USER GUIDE

BACKGROUND

Digital production, transmission and storage have revolutionized how we access and use information but have also made archiving an increasingly complex task required active, continuing maintenance of digital media. This challenge has focused some interest on DNA as an attractive target for information storage because of its capacity for high-density information encoding, longevity under easily achieved conditions and proven track record as an information bearer.

Although techniques for manipulating, storing and copying large amounts of existing DNA have been established for many years, one of main challenges for practical DNA-based information storage is the difficulty of synthesizing long sequences of DNA de novo to an exactly specified design.

OPERATING ENVIRONMENT

Computer equipment with normal network conditions.

MAIN INTERFACE AND FUNCTION INTRODUCTION



Encode Files and Decode Sequences Support editable DNA format

DNA is a basic and univeral data storage material in biology choosen by Nature. We can use artificial DNA sequences synthesized with commerically available oligonucleotide synthesis machines to store information and use DNA sequencing machines for retrival. As an information storage material, DNA has many advantages compared with electronic materials currently in use, such as high capacity, large density and long shelf-life, if held in right conditions.

To use DNA for information storage in this electronic computer era, we need to devise a protocol to convert traditional computer files into DNA molecules and retreive the information given the sequenced DNA data. Bio101:DNA Information Storage System is such a bridge between bits and nucleotids, i.e. between the current information techology (IT) world and the future biotechnology (BT) computing world, and it is designed for the information transformation between computer files and DNA sequences. It also provides a version responding to DNA editing, bringing information edit into reality.



© Copyright 2016. Designed | Developed by 2016 UESTC Software iGEM Team.

MAIN INTERFACE USER GUIDE

Fig.1 Main Interface

In the main interface, users can click 'Encode', 'Decode' or 'Edit' button to jump to corresponding function interface. At the bottom of home page, there is a navigation bar. Users can jump to any page quickly by clicking the corresponding buttons. They can also click 'About' button to learn our algorithm and design.

'ENCODE' INTERFACE

			HOM	E ENCODE DECODE ABOUT
Encode Ele	ctronic	Files to DNA	Sequences	
Choose a file (almost any kind), in a blink or a couple of minutes	enter a secret toker 5, depending on you	n for encryption purpose, and then clic r file size.	k Encode! The encoded DNA sea	quences will be ready for download
NOTES: Both the file name and t workload reason.	the token should no	ot be more than 256 characters. The up	ope file size limit is 50 MiB and t	his limit is imposed due to the
Select a File to Choose Upload: max. 50	File No file chosen MiB (megabytes).			
Enter a Token for max. 2 Encryption:	56 characters.			
Encod	e!			
		Encode Decode Abo	ut	
		Fig.2 Encode int	erface	
● ● ● Use Vim Like A Pro × ↓ 496 - → C ◎ bio101.uestc.edu.cn/trans	Favorites	■ Ⅲ	Carch Date Modified	R × New Tab × Dram
・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	E < > 88 Favorites 百度云同步盘 爲 All My Files	Name original.txt out.txt	① ① Q Search Date Modified 13 Oct 2016, 3:21 PM 13 Oct 2016, 3:20 PM	III × New Tab × Dram 248 с ☆ 😁 ☵ 📩 ⊙ : ические аэро ぐ ScribblerToo :: Mari »
Use Vim Like A Pr × 49 O bio101.uestc.edu.cn/trans 豆醇 一学习网站 一编程 一下载资源 UESTC-SOFTWARE DNA Information Storage System	■ < > 総 Favorites 百度云同步盘 品 All My Files Cicloud Drive へ Applications ■ Desktrop	Diamon and a constraint and a const	 Q Search Date Modified Q Search 3 Oct 2016, 3:21 PM 13 Oct 2016, 3:22 PM 13 Oct 2016, 3:13 PM 13 Oct 2016, 3:13 PM 13 Oct 2016, 3:13 PM 13 Oct 2016, 3:07 PM 	B// × New Tab x Dram 248 t Ф 至 № 0 Е 248 t Устівыетов :: Mari * 248 t 248 t CODE DECODE ABOUT 1 1
Use Vim Like A Pr × よ 49 O bio101.uestc.edu.cn/trans 国語 学习网站 単編程 日下載資源 UESTC-SOFTWARE DNA Information Storage System Encode Ele	E く S 20 Favorites 百度云同步盘 All My Files C ICloud Drive A Applications Desktop 合のcuments Occuments Downloads Movies	bio101core_v2	C C Search Date Modified C Search 13 Oct 2016, 3:21 PM 13 Oct 2016, 3:20 PM 13 Oct 2016, 3:20 PM 13 Oct 2016, 3:13 PM 13 Oct 2016, 3:13 PM 13 Oct 2016, 3:07 PM 13 Oct 2016, 3:06 PM 13 Oct 2016, 2:29 PM 13 Oct 2016, 2:29 PM	Bit × New Tab × Dram 248 t Image: Scribbler Too :: Mari > 248 t CODE DECODE ABOUT 1 1
Use Vim Like A Prox 249 O bio101.uestc.edu.cn/trans 国際 学习网站 当編程 日下教教派 UESTC-SOFTWARE DNA Information Storage System Encode Ele Choose a file (almost any kind), e in a blink or a couple of minutes	E く S S Favorites 百度太同步盘 A IM Wy Files A Cloud Drive A Applications Desktop Documents Occuments Occuments Movies	Image: Second state sta	C C Search Date Modified C 13 Oct 2016, 3:21 PM 13 Oct 2016, 3:20 PM 13 Oct 2016, 3:20 PM 13 Oct 2016, 3:13 PM 13 Oct 2016, 3:13 PM 13 Oct 2016, 3:07 PM 13 Oct 2016, 3:07 PM 13 Oct 2016, 2:29 PM 13 Oct 2016, 2:29 PM 13 Oct 2016, 2:29 PM	Bit x New Tab x Dram 248 t Image: Constraint of the second
Use Vim Like A Prox 2 49 C ① bio101.uestc.edu.cn/trans D W W W W W W	E く S 部 Favorites 百度云同步盘 品 II 物 Files C ICloud Drive A Applications Desktop Documents Documents Documents Documents Documents	Image: Second system	Cancel Ope	B: x New Tab x Dram 248 t x Image: ScribblerToo :: Marl ** 248 t x ScribblerToo :: Marl ** 248 t CODE DECODE ABOUT 1 as will be ready for download it is imposed due to the
Use Vim Like A Prox 249 O bio101.uestc.edu.cn/trans D bio101.uestc.edu.cn/trans	E く S S Favorites Favorites 百度之同步盘 A All My Files Could Drive A Applications Documents Documents Documonts Documonts Options File No file chosen MIB (megabytes).	Borner All Files	Cancel Ope	Image: New Tab X Dram 248 t Image: New Tab X Dram 248 t Image: New Tab X Image: New Tab X 248 t Image: New Tab Image: New Tab Image: New Tab Image: New Tab 248 t Image: New Tab Image: New Tab Image: New Tab Image: New Tab 248 t Image: New Tab Image: New Tab Image: New Tab Image: New Tab 248 t Image: New Tab Image: New Tab Image: New Tab Image: New Tab 248 t Image: New Tab Image: New Tab Image: New Tab Image: New Tab 248 t Image: New Tab Image: New Tab Image: New Tab Image: New Tab 248 t Image: New Tab Image: New Tab Image: New Tab Image: New Tab 248 t Image: New Tab Image: New Tab Image: New Tab Image: New Tab 248 t Image: New Tab Image: New Tab Image: New Tab Image: New Tab 248 t Image: New Tab Image: New Tab Image: New Tab Image: New Tab 248 t Image: New Tab Image: New Tab Image: New Tab Image: New Tab 248 t Image: New Tab Image: New Tab Image: New Tab Image: New Tab
● Use Vim Like A Prox 2 49 ● C ① bio101.uestc.edu.cn/trans 回题 単学习网站 単編程 下發激源 C USESTC-SOFTWARE DNA Information Storage System Encode Ele Choose a file (almost any kind), e in a blink or a couple of minutes NOTES: Both the file name and t workload reason. Select a File to Upload: max. 50 Enter a Token for Encryption:	E Favorites Favorites Fake and the set of the set	Image: Second state sta	Cancel Ope	B × New Tab × Dram 248 t * •
Use Vim Like A Pr × 49 O bio101.uestc.edu.cn/trans 団神 学习网络 単単目 下弦対系 UESTC-SOFTWARE DMA Information Storage System DMA Information Storage System Choose a file (almost any kind), e in a blink or a couple of minutes NOTES: Both the file name and t workload reason. Select a File to Upload: max.50 Enter a Token for Encryption: max.23	E く) 部 Favorites 百度之同步進 All My Files C ICloud Drive Applications Documents Documents Documents Documents Movies	Ime original.txt out.txt moonlight_recover.txt mo	Cancel Ope	B New Tab Can 248 t Соле ScribbleToo :: Mari 248 t CODE DECODE 248 t CODE DECODE 248 t ScribbleToo :: Mari ScribbleToo :: Mari 248 t CODE DECODE 248 t ScribbleToo :: Mari ScribbleToo :: Mari 25 t ScribleToo :: Mari ScribbleToo :
Use Vim Like A Prix 2 49 C ① bio101.uestc.edu.cn/trans D Wim Like A Prix 2 49 C ① bio101.uestc.edu.cn/trans D Wim Like A Prix 2 1 C ① bio101.uestc.edu.cn/trans USE Stransformer System USE Stransformer System D USE Stransformer System C Dook Information Storage System	E く > 33 Favorites Favorites Factorites Factorites Factorites All My Files Cicloud Drive A Applications Desktop Documents Downloads Movies Options Options File No file chosen MiB (megabytes). 56 characters. el	Image: Second state sta	 Q. Search Date Modified I 3 Oct 2016, 3:21 PM I 3 Oct 2016, 3:20 PM I 3 Oct 2016, 3:13 PM I 3 Oct 2016, 3:13 PM I 3 Oct 2016, 3:06 PM I 3 Oct 2016, 3:06 PM I 3 Oct 2016, 2:29 PM I 3 Oct 2016, 2:29 PM I 3 Oct 2016, 2:29 PM Cancel 	B X New Tab X Dram 248 Image: Constraint of the state of the sta

Fig.3 Select Local File

UESTC-SOFT DNA Information Store	WARE ge System	HOME	ENCODE	DECODE	EDIT	ABOUT
Encode	d Sequences					
Downloadable	Downloadable Files					
Text FASTA SBOL	Nucleotide sequences in simple text format Nucleotide sequences in FASTA format Nucleotide sequences in the Synthetic Biology Oper	n Langua	ge (SBOL) fo	rmat		
BioBrick Assembly Compatibility						
Assembly Report	Biobrick assembly standards					
	Encode Decode Edit A	bout				

Fig.4 Completion and Download

In 'Encode' interface, users need to choose a file for encoding, enter specific token for this file. Since token is the key for later progress, is should be remembered. After entering Token for this file, user can submit and get result. There are three different common formats for downloading. There is also an assembly report for Biosafety issues. **'DECODE' INTERFACE**



HOME ENCODE DECODE ABOUT

Decode a DNA Sequence File to Recover the Stored Information

Upload a DNA sequence file in Text, FASTA or SBOL format and your token to extract the stored information from the synthesized DNA. If you don't have a DNA sequences file, you may use the Encode page to generate one. NOTE 1: The file name and the token should be no more than 256 characters. NOTE 2: The file size should be less than 50 MiB due to the workload reason. NOTE 3: The text format file should contain A, T, C and G only and 215 bases per line; the FASTA and SBOL format file should contain all sequences in one file. Select a DNA Choose File No file chosen Sequences File are to Select a DNA Choose File No file chosen

quences File:	max. 50 MiB (megabytes).
the Token for Decryption:	max. 256 characters.
51	Decedel

Fig.5 'Decode' Interface



Enter

← → C ③ bio101.uestc.edu.cn/t	ans Favorites	Name	Date Modified	🔺 🚽 👌 🖓 🔛 📩 🔄 🖉
豆 豆瓣 🗀 学习网站 🗀 编程 📄 下载资源	▶ 百度云同步盘	original.txt	13 Oct 2016, 3:21 PM	248 в ические аэро 🤔 ScribblerToo :: Mari У
	All My Files	moonlight_recover.txt	13 Oct 2016, 3:20 PM 13 Oct 2016, 3:13 PM	248 t
	iCloud Drive	moonlight_nt.txt	13 Oct 2016, 3:13 PM	ICODE DECODE ABOUT
Diva momation Storage System	Applications	moonlight.txt	13 Oct 2016, 3:11 PM	248 t
	Desktop	c isnt2bit.c	13 Oct 2016, 3:07 PM	
Decode a	Documents	isbit2nt	13 Oct 2016, 2:29 PM	1 Stored
Decoue u	Downloads	c isbit2nt.c	13 Oct 2016, 2:29 PM	500100
Informatio	Movies			
Informatic)[-		
		Format: All Files		
	Ontinne		Ormeri	0.000
Upload a DNA sequence file i			Cancel	ed DNA.
lf you don't have a DNA sequ	ences file, you may use	e the Encode page to generate on	e.	
NOTE 1: The file name and th	e token should be no i	more than 256 characters.		
NOTE 2: The file size should b	e less than 50 MiB due	e to the workload reason.		
NOTE 3: The text format file s	hould contain A, T, C a	nd G only and 215 bases per line;	the FASTA and SBOL format fil	e should contain all sequences in one file.
Select a DNA Choo	out.txt			
Sequences File: max.	50 MiB (megabytes).			
Enter the Token for				
Decryption.				
Deo	:ode!			
		Fig.6 Select Lo	ocal File	
		•		
	F			
DNA Information Storage Syste	E m			HOME ENCODE DECODE ABOUT
	A. 2.30 LIIGIGUELS.			
Decryption:				
De	code!			
Docodod	Informa	tion Eilo		
Decoueu	IIIIOIIIId	lion file		
Downloadable Files	5			
Decoded File This	is the file stored in you	ur DNA sequences. The file was o	ompressed using the bzin ? corr	pressor Please decompress the file using
VOUI	favorite program. Eni	oy!	ompressed using the pzip2 com	ipressor, riease decompress the file using
		-		
		Encode Decode	About	
	8 C	at 2016 Deciment L Develop		
	© Copyrigi	nt 2016. Designed Developed by	2016 UESTC Software IGEM Tea	m.

Fig.7 Completion and Download

In this interface, users are supposed to upload an encoded file, and submit all necessary bits as a whole. Token should be right, and there will a download hyperlink.Download it, the user can extract the information he stored.



USER GUIDE



Edit Electronic Files to DNA Sequences

Submit the DNA fragment and your token to decode it firstly, then edit this content to acquire new segment for substitution. Result will be displayed as text directly, but you can also download it as SBOL file

NOTE 1:Length of sequence must keep same.

NOTE 2:The Token of default example is "123".

NOTE 3:Token should keep same since enter.

original fragment:	TACGTACGTACGTAGGCTTTGGTATACGAGCAGGTTGAATGCGCCTATGCATCATAGGCGTTCATTGTA GATGGGAACATATATCCATGCTCACACCCTGCCG	ACAGATCCATTGTGTGTGTACGTTAGCCCCGCAATTTGGCGGAC
Encrypte code:		
	Submit	
	Encode Decode Edit About	
	© Copyright 2016. Designed Developed by 2016 UESTC	C Software iGEM Team.
	Fig.8 'Edit' Interfac	:e
UESTC-SOFT DNA Information Storag	e System	HOME ENCODE DECODE EDIT ABOUT
Edit Elec	tronic Files to DNA Seque	ences
Submit the DNA fragme Result will be displayed	nt and your token to decode it firstly, then edit this content to acquire n as text directly, but you can also download it as SBOL file	new segment for substitution.
NOTE 1:Length of seque	nce must keep same.	
NOTE 2: The Token of de	stault example is "123".	
	some since encer.	
edit information:	A Quiet Night Tought Li Bai (Tan	
Enter the Token for edit:		
	Submit	

Fig.9 Decode Single Segment

USER GUIDE

UESTC-SOFT DNA Information Store	WARE gge System	HOME	ENCODE	DECODE	EDIT	ABOUT
	Li Bai (Tan				6	
Enter the Token for edit:	Submit					
Results						
original fragment: TACGTACGTACGTAGG	CTTTGGTATACGAGCAGGTTGAATGCGCCTATGCATCATAGGCGTTCATTGTACAGATCCAT	TGTGTTG	TACGTTAGC	CCCGCAAT	TGGCG	GACGATGGGAAG
TACGTACGTACGTAGG	CTTTGGTATACGAGCAGGTTGAATGCGCCTATGCATCATAGGCGTTCATTGTACAGCCAACG	GCTGAATA	AAACGCTCA	AGCGATCCO	ICTTACT	GGCGGGTGGAA
SBOL-XML File	This is the file stored in your DNA sequences. The file was compressed using the by your favorite program. Enjoy!	zip2 com	pressor. Plea	ase decomp	ress the	file using
	Encode Decode Edit About					

Fig.10 Completion and Download

In this interface, the user can put a single segment which he want to edit.With correct token, Bio101 will encode it independently.Since the user edits its content, we will provide a new segment for substitution.

The result will be show as text and be downloaded as SBOL-XML file. **'ABOUT' INTERFACE**

UESTC-SOFTWARE DNA Information Storage System

HOME ENCODE DECODE EDIT ABOUT

About BIO101: Design and Algorithm

Background

Living in an era of information explosion, digital production, transmission and storage so far have not only revolutionized the way information is accessed and used, but also have made archiving of information an increasingly complex task. Have you been perplexed by vast quantities of information? And have you ever image that there exists a practical, high-capacity, low-maintenance, and even self-copy information storage medium which would be still useful after thousands of years? It wouldn't be only a dreamy illusion anymore because of the appearance of great DNA storage technology.



DNA, one of the most miraculous masterpiece created by nature as the stable genetic material, hold a great promise for high-density, long-term and massive information. For example, human genome, just 3 billion base pairs, contains all of the complex biotic information of human being, including chemical molecules forming, cell growth, human developing and so on (Figure 1). Reversely why carit we utilize DNA to store messages? Researches indicate that it is extremely dense, and spectacular high-capacity with a raw limit of 1 exabyte/mm³(109 GB/mm³). In other word, every gram of DNA is equivalent to 14 thousand of 50GB blue-rav discs or 233 3TB hard disks heavy of 151 kilos. Meanwhile. compared with now comprehensively used information storage

Fig.11 'About' Interface

In 'About' page, we explain the background theory and algorithm in details. There are design details and enough references, but you can still click this cute panda to visit our WIKI page for massive information.

USER GUIDE