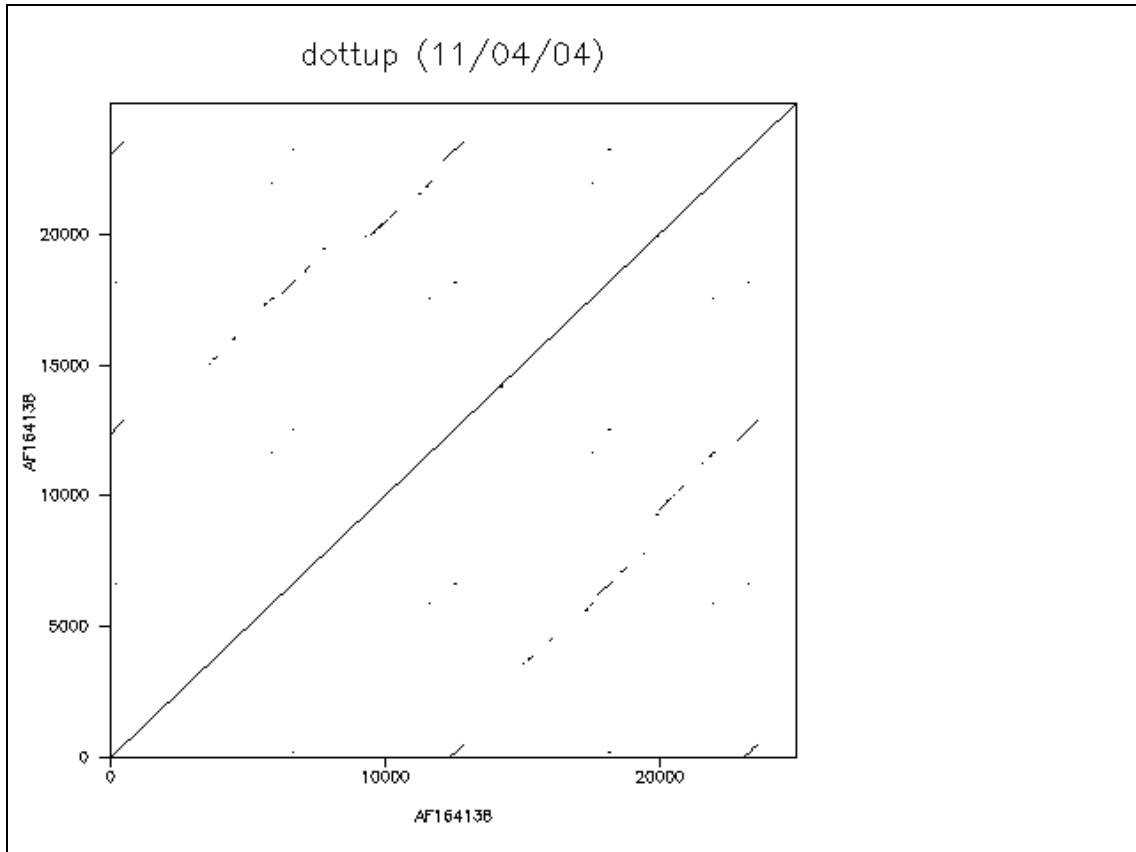


## EMBOSS 命令实例 - 序列比对

### dottup

```
dottup Fugu-Contig1.FASTA Fugu-Contig1.FASTA -word 15 -data N -graph x11
```

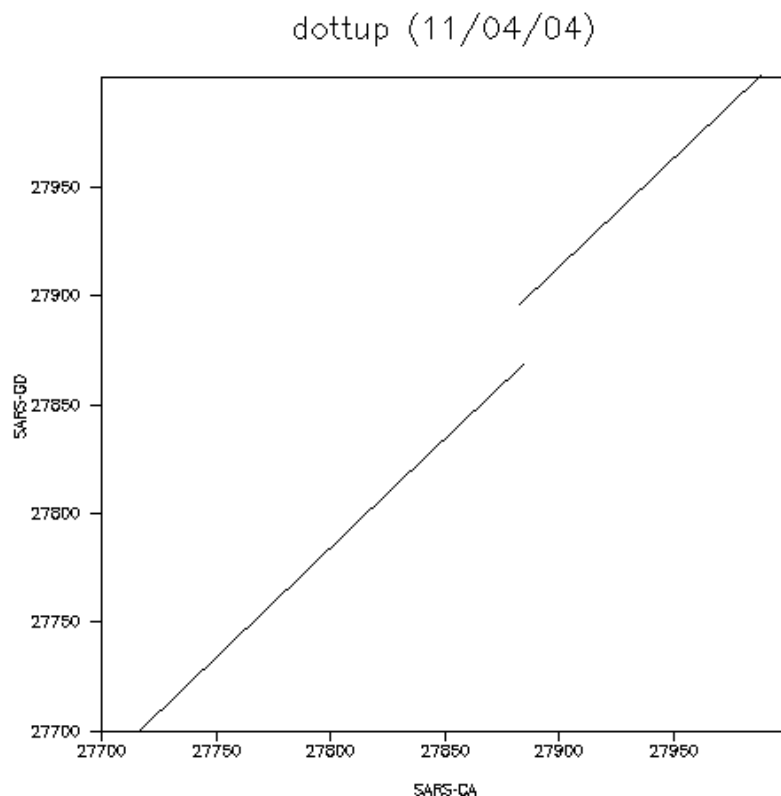
功能: 用图形方式显示 DNA 序列 “Fugu-Contig1.FASTA” 中的重复片段, 结果如下:



### dottup

```
dottup SARS-CA.FASTA -sbegin 27701 -send 28000 SARS-GD.FASTA -sbegin 27701  
-send 28000 -word 10 -data n -graph png
```

功能：用图形方式显示 SARS 冠状病毒基因组序列广东株 “SARS-GD.FASTA” 靠近 3' 端的插入片段，结果如下：



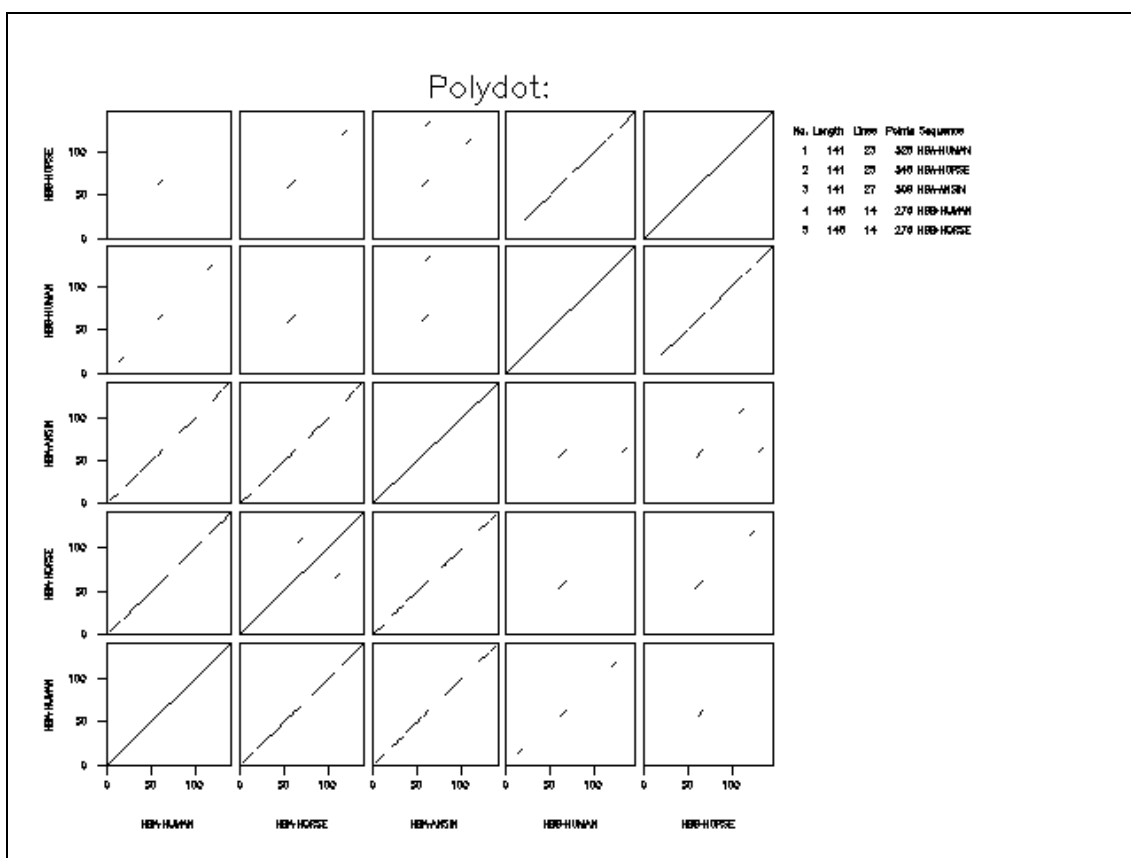
## polydot

```
polydot Hemo-5seq.FASTA -word 4 -graph x11
```

功能：用图形方式显示 5 个血红蛋白之间的相似性，输入数据如下：

```
>HBA_HUMAN - Hemoglobin alpha chain, Homo sapiens (Human) 141 AA
VLSPADKTNVKAAWGKVGAGHAGEYGAELERMFLSFPTTKTYFPHFDLSHGSAQVKGHGK
KVADALTNVAHVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPA
VHASLDKFLASVSTVLTISKYR
>HBA_HORSE - Hemoglobin alpha chain, Equus caballus (Horse) 141 AA
VLSAADKTNVKAAWSKVGGHAGEYGAELERMFLGFPTTKTYFPHFDLSHGSAQVKAHGK
KVGDALTLAVGHLLDDLPGALSNDLHAHKLRVDPVNFKLLSHCLLSTLAVHLPNDFTPA
VHASLDKFLSSVSTVLTISKYR
>HBA_ANSIN - Hemoglobin alpha-A chain, Anser indicus (Bar-headed goose) 141 AA
VLSAADKTNVKGVFSKISGHAEYGAETLERMFTAYPQTKTYFPHFDLQHGSAQIKAHGK
KVVAALVEAVNHIDDIAGALSKLSDLHAQKLRVDPVNFKFLGHCFVVAIHHPALTAE
VHASLDKFLCAVGTVLTAKYR
>HBB_HUMAN - Hemoglobin beta chain, Homo sapiens (Human) 146 AA
VHLTPEEKSAVTALWGKVNVDVGEALGRLLVVYPWTQRFFESFGDLSTPDVAVMGNPKV
KAHGKKVLGAFSDGLAHLNLDKGTFTLSELHCDKLVDPENFRLLGNVLVLCVLAHFFGK
EFTTPVQAAAYQKVAVGAVANALAHKYH
>HBB_HORSE - Hemoglobin beta chain Equus caballus (Horse) 146 AA
VQLSGEEKAAVLAALWDKVNVEEVGGEALGRLLVVYPWTQRFFDSFGDLSPGAVMGNPKV
KAHGKKVLHFSFGEGVHHLNLDKGTFAALSELHCDKLVDPENFRLLGNVLVVVLAHFFGK
DFTPELQASYQKVAVGAVANALAHKYH
```

输出结果如下:



## needle

```
needle sars_ca.fasta sars_gd.fasta -gapo 10 -gape 0.5 -out SARS-CA_SARS-GD.needle
```

```
#####
# Program: needle
# Runday: Sun Apr 11 23:19:34 2004
# Align_format: srspair
# Report_file: SARS-CA_SARS-GD.needle
#####
#-----
#
# Aligned_sequences: 2
# 1: SARS_CA
# 2: SARS_GD
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 329
# Identity: 270/329 (82.1%)
# Similarity: 270/329 (82.1%)
# Gaps: 58/329 (17.6%)
# Score: 1322.0
```

SARS_CA	1	ttaataatgcttattatatatttgggttttctactcgaaatccaggatctaga	50
SARS_GD	1	tattttgggttttctactcgaaatccaggatctaga	34
SARS_CA	51	agaaccttgtagcaaaagtctaaacgaacatgaaacttctcattggtttga	100
SARS_GD	35	agaaccttgtagcaaaagtctaaacgaacatgaaacttctcattggtttga	84
SARS_CA	101	cttgatatttctctatgacagttgcatatgcactgtagtacagcgctgtgca	150
SARS_GD	85	cttgatatttctctatgacagttgcatatgcactgtagtacagcgctgtgca	134
SARS_CA	151	tctaataaacctcatgtgcttgaagatccttg-----	182
SARS_GD	135	tctaataaacctcatgtgcttgaagatccttgctcctactggttaccaacc	184
SARS_CA	183	-----taaggtagcaaacactaggggtaataacttatagcactgctt	221
SARS_GD	185	tgaatggaatataaggtagcaaacactaggggtaataacttatagcactgctt	234
SARS_CA	222	ggctttgtgctctaggaaaggttttaccttttcatagatggcacactatg	271
SARS_GD	235	ggctttgtgctctaggaaaggttttaccttttcatagatggcacactatg	284
SARS_CA	272	gttcaaacatgcacacctaattgttactat	300
SARS_GD	285	gttcaaacatgcacac	300

## water

```
water sars_ca.fasta sars_gd.fasta -gapo 10 -gape 0.5 -out SARS-
CA_SARS-GD.water
```

#####			
# Program: water			
# Runday: Sun Apr 11 23:21:12 2004			
# Align_format: srspair			
# Report_file: SARS-CA_SARS-GD.water			
#####			
#=====			
#			
# Aligned_sequences: 2			
# 1: SARS_CA			
# 2: SARS_GD			
# Matrix: EDNAFULL			
# Gap_penalty: 10.0			
# Extend_penalty: 0.5			
#			
# Length: 300			
# Identity: 270/300 (90.0%)			
# Similarity: 270/300 (90.0%)			
# Gaps: 29/300 ( 9.7%)			
# Score: 1322.0			
SARS_CA	17	tattttgggttttctactcgaaatccaggatctagaagaacctgtacaaa	66
SARS_GD	1	tattttgggttttctactcgaaatccaggatctagaagaacctgtacaaa	50
SARS_CA	67	gtctaaacgaacatgaaacttctcattggtttgacttgatttctctatg	116
SARS_GD	51	gtctaaacgaacatgaaacttctcattggtttgacttgatttctctatg	100
SARS_CA	117	cagttgcatatgcactgtagtacagcgctgtgcatctaataaacctcatg	166

SARS_GD	101	.        cagttgcatagcactgtagtacagcgctgtgcatctaataaacctcatg	150
SARS_CA	167	tgcttgaagatccttg-----taagg 	187
SARS_GD	151	tgcttgaagatccttgctactggttaccaacctgaatggaatataagg	200
SARS_CA	188	tacaacactagggtaataacttatagcactgcttggctttgtgctctagg 	237
SARS_GD	201	tacaacactagggtaataacttatagcactgcttggctttgtgctctagg	250
SARS_CA	238	aaaggttttaccttttcatagatggcacactatggttcaaacatgcacac 	287
SARS_GD	251	aaaggttttaccttttcatagatggcacactatggttcaaacatgcacac	300