb.com

Australia
New England Biolabs (Australia) PTY
Telephone: +61 (1800) 934218
info.au@neb.com

Canada
New England Biolabs, Ltd.
Toll Free: 1-800-387-1095
info.ca@neb.com

China
New England Biolabs (Beijing), Ltd.
Telephone: 010-82378265/82378266
info@neb-china.com

France
New England Biolabs France
Telephone : 0800 100 632
info.fr@neb.com

Germany & Austria
New England Biolabs GmbH
Free Call 0800/246 5227 (Germany)
Free Call 00800/246 52277 (Austria)
info.de@neb.com

Japan
New England Biolabs Japan, Inc.
Telephone: +81 (0)3 5669 6191
info.jp@neb.com

Singapore
New England Biolabs Pte. Ltd.
Telephone +65 638 59623
sales.sg@neb.com

United Kingdom
New England Biolabs (UK), Ltd.
Call Free 0800 318486
info.uk@neb.com

For contacts in other countries, please visit
www.neb.com/international-support