

【小鼠大学问】2分钟让你看懂小鼠模型定制方案

一份基因工程小鼠定制方案一般包括以下几方面内容：

- 定制目的
- 目的基因信息
- 模型构建策略
- 定制周期
- 风险提示

定制目的

在进行方案分析前，首先我们需要清楚地了解你的需求以及目的。例如：我需要一个 Atg5 基因的条件性基因敲除小鼠来研究自噬。

靶基因信息

接下来我们会针对你感兴趣的靶基因，例如 Atg5 基因，来进行分析啦：

Atg5 基因在小鼠中的基因全名是 autophagy related 5，染色体位置是位于小鼠10号染色体，全长约96kb。



图1. Atg5 基因染色体大致位置

目前一共发现了两个转录本，分别为 Atg5-201 和 Atg5-202。其中202号转录本不编码蛋白；仅201号转录本编码功能蛋白产物，最终翻译成含有275个氨基酸的 Atg5 蛋白。

Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt	RefSeq	Flags
Atg5-201	ENSMUST00000039286.4	2352	275aa	Protein coding	CCDS23824	Q99J83	NM_001314013 NM_053069 NP_001300942 NP_444299	TSL:1 GENCODE basic APPRIS P1
Atg5-202	ENSMUST00000217412.1	415	No protein	Processed transcript	-	-	-	TSL:3

图2. Atg5 基因转录本信息

Atg5-201号转录本共有8个外显子（exon），翻译起始于第2个外显子中的起始密码子，终止于第8个外显子。在第7个内含子（intron）中还包含一个microRNA，Mir1929。

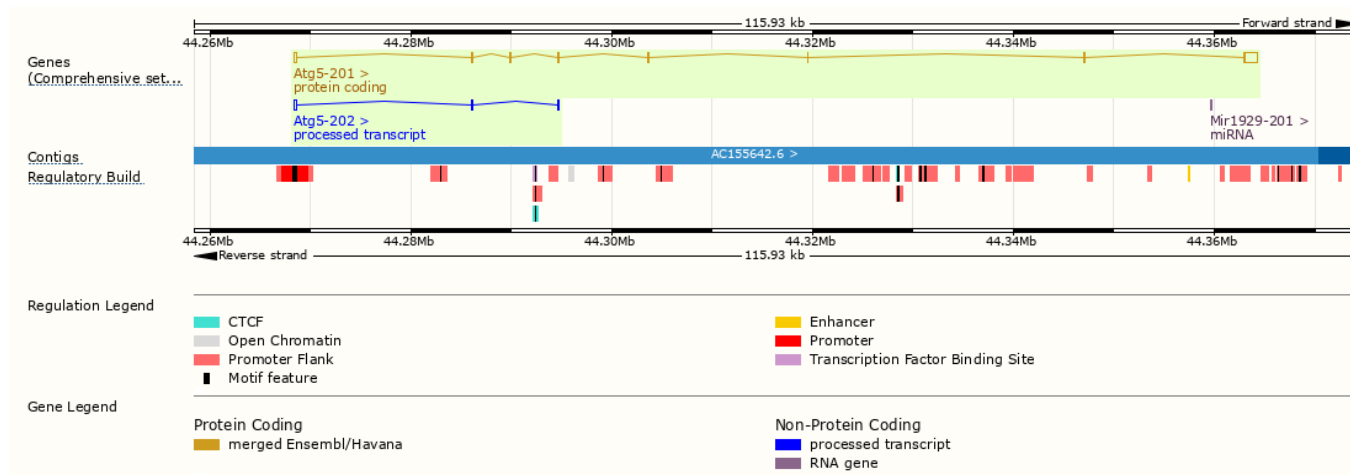


图3. Atg5 基因转录本结构信息

我们还会查询这个基因是否已经有 KO 或 CKO 的小鼠模型被文献报道过，以及它们的敲除策略是什么，有什么显著的表型。

比如，已经被报道的 Atg5 基因敲除小鼠模型有：

Allele Symbol Gene; Allele Name	Chr	Synonyms	Category	Abnormal Phenotypes Reported in these Systems
Atg5^{tm1.1Myok} autophagy related 5; targeted mutation 1.1, Minesuke Yokoyama	10		Targeted (Null/knockout)	
Atg5^{tm1Myok} autophagy related 5; targeted mutation 1, Minesuke Yokoyama	10	Atg5 ^{flox}	Targeted (Conditional ready, No functional change)	behavior, cardiovascular, cellular, digestive/alimentary, endocrine/exocrine, growth/size/body, homeostasis, liver/biliary, mortality/aging, muscle, nervous system, renal/urinary, respiratory, vision/eye
Atg5^{tm1Nmz} autophagy related 5; targeted mutation 1, Noboru Mizushima	10	Atg5 ⁻	Targeted (Null/knockout)	behavior, cardiovascular, cellular, endocrine/exocrine, hearing/vestibular/ear, hematopoietic, homeostasis, immune, liver/biliary, mortality/aging, nervous system

图4. Atg5 基因现有敲除小鼠模型信息

现有 Atg5 CKO 小鼠模型的敲除策略：

A targeting vector was designed to flank exon 3 with loxP sites. A neomycin resistance cassette

was included such that expression of cre recombinase would remove both exon 3 and the neo.
(Reference: Hara T, et al., Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. Nature. 2006 Jun 15;441(7095):885-9)

敲除或突变小鼠的已知表型有：

Mutation of this gene results in impaired autophagy due to absence of autolysosomes. Homozygotes die within 1 day of birth, have shorter survival times and reduced amino acid levels under fasting conditions. Homozygotes for a gene trap insertion mutation in this gene show no abnormal phenotype.

等一下，这些信息是从何而知的？**3个网站：NCBI、Ensembl、MGI**

模型构建策略

要敲除 Atg5 基因，我们可以选择 exon3 作为敲除区域，这与已经报道的 CKO 模型不谋而合。

为什么选 exon3 呢？

来看看 Atg5 基因的具体外显子序列信息吧。

No.	Exon / Intron	Start	End	Start Phase	End Phase	Length	Sequence
	5' upstream sequence					gagggcggagcgttgccacgtggtcttggtagcgtcatctcggggcgcc
1	ENSMUSE00000519951	44,268,358	44,268,584	-	-	227	GAGGGTGACTGGACCTACGGTGGGCGCCGGCGCTCGGTTCTCTGGTTCCGAGGGCGG GAAGTCCCCCTCGGCTGTTTTTCGACCCGGGCGGCAGCACTGTGCGGCTGCGCGCCAC TGGACTTCTGCTCCTGCGCTCTGCAGGACAGTGTGATCCGGCGACAGAACCCGAC GAGCGGCTTTCGCGCTGCGGGAAGCCGAGCAGAGCCGCACTCGG
	Intron 1-2	44,268,585	44,286,031			17,447	gtagggctcatgtgtgtcccgggcg.....ctaagtgtatttatttctgtttccag
2	ENSMUSE00000294761	44,286,032	44,286,197	-	0	166	TTTGGCTTTGGTTGAAGGAAGAACTTAGCCTATATGACTGCTTCATCCACTGGAAGAAT GACAGATGACAAAGATGTGCTTCGAGATGTGTGTTTGGACGAATCCAACTTGTCTTAC TCTCTATCAGGATGAGATAACTGAAAAGAGAAGCAGAACCACTACTAT
	Intron 2-3	44,286,198	44,289,862			3,665	gtaagtgtgcaaaaggggacctataa.....atggtttttcaaattttcatttcag
3	ENSMUSE00000521459	44,289,863	44,289,990	0	2	128	TTGCTTTGCCAAGAGTCAGCTATTGACGTGGTACTGACAAAGTGAAGAAGCACTTT CAGAAGGTTATGAGACAAGAAGATGTTAGTGAGATATGGTTGAATATGAAGGCACACC CTGAAATG
	Intron 3-4	44,289,991	44,294,595			4,605	gtgagttttttctctctttgtaagca.....tttaagatataacctttcgttttttag
4	ENSMUSE00000468884	44,294,596	44,294,674	2	0	79	GCATTATCCAATTGGTTTACTATTTCATCTTCTTGATCAAGTTCACTCTTCCCTGGAA CATCACAGTACATTTCAAG
	Intron 4-5	44,294,675	44,303,566			8,892	gtatagtttaaatagccattcctgt.....gttttggtagttatttcccttaaatag
5	ENSMUSE00000469939	44,303,567	44,303,729	0	1	163	AGTTTTCCGAAAAGGACCTTCTACACTGTCCATCCAAGGATCGGTTGAGGCTCACTTT ATGTCGTGTATGAAGAAGCTGATGCTTTAAAGCATAAAAAGTCAAGTATCAACGAATG CAGAAAAAGACCACAAGCAGCTCTGGATGGGACTGCAGATG
	Intron 5-6	44,303,730	44,319,459			15,730	gtaactacacagtgctgactcctg.....ataacagacttttatcatcttcag
6	ENSMUSE00000294728	44,319,460	44,319,554	1	0	95	ACAGATTTGACCAGTTTTGGGCCATCAACCGAAACTCATGGAATATCCCTCCAGAGAAA ATGGATTTCTGTATATCCCTTTAGAAATATATCAG
	Intron 6-7	44,319,555	44,346,956			27,402	gtaggaagtaacattacgaaaaatg.....acttgcttgggtttattccctatag
7	ENSMUSE00000294721	44,346,957	44,347,074	0	1	118	ACCAAGCAGCGCCCTTCATCAGAAAGCTGTTCCGGCTGTGGCCAGATGGACAG CTGCACACTTTGGAGATCTCCTCAGAGAACTCTGCTCCTCCGCACTGCCCTGAAG
	Intron 7-8	44,347,075	44,362,915			15,841	gtaaggtgttagcagctgagacac.....ctgttgatctcttctctggaacag
8	ENSMUSE00000365501	44,362,916	44,364,291	1	-	1,376	ATGGAGAGAAGAGGAGCCAGGTGATGATTCACGGGATGAGCCAAATGCTGGAACCCCTC TGCAGTGGCTGAGCGAGCATCTGAGCTACCCAGATAACTTTCTCATATTAGCATTGTCC CCCAGCCAAACAGATTGAAGACTGTGTCTCCTCGTATGGAACCACTTGAGTCAGG ACAACGAGGCGTGACACCCCTGCTTCAGTCAAGTTTCAGTGGAGGCAACAGAAACCCGGGC TGCTCAAGCCPAGGAGGAAGATTCATGAGAGATAGGSCCCGGGAGGGCTGAGT GTGCACCACTGCTTCGCTGAGACACACAGGACCACTGCAGCCCTCTCTCTGTAATG CAATGCAGCCGAGCCCTTTGCTCAATCAAAAAAAAAAATGGAATGTGCCACAGTTTG TATTTCTGATTAACAATAAAGTGGGTATTTCTAAGGTGACCATGTTGGAACCTTAGGC AGGAGAATGGAACATTTGTTGAATTTCAATAGAAATAGACTTAAGAAATGAAAGAGA AATTCGTTATTAATAACTTGCAGTAATTTTTTTGTAAGAATCAAAATACAGTAACCC ATCTTCTTAATGAGACTTTCCTGTTACAGTCAGTCTATTGTTATGCAAAACAACTTG TAACCTTGATTAATGACAGTGAAGATTTTTAATAAAGCCCTCAACTATGTTTTGTCTAT TTAATAACATACAGTTTGTCACTTTTCAAAAGCCCTCCTGATCTCATAGAGTAAGCCACT TTTTCTCTGTGTTCCATTCTCACTGGCATAAGGAGTGGGGCATAAGGCACATT GTCTCAGGAGCTGTACAGGATTTAATCTGTGACTGAAATACTGTCTCTATATACT AAAGGTATAAATAATCTATCTGCTTTGCTGTTAGTTGGTCACTGTAACCTGTCAAA TCTAGACTGCCAAGTATCTGCTATGATAATTTTTGAATATTTTGAATCTCCGTTCTCT TTCCAGTGTTTTTGTTGGTTTGGTTTTGGTTTTTTTACGTTTTTTGTTTTTGGCT CCTGATTAATGCTATTGGCCCTGCCAGCAGTCTTTAAAGCCCTGAGGTGACCTGTCT CTAGACTGCAGTAGCTTTTCTTATCATTACCAAAAACATCCAGAGTTACTGGAATC CTACACAGTAAGGAAAGTTTGTGCACTCTCTGATGCTGCTTGGAGACTCTGCTGT TGATTTGTAGCTAGCTTGTCTCCACTTGAATGTCAACCACTGAGTATGCTAAAAG ATGATATCAAAAATAATGGTTCTAGATTCAATAAAGATGAATGTTTTCTTAT
	3' downstream sequence						aaaaattatgtggtggctcaagtgggtcctgatagcctcctctctct.....

图5. Atg5 外显子序列信息

Atg5 起始密码子位于 exon2，而 exon3 是其后第一个不为3的整数倍的外显子（也就是说假设敲除 exon3 会导致 Atg5 的读码框发生移码突变，使翻译提前终止），而且 exon3 两侧的 intron 长度也适合 loxp 位点的插入。

如果靶基因没有会导致移码突变的 exon，那么可以选择敲除之后能移除重要结构域的 exon，从而实现蛋白突变。

决定了 Atg5 exon3 作为我们 flox 的区域后，我们会根据敲除后的剩余外显子序列推测出仍可能翻译出来的突变蛋白氨基酸序列，以供参考。

不过这个预测的突变蛋白是否具有功能，我们也无法预知呢。

Atg5 CKO 整个 flox 区域大约将有0.7kb，我们可以采用周期更短、费用更低的 Crispr-Cas9 技术来制备这个 CKO 模型。接着我们就要针对 loxp 位点将要插入的序列（intron2 和 intron3）来设计 guideRNA，并绘制打靶策略图。

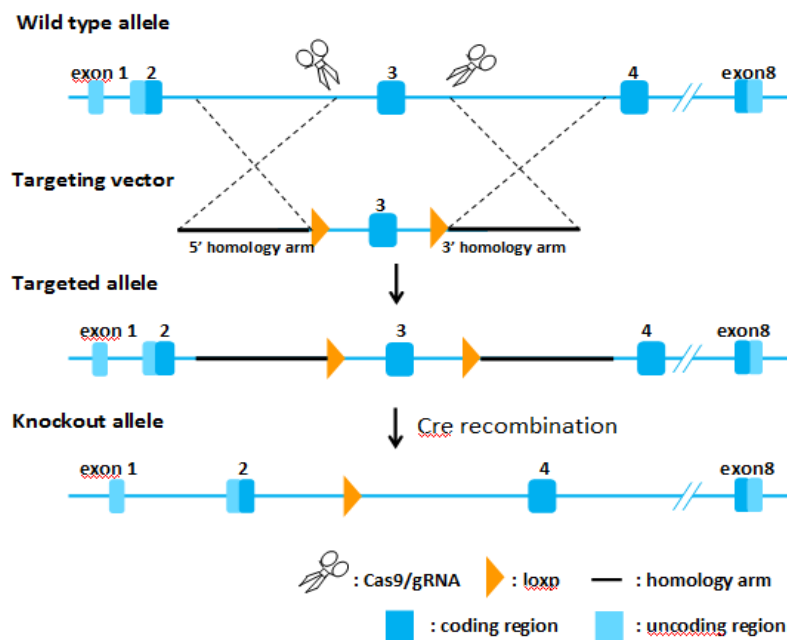


图7. Atg5 CKO 定制策略图

Intron2 序列信息如下：

```
aattgatagaactactagatctggattagatgccatgcaattatgtattagataaaggggt
agggaaaattatagattgtgtcaa
```

上述序列中的适合的 guide RNA 靶序列为：

Guide #1 ATGTATTAGATAAAGGGGTA GGG

Intron3 序列信息如下：

```
tatcaaataccaggaagctagcaccatacagcaggaccctctctcccaccgtcccccac
cgc
```

上述序列中的适合的 guide RNA 靶序列为：

Guide #2 GAAAGCTAGCACCATACAGC AGG

图8. 针对 Atg5 intron2 和 intron3 序列设计 guideRNA

如果 flox 区域非常庞大或需要重组的结构非常复杂，我们会建议采用 ES 细胞打靶的方式。

定制周期

然后就要告诉你完成这个小鼠模型的定制大约需要多少时间，也就是定制的周期。Crispr-Cas9 途径的定制周期一般为以下表格所示：

	流程	时间 (周)
1	GuideRNA 及 donor 载体构建	2-4
2	显微注射 (初次) *	1-2
3	F0 代小鼠出生	3
4	F0 代小鼠基因型鉴定*	4-6
5	F0 代小鼠繁育及 F1 代小鼠出生	8
6	F1 代小鼠基因型鉴定	4-5
7	报告撰写, 小鼠交付	2-5
	总计	24-40

而 ES 细胞打靶途径的定制周期如下:

	流程	时间 (月)
1	BAC 订购	1
2	同源重组载体构建	1-2
3	ES 细胞打靶	1-2
4	ES 细胞克隆鉴定	1
5	囊胚注射, 嵌合鼠获得	1-2
6	嵌合鼠繁育	2-3
7	杂合子小鼠获得与鉴定	1-2
8	报告撰写, 小鼠交付	1
	总计	9-14

风险提示

方案中还有一部分内容提示了这个小鼠模型在构建策略、技术路线等方面可能存在的风险或难点。预先了解风险也很重要哦!

比如:

loxp 位点插入的位置是否对基因外显子剪接有影响?

敲除的区域是否包含了其它基因的编码或调控区域?

基因敲除或突变后是否可能发生胚胎致死?

等等等等...

