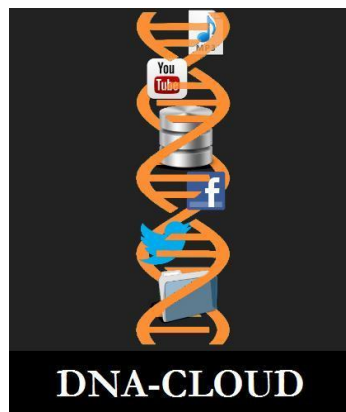


# User Manual



# DNA Cloud 2.0 User Manual

---

© 2015 Manish K Gupta,  
Laboratory of Natural Information Processing  
DA-IICT, Gandhinagar, Gujarat 382007  
<http://www.guptalab.org/dnacloud>

The software described in this book is furnished under an open source license agreement and may be used only in accordance with the terms of the agreement. Any selling or distribution of the program or its parts, original or modified, is prohibited without a written permission from

Manish K Gupta.

Documentation version 2.0

This file was last modified on Jan 27, 2015.

---

## Credits & Team

*Principle Investigator:* Manish K. Gupta, PhD.

*Graduate Mentor:* Dixita Limbachiya

*Developers:* Vijay Dhameliya, Madhav Khakhar

*Software Logo:* Foram Joshi

---

## Acknowledgments

We would like to thank Thorsten Weimann and Anand B Pillai whose open source libraries ([python-barcode](#) and [pytxt2pdf](#) respectively) have been used in the software.

# Table of contents

<b>General Information</b>	<b>Page No</b>
1. Introduction	<b>3</b>
2. Overview	
3. Product Scope	
4. Product Perspective	
<b>System Summary</b>	<b>4</b>
<b>Getting Started</b>	<b>4</b>
5. User Details	
6. Encode	
7. Decode	<b>5</b>
8. Estimator	<b>6</b>
9. Export	<b>8</b>
10. Help	<b>9</b>
<b>Support and Feedback</b>	<b>10</b>
<b>References</b>	<b>10</b>

## General Information

### 1. Introduction

The world data increase two fold every year. To store this huge data is burning issue. Since time immemorial, mankind has wanted to share and use information for later use. First, it was through the caveman paintings and symbols. Then we invented the alphabets, ideograms, numbers and other symbols. Using these, books were written and stored for future generations, in palm leaves, papyrus sheets or paper. The invention of printing brought the Gutenberg revolution, making multiple copies easily and spreading education to millions of people. Printed books occupy space. Libraries and archives are bursting at the seams. Enter the computer age and digitization using the binary code of combining zeros and ones (0,1) for alphabets and other such symbols, and reading them using the on-off electrical signals, which has made electronic storage possible, cutting down the size **and space for „hard copies“**. **Integrated circuits, processors and related electronic wizardry have** shrunk the size of computers and storage devices from room-size to finger nail size. But the storage and maintenance on this energy craving devices is burdensome. With the idea of using the natural hard drive DNA as storage device, Goldman 2013 *et al* [1] and Church 2012 *et al* [2] had stored file of size one Peta Bytes on 1 gram of DNA. Inspired from the work of Goldman, we have developed the software DNA Cloud for encoding and decoding the file to DNA. We have used the encoding and decoding method described by the Goldman in [1]. We have made the encoding and decoding procedure memory efficient.

### 2. Overview

DNA Cloud has been primarily developed to facilitate the storage of data in DNA. The software encourages the users to encode and decode the data. It converts any type of data into DNA strings and enables it to store in DNA. Also it helps to retrieve the data stored in DNA.

### 3. Product Scope

This software provides the user with:

- Encode the any type of data (text, image, audio, and video) to DNA string.
- Decode the data stored in the DNA string.
- Estimator is used to estimate the total length of DNA string used to convert the data and also the cost for the storage and biochemical properties of DNA.
- Export the converted data to various feasible formats for different application.

## 4. Product Perspective

This product's main aim is to make the process of making the computerization and computation of DNA storage user friendly.

### System Summary

Operating Systems – Windows 7, Mac OS X, Linux

This software requires the system to have specifications similar to: - Processor of .2.4 GHz Core 2 Duo, 100 MB Hard Disk, 1GB Random Access Memory and Basic peripherals like keyboard, mouse etc.

### Getting Started

Once you have successfully installed the software, open DNA Cloud by double clicking on the icon available on your desktop screen, a window with three menu options on the menu bar will appear. The dialogue box will appear with two buttons. If user wants to convert the data to DNA string, click on **File to DNA** button. If user wants to retrieve the data stored in DNA, select **DNA to File** options. Software has following main sections.

**File menu** includes following option:

- ✓ Encode (File to DNA)
- ✓ Decode (DNA to File)
- ✓ Estimator
- ✓ Export Generated Barcode
- ✓ Export DNA Synthesizer file
- ✓ Import DNA Sequencer file
- ✓ Export .dna to .PDF
- ✓ Clear temp files
- ✓ Exit

**Preferences menu** includes following option:

- ✓ Password protection
- ✓ Change password
- ✓ User Details

#### 1. Encode

This option is available under the File menu which will convert any type of data to DNA strings. When the user clicks on this option, the screen in as shown in figure 1 appears. User has to select the file to be converted into DNA string by clicking on **Choose File** button. In

DNA Cloud 2.0, there is additional feature for select the type of data encoding technique and compression technique. User can select **Golay code** type of encoding for compress and efficient data encoding. Also user can select the type of compression. There is option for **lossy and lossless compression**. If user wants compress data, he can choose lossless compression. Once the user selects the file to be encoded, the list of information for the encoded file will be displayed as below.

- Length of DNA string
- No of DNA oligonucleotides (chunks)
- Length of each DNA oligonucleotide
- File size in bytes

To save the encoded file, user can save the file with specific name on specific location by using **Encode your File** button. It will generate the file with extension .dna that has DNA string for the file selected. **RESET** button can clear the selected file and select new file by this option. It is like clear button.

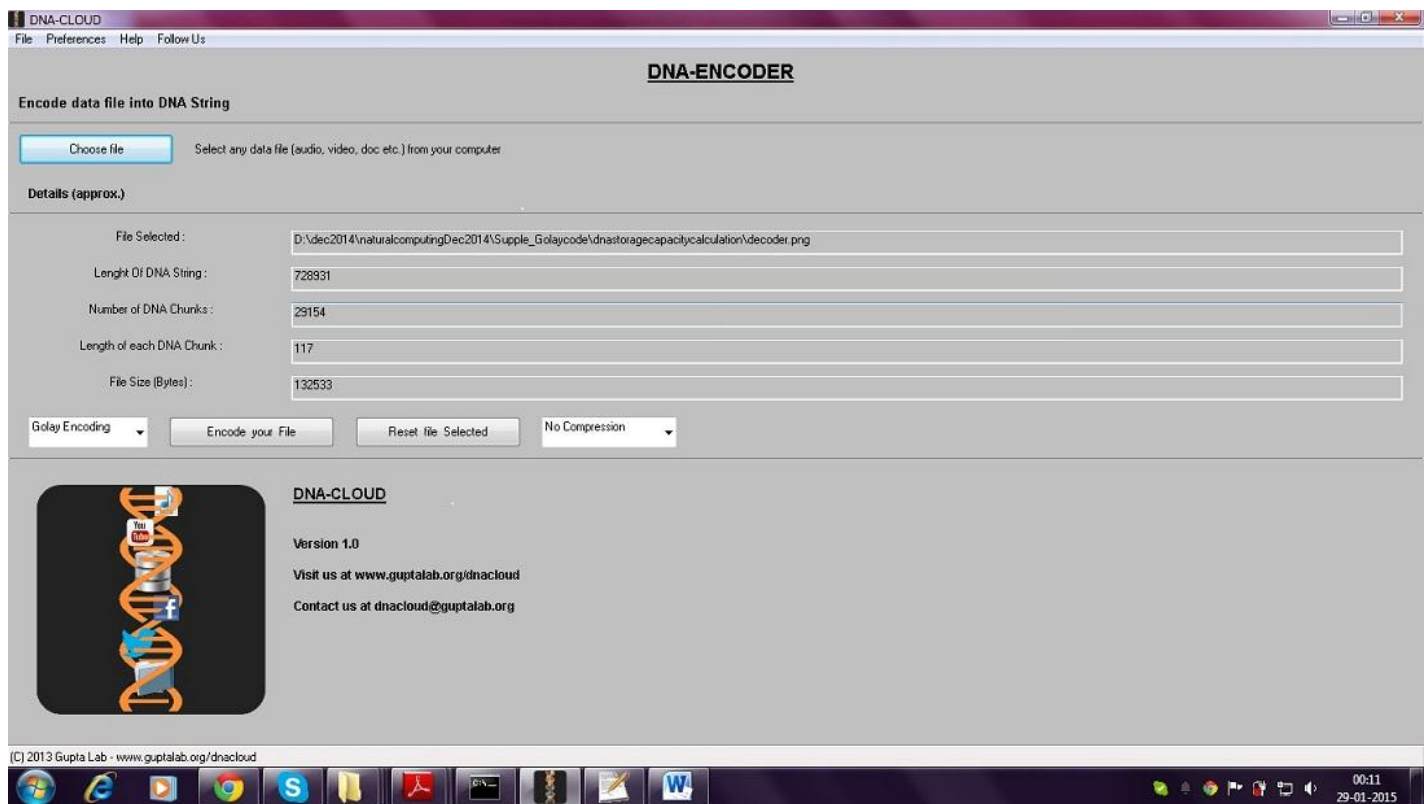


Figure 1 File to DNA (Encode)

## 2. Decode

This option is available under the file menu which will retrieve the data stored in DNA.

When the user clicks on this option, the screen in as shown in Figure 2 appears. User can enter the DNA string from which data is to be retrieved in the text box against **Please write DNA string** option. **Select .dna file** option. Once the file to be decoded is selected, click on **Decode Selected file** option. To save the decoded file, give the name and save the file at specific location. This will generate the original file that was encoded in DNA. **RESET** button can clear the selected file and select new file by this option. It is like clear button.

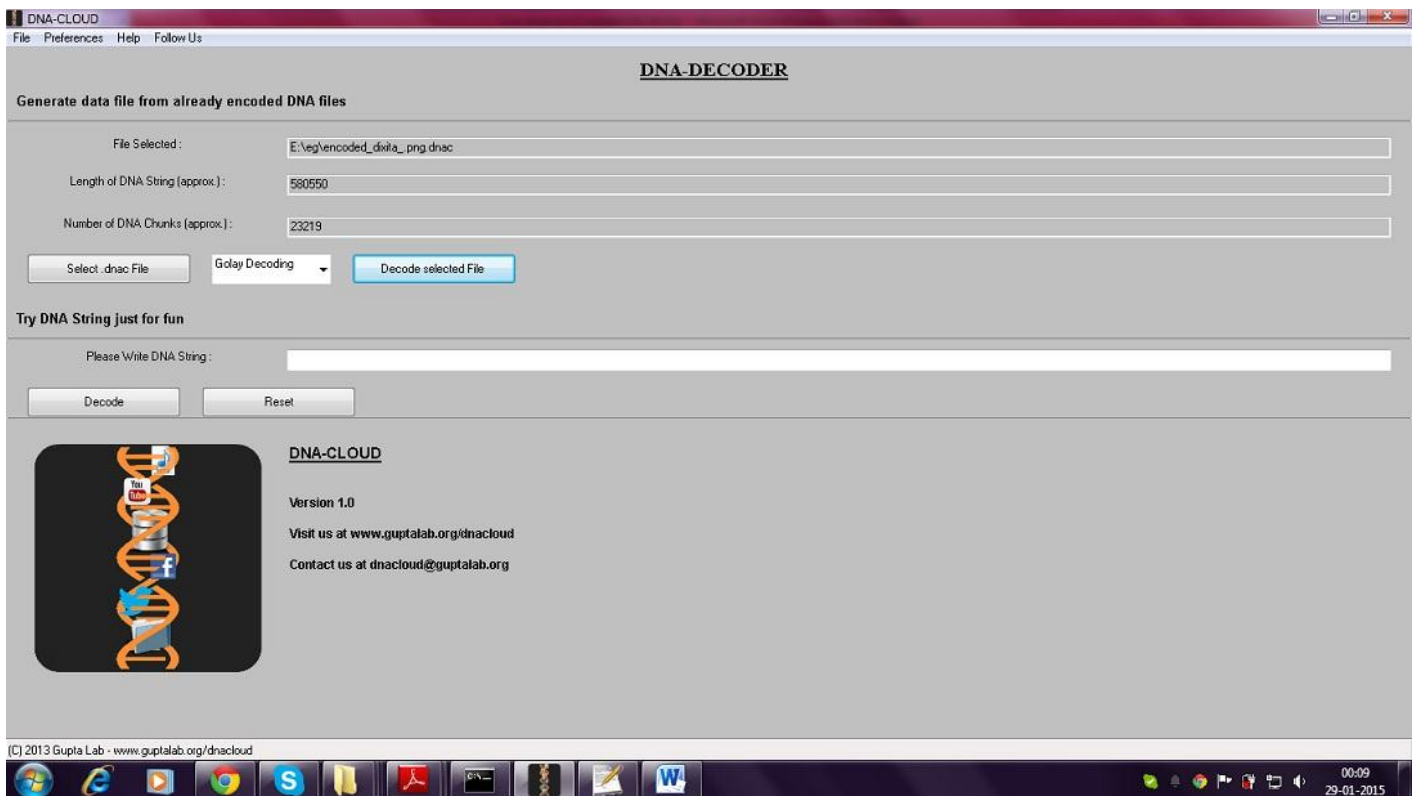


Figure 2 DNA to File (Decode)

### 3. Estimator

This will give various statistics for encoded file and biochemical properties of DNA of the decoded file. Estimator has two main sections.

#### **Memory Required**

User can select the file to be encoded from the system and the following values of the file are estimated as given in Figure 3. These values are approximated.

- File size in bytes
- Size of DNA string
- Free Memory Required

- To generate DNA chunks
- To generate DNA strings
- Amount of DNA Required

## Biochemical Properties

This will estimate the biochemical properties of the DNA sequence used to store the data. Select .dna file from the system which contains the DNA sequences to estimate the properties. This will take salt concentration (mM) and cost per base as an input. **Calculate** option will estimate the GC content of DNA, melting temperature of the DNA and total cost to store the file in DNA as shown in Figure 4. All the values are approximated. **Save** option can be used to save estimated information.

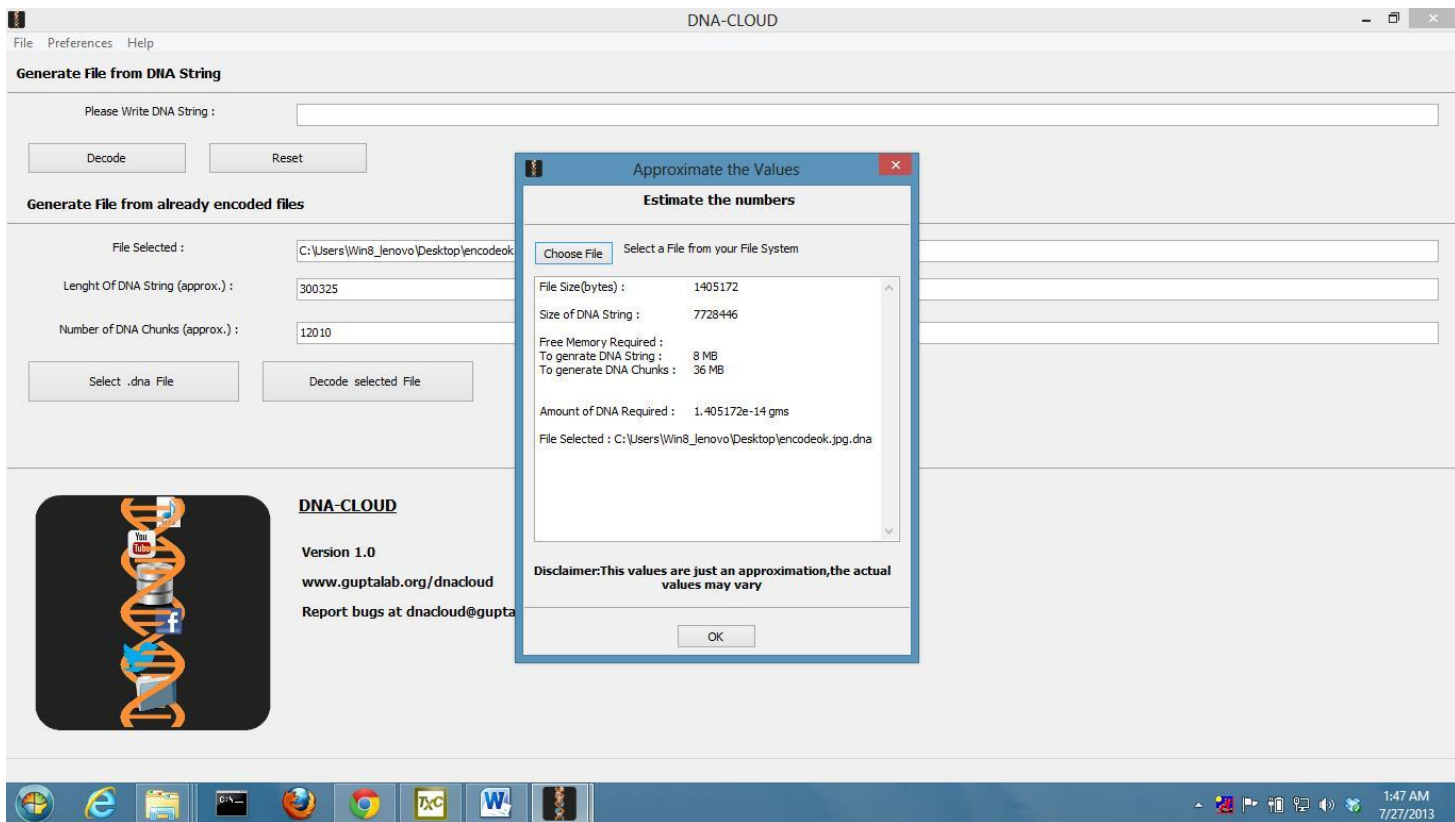


Figure 3 Estimate of Memory required for encoding the file to DNA



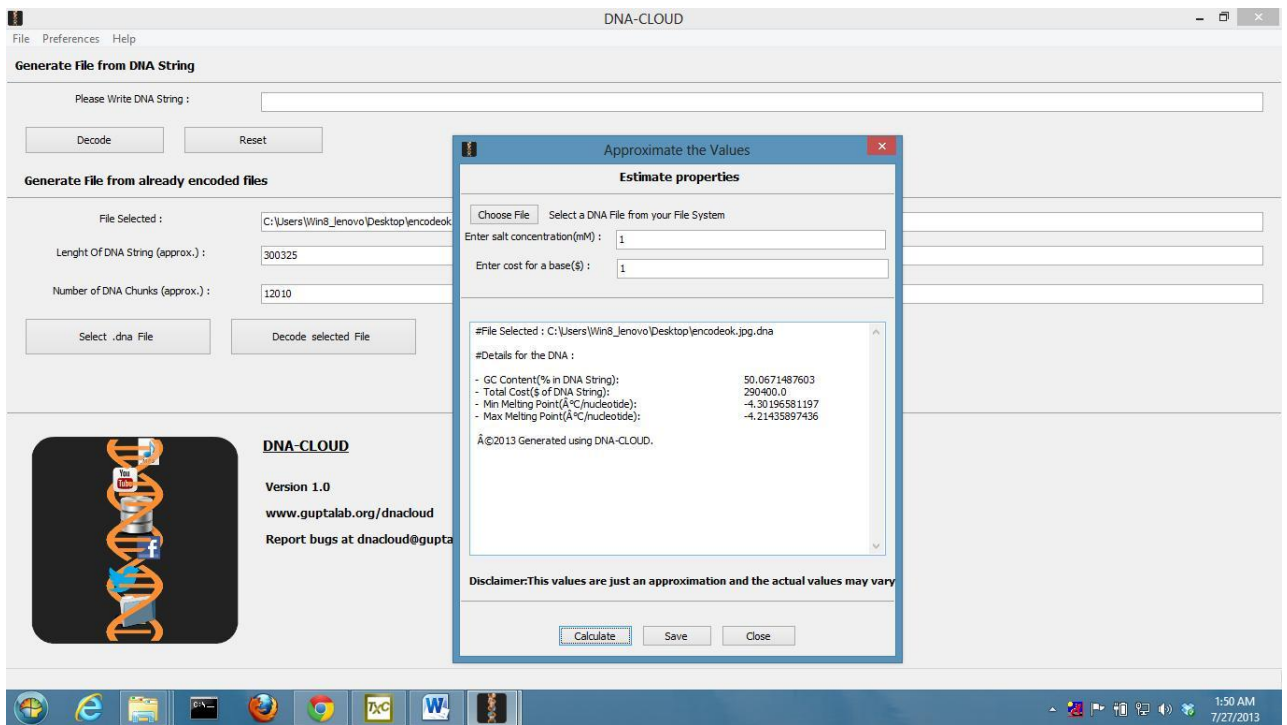


Figure 4 Estimate of Memory required for encoding the file to DNA

#### 4. Export

This will help the user to export the file generated to different formats. DNA strings can be exported to file format that can be used as input for the synthesizer. File can be exported to FASTQ file format that is required by the sequencer. These options are available in File menu. This will generate the feasible output of the DNA strings that is to be used by respective machines. To export the DNA file for the synthesizer to synthesize the DNA, use **Export DNA synthesizer File** option. To decode the file stored in DNA, use **Import DNA sequencer file** option to get the DNA sequences to be decoded. Above mentioned export options are under development, it will be available in next version. The .dna file can be exported to PDF with all software output details in single PDF by using the option **Export .dna to .PDF** in the File menu as shown in Figure 5.

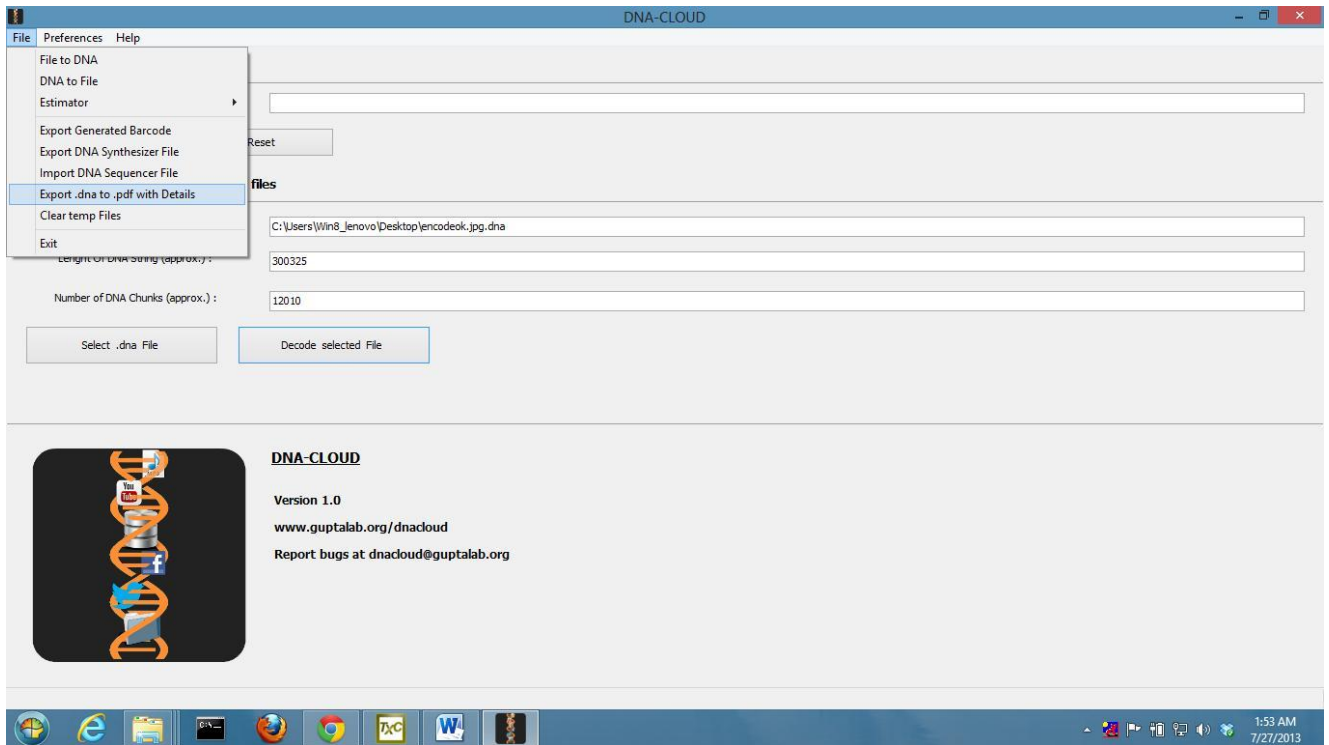


Figure 5 Export .dna to PDF

#### 5. Clear temp files

This will clear all your history of the software. It will remove all the files generated by the software.

#### 6. Exit

Exit will help to quit from the software.

#### 7. User details

This option is available in preferences menu. This is mandatory to make the data of the user stored in the DNA to avail only to the respective user. For the data security, user has to feed his details in the dialogue box as given below. Once the user feeds the data as shown in Figure 6, the dialogue box to enter the password appear. This help to user to retrieve his files converted to DNA safely. If user does not feed the information, the software will remind it again and dialogue box will appear again and again till the user enters the details. Barcode for the user information is generated by which one can retrieve the data easily. It includes the information like name of the user, contact number, email address. **Generate Barcode** will generate barcode for the user details along with the filename stored. Barcode can be exported to other image formats by using **Export Generated Barcode** option in File menu.

These barcode generated can be used by biotech companies to tag the DNA on which the particular file is to be stored.

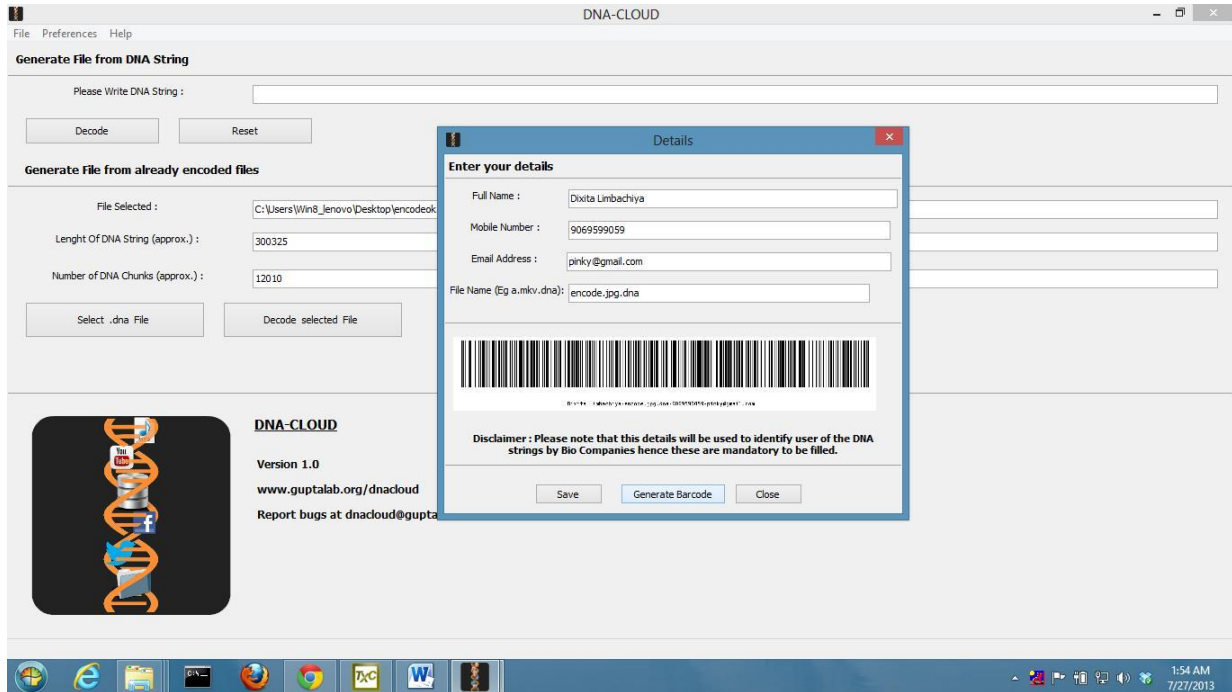


Figure 6 User details and barcode generated.

## 8. Help Menu

“Help” menu which tells about the software and developers. There are four options in the Help Menu.

1. User Manual
  2. Product Demo
  3. Product Feedback
  4. About Us
- First option will open default User Manual document in user's default PDF reader software
  - Second option product demo will open the homepage of DNA Cloud where the demo is available.
  - Third option will open the feedback form where you can give feedback for the software.

Finally about option recognize the contributors in this project. There is a dialog box that opens which contains information like logo of the software, version of the software, name of the software, License of the software, Credits button and URL of the software. On pressing “Credits”, it opens a PDF document for the credits.

## Support and Feedback

Users are requested to contact team at the email [dnacloud@guptalab.org](mailto:dnacloud@guptalab.org) for feedback and any other issues with the software. Two platform specific installers (Windows and Mac and Linux) are available on the project home page along with source code with open source license agreement.

## References

1. Goldman, Nick, et al. "Towards practical, high-capacity, low-maintenance information storage in synthesized DNA." *Nature* (2013).
2. Church, George M., Yuan Gao, and Sriram Kosuri. "Next-generation digital information storage in DNA." *Science* 337.6102 (2012): 1628-1628.

