Thermo Scientific Open Biosystems
cDNAs and ORFs

- World's largest collection of pre-made cDNAs and ORFs
- Fully sequenced, easy-to-transfer and cost-effective
- Genome-scale and expression-ready collections
The world’s largest and most complete collections of pre-made, full-length, sequence-verified cDNA and open reading frame (ORF) clones for reliable gene over-expression

**Expression-ready**
- Lentiviral ORFs
- Expression-ready cDNAs

**Easy-to-transfer**
- Gateway®-adapted ORFs
- Expression-ready cDNAs

**Genome-scale Coverage**
- Fully sequenced cDNAs
- Gateway-adapted ORFs
- Expression-ready ORFs

**Viral Format**
- Lentiviral ORFs

**Table of contents**

2 Pre-made cDNAs and ORFs as tools for gene expression
3 Thermo Scientific Open Biosystems pre-made cDNA and ORF collections
4 Gene expression solutions and formats
5 Thermo Scientific Open Biosystems cDNA collections
6 The importance of full insert sequencing
7 Thermo Scientific Open Biosystems Open Reading Frame (ORF) collections
8 Thermo Scientific Open Biosystems Precision LentiORFs
9 Pre-defined cDNA and ORF gene families
10 Delivery and viral packaging
11 Gene Search and Online BLAST
12 References
Pre-made cDNAs and ORFs are versatile tools for gene analysis and protein science

cDNA and ORF clones provide the strongest evidence for the structure of RNA transcripts and are indispensable tools for determining gene function.

Thermo Scientific Open Biosystems products include the world’s largest collection of tools for gene analysis. These include fully sequenced cDNA clones for native protein expression, ORF clones that can be transferred to your expression vector of choice, and expression-ready lentiviral ORFs providing the fastest path to your gene analysis results.

Pre-made cDNA and ORF collections facilitate applications across many fields of biological research

cDNAs and ORFs offer experimental solutions for both cell biologists and protein scientists. cDNAs mirror the native RNA sequence including the protein coding region and the regulatory sequences often found in the UTR.

The non-coding regions are removed in the ORF clones and replaced with flexible cloning sites. This facilitates a variety of experiments including: tracking subcellular localization, protein-protein interaction studies, purification for in vitro assays or antibody production.

Full-Length cDNA Collections

To be "full-length", a cDNA clone must contain the complete open reading frame (ORF) and typically includes part or all of the 5’ and 3’ untranslated regions (UTRs).

Full-Length ORF Collections

ORF clones are derived from cDNA clones by removing the untranslated region (UTR) leaving just the protein coding sequence known as the ‘Open Reading Frame’ or ORF. The UTRs are replaced by recombinational cloning sites (att sites) for easy transfer.

cDNA clones provide the user with the protein coding sequence with its native stop codon and regulatory elements often found in the UTRs. By replacing the UTRs and stop codon with Gateway sites, the ORF clones offer increased flexibility when transferring to expression vectors.
Fully sequenced, easy-to-transfer and expression-ready cDNAs and ORFs

**Full-Length cDNA Collections**

Mammalian Gene Collection (MGC)
- High-quality human, mouse and rat clones
- Fully sequenced and verified to contain a complete coding sequence (CDS)

Incyte Clone Collection
- Supplements MGC gene coverage
- Provides additional unique gene coverage and splice variants
- Exclusive to Thermo Scientific

Expression-ready MGC subset
- Fully sequenced
- Robust CMV promoter

**Gateway-adapted ORF Collections**

ORFeome Collaboration Collection
- Fully sequenced
- With or without stop codon

Human ORFeome
- Genome-scale coverage
- Derived from fully sequenced clones from the MGC

Precision LentiORFs
- Expression-ready lentiviral ORFs
- High-titer purified lentiviral particles for hard-to-transfect cell lines

**Save time and money with pre-made cDNAs and ORFs**

Determine your gene of interest

- Gene search
- Full-length clone delivered

Start your experiment in **DAYS**
- Harvest cells
- RNA purification
- cDNA synthesis
- Gene-specific PCR
- Gel sizing and purification
- Ligate and transform
- Screen and fully sequence insert

Start your experiment in **WEEKS**

Steps involved in *de novo* cloning versus ordering pre-made cDNA or ORF clones.
**Which gene expression solution is suitable for your research?**

<table>
<thead>
<tr>
<th>Format</th>
<th>Species</th>
<th>cDNA</th>
<th>ORFs</th>
<th>LentiORFs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacterial</td>
<td>Mouse</td>
<td>MGC</td>
<td>Incyte</td>
<td></td>
</tr>
<tr>
<td>glycerol stock</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mouse</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Arrayed format</td>
<td>Human</td>
<td>Subsets</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Human</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arrayed library</td>
<td>Rat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arrayed format</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene families</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>and pathways</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluorescent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>marker</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Fully sequenced ORFs from the ORFeome collaboration.

---

**Follow the icons to your perfect product solution.**

- **Bacterial glycerol stock:** Bacterial culture transformed with a plasmid vector, grown in presence of glycerol for freezer storage and supplied in single tubes.

- **Arrayed format:** Bacterial glycerol stocks, custom arrayed in 96-well microtiter plates.

- **Arrayed library arrayed format:** Genome-scale libraries of bacterial glycerol stocks arrayed in 96-well microtiter plates.

- **Gene families and pathways:** Bacterial glycerol stocks of cDNA and ORF pre-defined libraries arranged by gene family and/or function and arrayed in 96-well microtiter plates.

- **Fluorescent marker:** TurboGFP as marker of expression.

- **High-titer viral particles:** Transduction-ready, high-titer (10⁸ TU/mL) lentiviral particle format. Turnaround time is 6-8 weeks.

- **Precision LentiORF Starter Kits:** All reagents and protocols necessary to start a gene expression experiment are provided in a single, convenient package.

- **Guaranteed** to express.

---

**Species**

- **Mouse**
- **Human**
- **Rat**

**Format**

- **High-titer viral particles**: When used as part of a LentiORF starter kit according to kit protocols.
Guaranteed to match Genbank published sequence

Genome-scale coverage for Human, Mouse and Rat

At the onset of the ‘Genomics Revolution’ the National Institutes of Health (NIH) spearheaded the creation of genome-scale cDNA collections for the human, mouse and rat. In 2009, this multi-institute project was completed, creating the most extensive, rigorously sequenced collection of mammalian cDNAs (<1 base in 50,000 nucleotide error rate)\textsuperscript{1,2,3}.

- Pre-made genome-scale collections
- Sequence-verified to contain a complete ORF
- Guaranteed sequence*

Exclusive supplemental gene coverage is offered through the Incyte cDNA Collection

Over 27,000 additional fully sequenced clones add depth and breadth to the genome-scale MGC cDNA resource.

- Provides rare gene coverage not included in the MGC collection
- Use our BLAST search to find cDNAs not covered by other collections

To be “full-length”, a cDNA clone must contain the complete open reading frame (ORF) and typically includes part or all of the 5’ and 3’ untranslated regions (UTRs).

A)

<table>
<thead>
<tr>
<th></th>
<th>Human</th>
<th>Mouse</th>
<th>Rat</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGC cDNA clones</td>
<td>30,039</td>
<td>27,457</td>
<td>6,769</td>
</tr>
<tr>
<td>MGC Non-redundant genes</td>
<td>17,421</td>
<td>17,285</td>
<td>6,363</td>
</tr>
</tbody>
</table>

B)

<table>
<thead>
<tr>
<th></th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incyte Collection clones</td>
<td>27,000</td>
</tr>
</tbody>
</table>

Number of clones and unique genes covered by the MGC (A), and the number of additional Incyte clones (B) providing supplemental transcript variants.

*Guaranteed to match Genbank published sequence
The importance of full-insert sequencing

Following full-insert sequencing, approximately 45% of putative full-length clones in the MGC pipeline were eliminated for various reasons including incomplete ORFs, chimeric sequences, and frameshifts.

The wide-spread phenomenon of alternative splicing provides further motivation for full-insert sequencing of cDNA clones. It is estimated that 74% of multi-exon genes are alternatively spliced. Thus, clones that are only end-sequenced have a significant risk of being assigned an incorrect reference accession.

Multiple transcript variants of the DNA methylase gene (DNMT3B) can be expressed from a single cell line. Despite only small transcript size differences these variants have dramatic effects on cells. Expression of DNMT3B has been shown to change gene expression profiles and alter genome methylation states of HEK293 cells to mimic those seen in cancer cell types. Without full-insert sequencing this splice variant would be indistinguishable from the wild-type transcript.

45% Putative full-length clone failure rate in the MGC pipeline

Cumulative numbers for putative and confirmed full-length clones over the history of the MGC (from the MGC weekly status report, August 28, 2005).

Alternative splicing events results in two variant cDNAs for the DNMT3B gene. (a) The wild-type transcript produces the full-length protein. (b) The variant occurs when an aberrant splicing event excludes exon 10 and retains intronic sequence. Exons outside of this area are unaffected, and the two isoforms would be indistinguishable via end sequencing and gel electrophoresis.
The ORFeome Collaboration has created a collection of fully sequenced ORFs that will represent all curated genes in the human genome. These clones are verified to have a full-length ORF and are easily transferable into any Gateway-adapted expression vector. Additional clones for this expanding collection are offered as they become available.

- Fully sequenced
- Easy-to-transfer - Gateway-adapted
- Native and fusion-ready options – with or without stop codon

Visit www.orfeomecollaboration.org for more details.

The Human ORFeome Collection was developed by Dr. Marc Vidal’s lab at the Dana-Farber Cancer Institute. Using the MGC cDNA collection as templates, these ORFs were created by high-throughput PCR providing broad gene coverage. The resulting PCR product was cloned into an entry vector and was confirmed to contain the correct MGC insert via 5' end sequencing read.

- Broad gene coverage
- Easy-to-transfer – Gateway-adapted
- Maximum flexibility – Stop codon removed to allow fusion-tags

ORF clones are derived from cDNA clones by removing the untranslated region (UTR) leaving just the protein coding sequence known as the ‘Open Reading Frame’ or ORF.

Gateway cloning provides fast and efficient gene content portability without the need for PCR or re-sequencing. This allows flexible cloning strategies to quickly incorporate the cloned ORF into many different experimental designs.
The Precision LentiORF Collection is a library of full-length ORFs in an expression-ready lentiviral vector generated from the fully sequenced ORFeome Collaboration Collection. These ready-to-use ORFs are high-quality, versatile and guaranteed* to express.

- Expanded delivery options – use as plasmid or viral particles for hard-to-transfect cells
- Efficient expression - even at low multiplicity of infection (MOI)
- Flexible applications - cloning site for adding a custom fusion-tag

### Vector Element | Utility
--- | ---
CMV Promoter | Efficient expression
TurboGFP(nuc) | Visual tracking of transduction and expression
MultiTag Cloning Site | Convenient cloning site for the addition of a purification or tracking tag
Blasticidinr | Selection for pure populations
SIN LTR 3’ | Self-inactivating long terminal repeat provides additional safety feature

### Controlled ORF expression levels with increasing viral titers.

<table>
<thead>
<tr>
<th>MOI</th>
<th>1</th>
<th>3</th>
<th>10</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viral particles per cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

HEK293T cells transduced at a range of MOIs. Western blot analysis depicts the expected increase in protein expression correlating with an increase in the number of LentiORF viral particles applied to the cell.

### Easily track transduction and expression.

Precision LentiORF transduced HEK293T cells expressing nuclear localized TurboGFP (TurboGFPnuc).
The Druggable Genome set is a focused sub-set of the MGC collection consisting of genes considered to be candidates for therapeutic development. The resulting proteins from these cDNAs are predicted to be targets for small molecules. As a cDNA collection, it can be transferred to an expression vector for over-expression screening or for protein production in screening validations.

**Pre-defined cDNA and ORF gene family sub-libraries** have been assembled using innovative bioinformatic tools to reflect the latest genetic and ontological information from peer-reviewed publications and data-bases. These collections represent the most recent publications, providing the highest confidence for your over-expression experiments.

Please visit www.thermo.com/genefamilies for more information or contact openbiosystems@thermofisher.com to build your custom gene set.
Maximize delivery with Thermo Scientific Open Biosystems Express-In transfection reagent

Express-In transfection reagent is a proprietary lipopolymeric formulation developed and optimized for transfection of plasmid DNA into the nucleus of cultured eukaryotic cells. Polymers, not cationic lipids, protect DNA in the cytoplasm and promote entry into the nucleus of transfected cells.

- Polymeric formulation maximizes delivery into the nucleus
- High efficiency delivery to adherent and suspension cell lines
- Minimal cytotoxicity

* Comparison experiments were performed using manufacturers recommended protocols.

Thermo Scientific Open Biosystems Trans-Lentiviral packaging system for high efficiency and biosafety

The most effective strategy for efficiently expressing your gene of interest.

VERSATILE: Effective packaging of second or third generation lentiviral vectors

UNSURPASSED BIOSAFETY: Genes encoding the components required for packaging the viral genome are separated across five plasmids minimizing the threat of recombinant replication competent virus production

BROAD TROPISM: Generate lentiviral particles that effectively transduce both dividing and non-dividing mammalian cells in vitro and in vivo

HIGH-TITERS: Generate titers of 1-5 x 10⁶ TU/mL; virus can be further concentrated up to 1 x 10⁸ TU/mL

Superior transfection efficiency seen with Express-In transfection reagent compared* to commercially available transfection reagents in various cell lines.
Use our gene-based search to find the right product for your research

1. What is your gene of interest?
   Type the official symbol or accession related to your gene of interest into the gene search that is located at the top of every webpage.

2. Search returns the hits organized by product tab.
   The blue bar under each tab shows the number of clones returned for each.

3. Click the product tab you are interested in to view the specific clones for each.
   Genbank accession and a brief description of the clone, species and vector are available for each clone.
   Clicking on Clone ID brings up additional clone details including sequence.

4. If additional filtering is required then click ‘Change Filters’ to narrow the results by species or vector type.
   The number of clones shown in each tab’s count will decrease accordingly.

5. Add a clone to your cart by clicking on the cart icon.
   A check mark will appear when it has been added.

To quickly find Incyte Clones, ESTs and clone homologs from different species, BLAST your sequence of interest against all 11 million sequences in our clone collection.

Gene search and Online BLAST: www.thermo.com/openbiosystems
References

1. MGC (Mammalian Gene Collection) Program Team (2002) Generation and Initial Analysis of more than 15,000 Full-Length Human and Mouse cDNA Sequences. PNAS 2002 99: 26; 16899-16903


Thermo Scientific Open Biosystems
cDNAs and ORFs

Contact Information

Thermo Scientific Gene Expression Products

North America
601 Genome Way, Suite 2100
Huntsville, AL 35806
Toll free: 888-412-2225
Tel: 256-704-4848
Fax: 256-704-4849

www.thermo.com/openbiosystems

US e-mail: openbiosystems@thermofisher.com

Europe e-mail: techservice.emea.genomics@thermofisher.com

Belgium
Tel: 0800 80543
Fax: 0800 80178

France
Tel: 08009 14294
Fax: 08009 11885

Germany
Tel: 0800 1830746
Fax: 0800 1810366

The Netherlands
Tel: 08000 223464
Fax: 08000 223291

United Kingdom
Tel: 08009 171501
Fax: 01372 840545

Switzerland
Tel: 08005 63505
Fax: 08008 38669

Other Countries:
Technical support: +44 1372 840410