

Question 2

How can sequence-tagged sites within a DNA sequence be identified?

doi:10.1038/ng967

The NCBI's electronic PCR (e-PCR) tool¹², which is part of the UniSTS resource, can be used to find STS markers within a DNA fragment of interest. UniSTS (<http://www.ncbi.nih.gov/genome/sts/>) contains all the available data on STS markers, including primer sequences, product size, mapping information and alternative names. Links to other NCBI resources such as Entrez, LocusLink and the MapViewer are also provided. e-PCR looks for potential STSs in a DNA sequence by searching for subsequences with the correct orientation and distance that could represent the PCR primers used to generate known STSs.

The e-PCR home page can be found by going to the NCBI home page, at <http://www.ncbi.nlm.nih.gov>, and then following the *Electronic PCR* link in the right-hand column. On the e-PCR home page, paste the sequence of interest or enter an accession number into the large text box at the top of the page. The accession number of the sequence for this example is AF288398. This sequence contains only one STS, stSG47693, which is located between nucleotides (nt) 2102 and 2232 of the sequence under study (Fig. 2.1).

Click on the marker name to bring up details of the STS from UniSTS (Fig. 2.2). The primer information and PCR product size are listed at the top of the page, along with alternative names for the marker. Often STSs are known by different names on differ-

ent maps. Cross-references to LocusLink, UniGene and the Genebridge 4 map to which this STS was mapped are shown next. The mapping information section contains links to the NCBI's MapViewer. At the bottom of the page, the Electronic PCR results show other sequences, including contigs, mRNAs and ESTs that may contain this STS marker.

To see the genomic context of the STS marker in all maps to which it has been mapped, click on the link labeled *MapViewer* at the top of the *Mapping Information* section. This map view (Fig. 2.3) shows two maps. Note that, in this view, the STS stSG47693 is called RH92759 (highlighted in pink). Gene Map '99–Genebridge 4 (GM99_GB4, left) has 46,000 STS markers mapped onto the GB4 RH panel by the International Radiation Hybrid Consortium. The STS map (right) shows the NCBI's placement of STSs onto the genome sequence assembly using e-PCR. Gray lines connect markers that appear in both maps, whereas the red line denotes where the STS RH92759 appears on both maps. In the region shown, there are a total of 211 STSs on the STS map, but only 20 are labeled in this view. To the right of the STS map, the green and yellow circles show the maps on which the STS markers have been placed. One can zoom in or out of this view by clicking on the lines of the zoom tool in the left sidebar.

Figure 2.1

Location: <http://www.ncbi.nlm.nih.gov/genome/sts/eper.cgi>

Search UniSTS for Go

Electronic PCR

Query sequence: g|12620204|gb|AF288398.1|AF288398, 2853 bases

Homo sapiens C1orf14 mRNA, complete cds

Site (bases)	Marker	Chr.	Organism
2102..2232	stSG47693	1	Homo sapiens

Questions or Comments? Write to the NCBI Service Desk

Disclaimer Privacy statement

Figure 2.2

Location: <http://www.ncbi.nlm.nih.gov/genome/sts/sts.cgi?uid=92492>

Search UniSTS for Go

stSG47693 UniSTS:92492

Primer Information

Forward primer: **TTCTGATATCCCCCTTGACG**

Reverse primer: **AGAAGCTCTACAAAAGGGC**

PCR product size: 131 (bp), Homo sapiens

Homo sapiens

Name: stSG47693

Also known as: RH92759

Cross References

LocusLink

LocusID: 81626

Symbol: C1orf14

Description: chromosome 1 open reading frame 14

Position: 1q25

UniGene

Hs.112017 chromosome 1 open reading frame 14

RH details

RH92759 Genebridge4

Mapping Information

View all results using the Map Viewer

stSG47693	Sequence Map:	Chr 1	mv
	Position:	178248643-178248773 (bp)	
stSG47693	GM99-GB4 Map:	Chr 1	mv
	Position:	635.56 (cRa000)	
	Lod score:	0.02	
	Reference Interval:	D1S2640-D1S481 (199.7-208.9 cM)	

Electronic PCR results

RefSeq mRNA (2)

NM_030933.1	2102..2232	Homo sapiens GE36 gene (GE36), mRNA (2363 bp)
XM_046428.4	2103..2233	Homo sapiens chromosome 1 open reading frame 14 (C1orf14), mRNA (2364 bp)

mRNA (2)

AF288397.1	1761..1891	Homo sapiens C1orf14 mRNA, complete cds (2022 bp)
AF288398.1	2102..2232	Homo sapiens C1orf14 mRNA, complete cds (2363 bp)

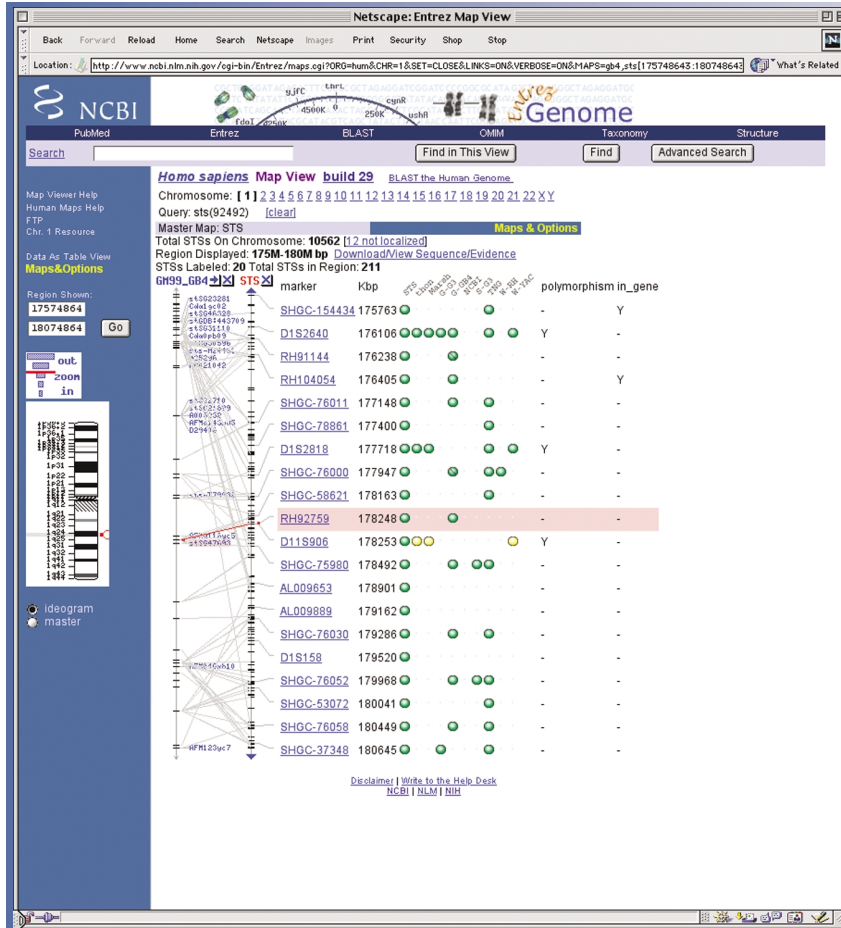


Figure 2.3