

## gRNA\_Prediction 程序

### 一、 程序功能：

对目标基因进行 gRNA 预测，并对给 gRNA 加上指定 overhang。

### 二、程序文件目录：

```
bin
├── DNA.jpg
├── Get_Gene_info_genbank.exe
├── bowtie-align-l.exe
├── bowtie-build-l.exe
├── chopchop.exe
├── gRNA_GUI.exe
├── gRNA_Prediction.exe
└── src
```

### 程序文件说明

bin 目录中的文件要保持在同一目录中。

DNA.jpg 是帮助框旁边的背景缩图

Get\_Gene\_info\_genbank 是处理 genbank 文件获取坐标及序列信息的程序

bowtie-align-l 是 bowtie 比对程序

bowtie-build-l 是建立 bowtie index 索引程序

chopchop 是预测 gRNA 的主程序

gRNA\_GUI 是 GUI 界面程序，双击即可打开用户界面

gRNA\_Prediction 是出来 GUI 界面参数及 chopchop 预测结果的主程序

src 是 primer3 的依赖库文件

### 三、程序输入文件和参数

**输入文件：** genbank 格式的基因组文件和目标基因的 configure 文件

**参数：**

名称	示例	说明
Genome	GCF_000146045.2_R64_genomic.gbff	目标基因所在基因组 genbank 文件
Configure	example.xls	目标基因 txt 文件，至少提供染色体和基因 locus_tag 两列，table 分隔。 (也可自定义基因名同时提供染色体坐标信息)
Outdir	Path/result	输出结构目录路径，result 目录可自动创建
gRNA_type	ORF	用于预测 gRNA 的区域，只可选 ORF、Upstream、Downstream。
enzyme_name	BsaI,BsmBI	限制性 DNA 酶，多个可用逗号隔开。酶的名称须和附表支持的 Enzyme List 一致。
enzyme_site	GGTCTC,CGTCTC	酶切识别位点，反向互补序列无需填写，多个可用逗号隔开
forward_overhang	GCGA	加到 forward gRNA 的 overhang
reverse_overhang	AAAT	加到 reverse gRNA 的 overhang
gene_region_for_gRNA	default whole orf	gRNA_type 是 ORF 时的参数，默认不需设置
gene_upstream_len	2000	gRNA_type 是 Upstream 时的参数
gene_downstream_len	2000	gRNA_type 是 Downstream 时的参数
Design_Type	CRISPR	gRNA 设计类型，只可选 CRISPR
Model	Cas9	选用的酶类。只可选 Cas9 或 Cpf1
Score_gc	True	是否对 GC content 打分
ScoringMethod	DOENCH_2016	选择对 CRISPR 的打分算法。Cas9 Default is DOENCH_2016,while Cpf1 is KIM_2018。
Target_type	gene	在 genbank 中的 feature 类型，gene、CDS、mRNA 等。
Guid_size	20	gRNA 的长度
PAM	NGG	Pam 序列
gRNA_num	3	最多输出 gRNA 数目
Max_offtargets	300	最大允许脱靶数
Max_mismatches	3	最大允许错配位数

**Configure 文件:**

至少要存在 ChrID 和 GeneID, 当 GeneID 不是基因组上已注释的 locus\_tag 时, 需要自定义 GeneID 并提供 Start、End、Strand, 如: Fragment1 2480 2707 -

**example.xls: 目标基因列表文件 (table 分隔的文本文件)**

```

Chrname Locus_tag      Start   End     Strand
NC_001133      Fragment1      2480    2707    -
NC_001224      Q0020
NC_001224      Q0045
NC_001224      Q0050
NC_001224      Q0055
NC_001224      Q0060
NC_001224      Q0065
NC_001224      Q0070
NC_001224      Q0075
NC_001224      Q0080
NC_001224      Q0085

```

**四、使用示例:**

双击 bin/ gRNA\_GUI.exe ,弹窗如下窗口, 选择输入文件 and 对应数据点击 Run 即可。

**gRNA\_Prediction**

**输入文件**

Genome:

Configure:

Outdir:

**gRNA预测区域及酶相关参数**

gRNA\_type:

enzyme\_name:

enzyme\_site:

forward\_overhang:

reverse\_overhang:

gene\_region\_for\_gRNA:

**gRNA设置参数**

Design\_Type:

Model:

Score\_gc:

ScoringMethod:

Target\_type:

Guid\_size:

PAM:

gRNA\_num:

Max\_offtargets:

Max\_mismatches:

**参数说明**

The help of Program:

Genome: the genome file in genbank format.

Configure: the target genes for gRNA predicting in a table file.  
eg: Chr Gene Start End Strand

Outdir: the output path of results.

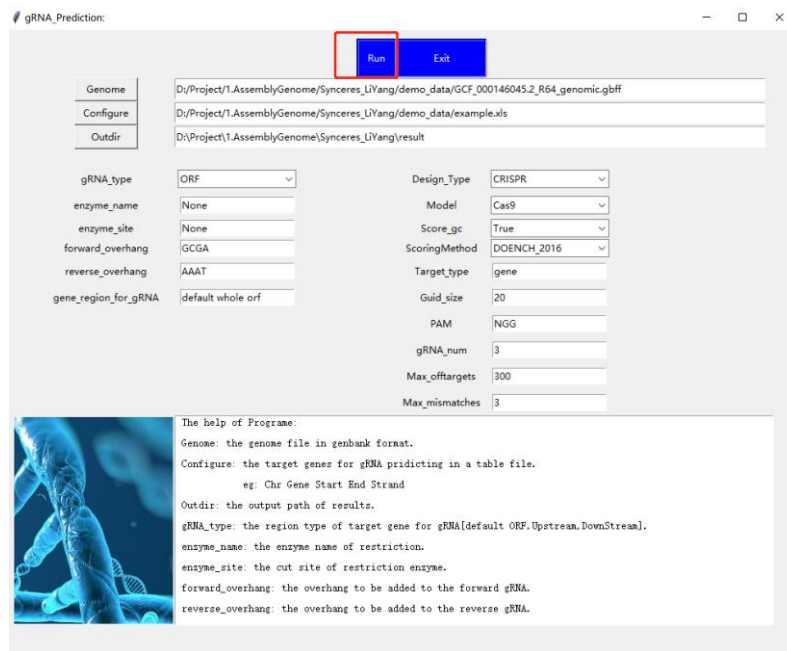
gRNA\_type: the region type of target gene for gRNA[default ORF,Upstream,DownStream].

enzyme\_name: the enzyme name of restriction.

enzyme\_site: the cut site of restriction enzyme.

forward\_overhang: the overhang to be added to the forward gRNA.

reverse\_overhang: the overhang to be added to the reverse gRNA.



## 五、输出结果：

result

```

├── All_gRNA.xlsx
├── gRNA.conf
├── gRNA_result
│   ├── Fragment1
│   ├── Q0020
│   ├── Q0045
│   ├── Q0050
│   ├── Q0055
│   ├── Q0060
│   ├── Q0065
│   ├── Q0070
│   ├── Q0075
│   ├── Q0080
│   ├── Q0085
│   └── bowtie_db
├── input
│   ├── GCF_000146045.2_R64_genomic.gbff
│   └── example.xls
└── run.log
  
```

All\_gRNA.xlsx 是所有目标基因 gRNA 汇总结果, 包含 forward 和 reverse gRNA 及 overhang, 每个 sheet 是一种 gRNA, 有多少 gRNA\_num 就有多少 sheet。

gRNA.conf 是参数配置文件。

gRNA\_result 是各目标基因 gRNA 预测结果文件。

input 是原始输入数据。

run.log 是程序运行记录文件

**附表：支持的 Enzyme List:**

['AanI', 'AarI', 'AasI', 'AatII', 'Aba6411II', 'AbaB8342IV', 'AbaCIII', 'AbaSI', 'AbaUMB2I', 'Abr4036II', 'AbsI', 'Acc16I', 'Acc36I', 'Acc65I', 'Acc65V', 'AccB1I', 'AccB7I', 'AccBSI', 'Accl', 'AcclI', 'AcclII', 'AcclIII', 'AchA6III', 'Acil', 'AclI', 'AclWI', 'Aco12261II', 'AcoI', 'AcoY31II', 'Acsl', 'Acul', 'AcvI', 'Acyl', 'Adel', 'Adh6U21I', 'Afal', 'AfeI', 'Afil', 'AflII', 'AflIII', 'AgeI', 'AgsI', 'AhalII', 'AhdI', 'AhlI', 'AhyRBAHI', 'AhyYL17I', 'AjlI', 'AjlI', 'AjlI', 'AleI', 'AlfI', 'Alol', 'AluBI', 'Alul', 'Alw21I', 'Alw26I', 'Alw44I', 'AlwFI', 'AlwI', 'AlwNI', 'Ama87I', 'AmaCSI', 'Aod1I', 'Aor13HI', 'Aor51HI', 'AoxI', 'ApaBI', 'Apal', 'ApaLI', 'ApeKI', 'ApoI', 'ApyPI', 'Aqull', 'AqullI', 'AqulV', 'ArsI', 'Ascl', 'Asel', 'Asi256I', 'AsiGI', 'AsiSI', 'Asp103I', 'Asp114pII', 'Asp337I', 'Asp700I', 'Asp718I', 'AspA2I', 'AspAMDIV', 'AspBHI', 'AspDUT2V', 'AspJHL3II', 'AspLEI', 'AspNIH4III', 'AspS9I', 'AspSLV7III', 'Asu14238IV', 'AsuC2I', 'AsuHPI', 'Asul', 'Asull', 'AsuNHI', 'AteTI', 'Aval', 'Avall', 'AvalII', 'AvrII', 'Awo1030IV', 'Axyl', 'BaeGI', 'BaeI', 'Bag18758I', 'Ball', 'BamHI', 'BanI', 'BanII', 'BanLI', 'BarI', 'Bau1417V', 'Baul', 'Bbr52II', 'Bbr57III', 'Bbr7017II', 'Bbr7017III', 'BbrPI', 'BbsI', 'BbuB31I', 'BbuB31II', 'Bbv12I', 'BbvCI', 'Bbvl', 'BbvII', 'Bccl', 'Bce3081I', 'Bce83I', 'BceAI', 'BceSIV', 'Bcefl', 'Bcgl', 'BciT130I', 'BciVI', 'BclI', 'BcnI', 'BcoDI', 'Bcul', 'Bdal', 'BetI', 'Bfal', 'BfaSI', 'Bfil', 'Bfml', 'Bfol', 'Bfrl', 'BfuAI', 'Bful', 'Bga514I', 'BgII', 'BgIII', 'BinI', 'BisI', 'BkrAM31DI', 'Ble402II', 'BlnI', 'BloAI', 'Blpl', 'BlsI', 'BmcAI', 'Bme1390I', 'Bme18I', 'BmeDI', 'BmeRI', 'BmeT110I', 'BmgBI', 'BmgI', 'BmgT120I', 'Bmil', 'BmrFI', 'BmrI', 'BmsI', 'BmtI', 'Bmul', 'BoxI', 'Bpil', 'BplI', 'Bpml', 'Bpu10I', 'Bpu1102I', 'Bpu14I', 'BpuEI', 'BpuMI', 'Bsa29I', 'BsaAI', 'BsaBI', 'BsaHI', 'Bsal', 'BsaJI', 'BsaWI', 'BsaXI', 'BsbI', 'Bsc4I', 'BscAI', 'BscGI', 'Bse118I', 'Bse1I', 'Bse21I', 'Bse3DI', 'Bse8I', 'BseAI', 'BseBI', 'BseCI', 'BseDI', 'BseGI', 'BseJI', 'BseLI', 'BseMI', 'BseMII', 'BseNI', 'BsePI', 'BseRI', 'BseSI', 'BseX3I', 'BseXI', 'BseYI', 'BsgI', 'Bsh1236I', 'Bsh1285I', 'BshFI', 'BshNI', 'BshTI', 'BshVI', 'BsiEI', 'BsiHKA', 'BsiHKCI', 'Bsil', 'BsiSI', 'BsiWI', 'BsiYI', 'BslFI', 'BslI', 'BsmAI', 'BsmBI', 'BsmFI', 'BsmI', 'BsnI', 'Bso31I', 'BsoBI', 'Bsp119I', 'Bsp120I', 'Bsp1286I', 'Bsp13I', 'Bsp1407I', 'Bsp143I', 'Bsp1720I', 'Bsp19I', 'Bsp24I', 'Bsp3004IV', 'Bsp460III', 'Bsp68I', 'BspACI', 'BspANI', 'BspCNI', 'BspD6I', 'BspDI', 'BspEI', 'BspFNI', 'BspGI', 'BspHI', 'BspLI', 'BspLU11I', 'BspMAI', 'BspMI', 'BspMII', 'BspNCI', 'BspOI', 'BspPI', 'BspQI', 'BspT104I', 'BspT107I', 'BspTI', 'BspTNI', 'BsrBI', 'BsrDI', 'BsrFI', 'BsrGI', 'BsrI', 'BssAI', 'BssECI', 'BssHII', 'BssMI', 'BssNAI', 'BssNI', 'BssSI', 'BssT1I', 'Bst1107I', 'Bst2BI', 'Bst2UI', 'Bst4CI', 'Bst6I', 'BstACI', 'BstAFI', 'BstAPI', 'BstAUI', 'BstBAI', 'BstBI', 'BstC8I', 'BstDEI', 'BstDSI', 'BstEII', 'BstENI', 'BstF5I', 'BstFNI', 'BstH2I', 'BstHHI', 'BstKTI', 'BstMAI', 'BstMBI', 'BstMCI', 'BstMWI', 'BstNI', 'BstNSI', 'BstPAI', 'BstPI', 'BstSCI', 'BstSFI', 'BstSLI', 'BstSNI', 'BstUI', 'BstV1I', 'BstV2I', 'BstX2I', 'BstXI', 'BstYI', 'BstZ17I', 'BstZI', 'Bsu15I', 'Bsu36I', 'Bsul', 'BsuRI', 'BsuTUI', 'Btgi', 'BtgZI', 'BthCI', 'BtrI', 'BtsCI', 'BtsI', 'BtsIMutI', 'BtuMI', 'Bve1B23I', 'Bvel', 'Cac8I', 'Cail', 'Cal14237I', 'CalB3II', 'Cau10061II', 'Caull', 'Cba13II', 'Cba16038I', 'Cbo67071IV', 'Cch467III', 'CchII', 'CchIII', 'Ccil', 'CciNI', 'Cco14983V', 'Cco14983VI', 'CcrNAIII', 'Cdi11397I', 'Cdil', 'Cdpl', 'Cdu23823II', 'Cfol', 'Cfr10I', 'Cfr13I', 'Cfr42I', 'Cfr9I', 'CfrI', 'CfrMH13II', 'CfrMH16VI', 'Cfupf3II', 'Cgl13032I', 'Cgl13032II', 'Chal', 'Cje265V', 'Cje54107III', 'CjeFIII', 'CjeFV', 'CjeI', 'CjeNII', 'CjeNIII', 'CjeNV', 'CjeP659IV', 'CjePI', 'Cjul', 'Cjull', 'Cla11845III', 'Clal', 'Cly7489II', 'Cma23826I', 'Cpol', 'Csel', 'Csil', 'Csp2014I', 'Csp6I', 'CspAI', 'CspCI', 'Cspl', 'CstMI', 'CviAI', 'CviII', 'CviKI\_1', 'CviQI', 'CviRI', 'Dde51507I', 'Ddel', 'DinI', 'Dpnl', 'DpnII', 'Dral', 'Drall', 'DralII', 'DraRI', 'DrdI', 'DrdII', 'DrdIV', 'Dril', 'Dsal', 'DseDI', 'DvullI', 'Eael', 'Eagl', 'Eam1104I', 'Eam1105I', 'EarI', 'Ecil', 'Ecl136II', 'Ecl234I', 'Ecl35734I', 'EclXI', 'Eco105I', 'Eco130I', 'Eco147I', 'Eco24I', 'Eco31I', 'Eco32I', 'Eco43896II', 'Eco4465II', 'Eco47I', 'Eco47III', 'Eco52I', 'Eco53kl', 'Eco57I', 'Eco57MI', 'Eco72I', 'Eco81I', 'Eco88I', 'Eco91I', 'EcoBLMcX', 'EcoE1140I', 'EcoHI', 'EcoHSI', 'EcoICRI', 'EcoMVI', 'EcoNI', 'EcoNIH6II', 'EcoO109I', 'EcoO65I', 'EcoRI', 'EcoRII', 'EcoRV', 'EcoT14I', 'EcoT22I', 'EcoT38I', 'Egel', 'Ehel', 'Eli8509II', 'ErhI', 'EsaBC3I', 'EsaSSI', 'Esp3007I', 'Esp3I', 'EspI', 'Fael', 'Fail', 'Fall', 'Faql', 'FatI', 'Faul', 'FauNDI', 'Fba202Z8II', 'Fbal', 'FbII', 'Fco1691IV', 'FinI', 'Fmul', 'Fnu4HI', 'FnuDII', 'FokI', 'FriOI', 'Fsel', 'Fsp4HI', 'FspAI', 'FspBI', 'FspEI', 'Fspl', 'FspPK15I', 'FtnUV', 'GauT27I', 'Gba708II', 'GdiII', 'GlaI', 'GluI', 'GsaI', 'GsuI', 'Hael', 'HaelI', 'HaeIII', 'HapII', 'HaulI', 'Hball', 'HdeNY26I', 'HdeZA17I', 'Hgal', 'HgiAI', 'HgiCI', 'HgiEII', 'HgiJII', 'Hhal', 'Hin1I', 'Hin1II', 'Hin4I', 'Hin4II', 'Hin6I', 'HinP1I', 'HindI', 'HindII', 'HindIII', 'HinfI', 'Hpal', 'Hpall', 'HphI', 'Hpy166II', 'Hpy178III', 'Hpy188I', 'Hpy188III', 'Hpy300XI',

'Hpy8I', 'Hpy99I', 'Hpy99XIII', 'Hpy99XIV', 'Hpy99XIV\_mut1', 'Hpy99XXII', 'HpyAV', 'HpyAXIV', 'HpyAXVIII', 'HpyAXVI\_mut1', 'HpyAXVI\_mut2', 'HpyCH4III', 'HpyCH4IV', 'HpyCH4V', 'HpyF10VI', 'HpyF3I', 'HpySE526I', 'HpyUM032XIII', 'HpyUM032XIII\_mut1', 'HpyUM032XIV', 'HpyUM037X', 'Hso63250IV', 'Hso63373III', 'Hsp92I', 'Hsp92II', 'HspAI', 'HspMHR1II', 'Jma19592I', 'Jma19592II', 'Jsp2502II', 'KasI', 'Kfil', 'Kor51II', 'Kpn156V', 'Kpn2I', 'Kpn327I', 'Kpnl', 'KpnNH25III', 'KpnNIH30III', 'KpnNIH50I', 'Krol', 'Ksp22I', 'Ksp632I', 'KspAI', 'Kspl', 'Kzo9I', 'Lba2029III', 'Lbr124II', 'Lde4408II', 'Lgul', 'LlaG50I', 'Lmnl', 'Lmo370I', 'Lmo911II', 'Lpl1004II', 'Lpnl', 'LpnPI', 'Lra68I', 'LsaDS4I', 'Lsp1109I', 'Lsp48III', 'Lsp6406VI', 'Lwel', 'Mabl', 'Mael', 'Maell', 'MaellI', 'Mall', 'Maql', 'MauBI', 'Mba11I', 'Mbil', 'Mbol', 'MbolI', 'McaTI', 'Mcr10I', 'McrI', 'Mfel', 'Mfil', 'MhII', 'MjalV', 'MkaDII', 'MlsI', 'Mlu211III', 'MluCI', 'MluI', 'MluNI', 'Mly113I', 'MlyI', 'Mmel', 'MnII', 'Mox20I', 'Mph1103I', 'Mrel', 'Mrol', 'MroNI', 'MroXI', 'MscI', 'Msel', 'MslI', 'Msp20I', 'MspA1I', 'MspCI', 'MspGI', 'Mspl', 'MspI7II', 'MspI7IV', 'MspJI', 'MspR9I', 'MspSC27II', 'Mssl', 'MstI', 'Mtel', 'MtuHN878II', 'MunI', 'Mva1269I', 'Mval', 'Mvnl', 'Mwol', 'Nael', 'Nal45188II', 'Narl', 'Nbr128II', 'Ncil', 'Ncol', 'Ndel', 'NdelI', 'NgoAVII', 'NgoAVIII', 'NgoMIV', 'NhaXI', 'Nhel', 'Nhol', 'NlaCI', 'NlallI', 'NlalIV', 'Nli3877I', 'NmeA6CIII', 'NmeAIII', 'NmeDI', 'NmuCI', 'NotI', 'NpeUS61II', 'Nrul', 'Nsbl', 'Nsil', 'NspBI', 'Nspl', 'NspV', 'ObaBS10I', 'OgrI', 'OliI', 'OspHL35III', 'Pabl', 'Pac19842II', 'PacI', 'PacII', 'Pael', 'PaePA99III', 'PaeR7I', 'PagI', 'Pal408I', 'PalAI', 'Pasi', 'Paul', 'Pba2294I', 'Pcall', 'Pcel', 'Pcil', 'PciSI', 'Pcr308II', 'Pcsl', 'Pctl', 'Pdi8503III', 'Pdii', 'Pdml', 'Pdu1735I', 'PenI', 'Pfel', 'Pfi1108I', 'Pfi23II', 'Pfi8569I', 'PfiFI', 'PfiMI', 'PfiPt14I', 'Pfol', 'PfrJS12IV', 'PfrJS12V', 'PfrJS15III', 'Pin17FIII', 'PinAI', 'PinP23II', 'PinP59III', 'Pkrl', 'PlaDI', 'Ple19I', 'PleI', 'PliMI', 'PluTI', 'PmaCI', 'Pmel', 'PmII', 'Ppil', 'PpiP13II', 'PpsI', 'Ppu10I', 'Ppu21I', 'PpuMI', 'Psci', 'Pse18267I', 'PshAI', 'PshBI', 'Psil', 'Psp0357II', 'Psp03I', 'Psp124BI', 'Psp1406I', 'Psp5II', 'Psp6I', 'PspCI', 'PspEI', 'PspFI', 'PspGI', 'PspLI', 'PspN4I', 'PspOMI', 'PspOMII', 'PspPI', 'PspPPI', 'PspPRI', 'PspXI', 'PsrI', 'Pssl', 'Pst14472I', 'Pst145I', 'Pst273I', 'PstI', 'PstNI', 'PsuGI', 'Psul', 'Psyl', 'Ptel', 'Pvul', 'Pvull', 'Rba2021I', 'Rcel', 'RdeGBI', 'RdeGBII', 'RdeGBIII', 'RfilIII', 'Rgal', 'Rigl', 'Rlal', 'Rlall', 'RleAI', 'Rmu369III', 'RpaB5I', 'RpaBI', 'Rpal', 'RpaTI', 'Rrul', 'Rsal', 'RsaNI', 'Rsel', 'Rsp008IV', 'Rsp008V', 'Rsp531II', 'RspPBT52III', 'Rsr2I', 'RsrII', 'Rtr1953I', 'Sacl', 'SaclI', 'Saf8902III', 'Sag901I', 'Sall', 'SanDI', 'SapI', 'SaqAI', 'SatI', 'Sau3AI', 'Sau64037IV', 'Sau96I', 'Saul', 'Sba460II', 'Sbfi', 'Sbo46I', 'Scal', 'SchI', 'Scil', 'ScoDS2II', 'ScrFI', 'Sdal', 'SdeAI', 'SdeOSI', 'Sdul', 'Secl', 'Sell', 'Sen17963III', 'SenA1673III', 'SenSARA26III', 'SenTFIV', 'SetI', 'SexAI', 'SfaAI', 'SfaNI', 'Sfcl', 'Sfel', 'Sfil', 'Sfol', 'Sfr274I', 'Sfr303I', 'Sful', 'Sgel', 'Sgfl', 'SgrAI', 'SgrAll', 'SgrBI', 'SgrDI', 'SgrTI', 'Sgsl', 'SimI', 'SinI', 'Slal', 'Smal', 'SmaUMH5I', 'SmaUMH8I', 'Smil', 'SmiMI', 'SmII', 'Smol', 'Sna507VIII', 'SnaBI', 'Snal', 'Sno506I', 'Spe19205IV', 'SpeI', 'SphI', 'SplI', 'SpnRII', 'SpoDI', 'Srfl', 'Sse232I', 'Sse8387I', 'Sse8647I', 'Sse9I', 'SseBI', 'Ssil', 'Ssp6803IV', 'Ssp714II', 'SspD5I', 'SspDI', 'Sspl', 'SspJOR1II', 'SspMI', 'SstE37I', 'SstI', 'Sth132I', 'Sth20745III', 'Sth302II', 'SthSt3II', 'Stsl', 'Stul', 'StyD4I', 'Styl', 'SurP32aII', 'Swal', 'Taal', 'Tail', 'TaqI', 'TaqII', 'TaqIII', 'Tasl', 'TatI', 'Taul', 'Tfil', 'TpyTP2I', 'Tru1I', 'Tru9I', 'TscAI', 'TseFI', 'Tsel', 'Tsol', 'Tsp45I', 'Tsp4CI', 'TspARh3I', 'TspDTI', 'TspEI', 'TspGWI', 'TspMI', 'TspRI', 'Tssl', 'TstI', 'Tsul', 'Tth111I', 'Tth111II', 'UbaF11I', 'UbaF12I', 'UbaF13I', 'UbaF14I', 'UbaF9I', 'UbaPI', 'UcoMSI', 'Unbl', 'Van9116I', 'Van91I', 'VchE4II', 'Vdi96II', 'Vha464I', 'Vnel', 'VpaK11AI', 'VpaK11BI', 'Vspl', 'Vtu19109I', 'Wvil', 'Xagl', 'XapI', 'Xbal', 'Xca85IV', 'Xcel', 'XcmI', 'Xhol', 'XholI', 'Xmal', 'XmalII', 'XmaJI', 'Xmil', 'Xmnl', 'Xspl', 'Ykrl', 'Zral', 'Zrml', 'Zsp2I']