一种基于重组 DNA 技术的密码方案

饶妮妮
(电子科技大学生命科学与技术学院, 四川成都 610054)

摘 要: DNA 密码体制的设计、分析和应用在国际上尚处于探索性研究中. 本文提出了一种基于重组 DNA 技术的密码体制方案; 设计了新密码体制的加/ 解密方法以及 DNA 序列的数字编码方法; 分析了新密码体制的保密强度. 本文工作具有一定的原始创新性.

关键词: DNA 重组技术; 加密; 解密; 密钥; 保密强度
中图分类号: TN918 文献标识码: A 文章编号: 0372-2112 (2004) 07-1216-03

A Cryptosystem Based on Recombinant DNA Technique

RAO Ni-ni
(School of Life Science and Technology, UESTC, Chengdu, Sichuan 610054, China)

Abstract: The design, analysis and application of DNA cryptosystem just begin to be explored all over the world. A cryptosystem based on recombinant DNA technique is presented. The encryption/ decryption and digital DNA coding are designed. The security is analyzed. This work is inventive.

Key words: Recombinant DNA technique; encryption; decryption; key; security

1 引言

在信息时代, 知识即意味着财富和实力. 伴随着信息技术与信息产业的发展, 网络与信息安全问题及其对经济发展、国家安全和社会稳定的重大影响, 正日益突出地显现出来, 受到越来越多国家的关注. 密码技术是信息安全技术的核心, 它的主要任务是寻求产生安全性高的有效密码算法和协议, 以满足对消息进行加密或认证的要求. 学者们把密码理论与技术分成两大类, 一类是基于数学的密码理论与技术, 包括公钥密码、分组密码、序列密码、密钥管理、认证码、数字签名、Hash 函数、身份识别、PKI 技术、VPN 技术等[1~3]; 另一类是非数学的密码理论与技术, 包括信息隐藏、量子密码、基于生物特征的识别理论与技术等[4~6]. 随着计算技术和数学理论的发展, 基于数学的密码技术受到了威胁, 如 56bit 分组密码体制 DES 已被破译, 著名公钥体制 RSA 的安全受到威胁. 于是学者们希望寻求更安全更简单的加密体系. 在此同时基因工程等生物技术获得了飞速发展[7,8], 尤其是人类基因组测序计划完成后, 产生了数量巨大的 DNA 序列, 这诱发了人们用生物技术的方法对信息进行加密的思想[9~12]. 目前已有学者提出了 DNA 加密技术[13], 本文则提出一种基于重组 DNA 技术的加密方案. 本文的目的除了寻求学者们对该密码体制进行测试、分析和评估以外, 更重要的是想以此推动基于生物技术的密码体制的研究进程.

2 重组 DNA 过程与加/ 解密过程

重组 DNA 过程如图 1 所示, 其步骤如下[14]:

(1) 从复杂的生物有机体基因组中, 经过酶切消化或 PCR 扩增等步骤, 分离出带有目的基因的 DNA 片段;

(2) 在体外, 将带有目的基因的外源 DNA 片段连接到能够自我复制的并具有选择记号的载体分子上, 形成重组 DNA 分子;

(3) 将重组 DNA 分子转导入适当的宿主细胞, 并与之一起增殖;

(4) 从大量的细胞繁殖群体中, 筛选出获得了重组 DNA 分子的宿主细胞; 从这些筛选出来的宿主细胞中, 提取已经得到扩增的目的基因.

常规密码系统的加密过程是:

$$C = E_K(p) \tag{1}$$

其中, E_K 是实现加密变换的带有参数 K 的加密变换函数, 它把明文 p 从明文信息空间 SP 变换到密文信息空间 SC, 既

$$E_K: SP \to SC$$

解密过程是:

$$D_K(C) = D_K(E_K(p)) = p \tag{2}$$

其中, D_K 是解密变换函数, 实现把密文信息空间 SC 逆变换到明文信息空间 SP:

$$D_K: SC \to SP$$

收稿日期: 2002-10-22; 修回日期: 2003-06-30
基金项目: 电子科技大学青年科技基金(No. YF021201); 教育部科学技术重点研究项目(No. 03146)

E_K 和 D_K 是一对可逆变换. 在不同的密码体制中, 如 DES、RSA、椭圆函数等, E_K 和 D_k 是不同意义的数学变换.

比较 DNA 重组原理与加/解密原理不难发现, 从目的基因到重组获得环状 DNA 分子, 再经筛选和分离、纯化获得目的基因的过程实质就是从明文加密获得密文, 再解密密文获得明文的过程, 因而可以利用重组 DNA 技术构造密码体制.

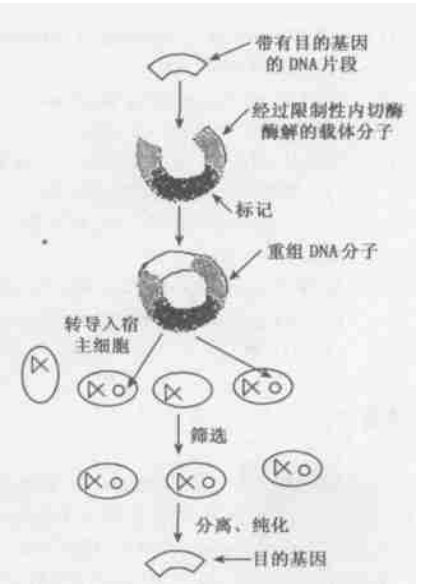

图 1　重组 DNA 过程示意图

3　新密码体制设计

311　加/解密设计

新密码体制加/解密方案设计如下:

加密过程方框图如图 2 所示. 首先将明文信息编码成目的基因, 通过重组 DNA 实验, 把带有目的基因的 DNA 片段及载体分别选择合适条件(该条件作为密钥)进行剪接处理, 把带有目的基因的 DNA 片段与载体连接获得重组环状 DNA 分子(载体性质和连接位置作为密钥), 该重组环状 DNA 分子即为密文; 与此同时还有大量其他多种类型的 DNA 分子形成(如仅由载体自身再连接形成的环形 DNA 分子和多聚 DNA 分子等), 再将各种类型的环状 DNA 分子进行二进制代数编码, 形成适合于通信信道传输的二进制信息格式后, 通过通信信道传送到接收端.

图 2　加密过程方框图

为了确保加密过程的安全性, 编码明文信息的目的基因应作为私有密钥进行保密, 实现重组 DNA 的实验条件(包括载体质粒性质和接入位置等)作为加密密钥. 为便于译码, 编码方法或规则给予公开. 加密密钥通过安全通道传送到接收端.

解密过程方框图如图 3 所示. 首先将接收的二进制密文信息进行代数解码, 获得多种类型的环状 DNA 分子. 利用来自于发送端的实验条件, 通过分子生物学实验将各种类型的 DNA 分子转导入宿主细胞, 通过在合适条件下培养, 既可大量扩增、繁殖, 若培养条件选择得适当, 则只有含目的基因的细胞才能生长, 进一步再用像颜色反应这样的指标, 对能生长的细胞进行筛选, 就可筛选出仅仅含有目的基因的重组质粒的细胞. 经过分离纯化等实验, 提取出目的基因, 将目的基因译码便获得明文信息.

图 3　解密过程方框图

接收端使用的解密密钥也是重组实验条件, 可见加/解密密钥相同. 因此, 新密码体制是一种对称密码.

312　代数编码方法

现有的 DNA 序列代数编码方法较多, 如 AA、TT 法等. 经过分析, 本文提出一种简单的适合于 DNA 序列的代数编码方法[15]. 编码原理如下:

单链 DNA 序列由四种碱基 A、C、G、T 组成, 其中 A 与 T、C 与 G 互补. 有关定理表明: 设 p 是任一素数, 而 n 是任一正整数, 那么总存在着一个恰含 p^n 个元素的有限域. 若设 p= 2, n = 2, 则 GF(2^2) 是有限域. 因此, 用 00、01、10 和 11 分别编码 DNA 序列中的 A、C、G、T, 就可得到有限域 GF(2^2) 上源于 DNA 序列的一类二进制序列, 为区别普通的二进制序列, 称该序列为 DNA 编码序列.

4　保密性分析

新密码体制不同于常规密码体制之处在于: 除要进行数学变换以外, 还需要依赖于分子生物学实验来实现信息的加密/解密变换, 加密过程需要进行两次编码. 因此, 新密码体制实质是一种二次加密体制. 第一次加密是将明文信息变成目的基因, 此目的基因作为私有密钥严格保密. 密码分析者要想破译出此目的基因首先需要具备生物技术知识和相关的分子生物实验条件, 再从数量巨大的 DNA 序列数据库中依靠生物技术实验检测出基因序列, 再用穷举搜索法从种类繁多的基因序列中猜测目的基因. 从实际保密的角度来看, 分析者成功破译获得目的基因的概率几乎是零. 第二次加密是基于重组 DNA 原理的分子生物技术实验, 既将目的基因序列转变为重组的环状 DNA 分子作为密文公开发送. 由于大量其他类型的环形 DNA 分子与密文 DNA 分子同时产生, 对密文起到了很好的隐藏作用, 再加上重组实验条件作为密钥进行保密, 所以在此条件下密码攻击者要想通过生物技术实验获得准确的目的基因的机率是很小的. 通常, 分子生物学的实验设备较为昂贵, 这对密码攻击者也是一种制约. 进一步, 第二次加密完全依靠生物学理论和技术实现, 因此, 建立在数学理论基础上的密码分析方法(如强力攻击、差分密码分析、插值攻击、错误攻击、定时攻击等)对本次加密失去作用, 需要重新寻求破译第二次加密的密码分析方法. 不难预测, 新密码体制具有很强的保密强度.

5　结束语

提出基于重组 DNA 技术的密码方案虽然是一项具有原创性的重要研究工作, 但却仅仅是一个完善加密体系构建的第一步. 关于该密码体制仍然有许多研究工作有待进一步开展, 例如, 它的安全性分析、密钥管理、实现原理以及应用前景等. 为了充分论证新密码方案的可行性, 我们真诚希望同行学者对此密码方案进行测试、分析和评估.

参考文献:

[1] SCHNEIER B. Applied Cryptography, Protocols, Algorithm and Source Code in C[M]. New York: John Wiley and Sons, 1996.

[2] ADAMS C, LLOYD S. Understanding Publickey Infrastructure: Stan2dard and Deployment Considerations[M]. Indiana: Macmillan Technical Publishing, 1999.

[3] 卢开澄. 计算机密码学(第 2 版) [M]. 北京: 清华大学出版社, 1998.

[4] KATZENBEISSER S, PETITCOLAS F A P. Information Hinding Tech2niques for Steganography and Digital Watermarking [M]. Boston: ARTECH HOUSE, 2000.

[5] BENNETT C H, BESSETTE F, BRASSARD G, et al. Experimental quantum cryptography[J]. J Cryptology 1992, 5(3): 3- 28.

[6] EKERT A K. Quantum cryptography bases on Bell s theorem[J]. Phys Rev Lett, 1991, 67: 661.

[7] 寿天德. 现代生物学导论[M]. 合肥: 中国科技大学出版社, 2001.

[8] 美国国家生物医学中心[DBOL] http://www.ncbi.nlm.nih.gov

[9] D. Anastassiou. Genomic Signal Processing[J]. IEEE Signal Processing Magazine 2001, 18(4): 8- 20.

[10] D. Anastassiou. Frequency2domain analysis of biomolecular sequence [J]. Bioinformatics, 2000, 16(12): 1073- 1081.

[11] B D Silverman, et al. A measure of DNA periodicity [J]. J Theor Biolo2gy, 1986, 65: 295- 300.

[12] W Li, et al. Understanding long2Range correlation in DNA sequence [J]. Phys D, 1994, 75: 392- 416.

[13] 陈国龙, 刘传才. DNA 加密技术与数字水引技术的综合集成[J]. 引进与咨询, 2000, (2): 57- 59.

[14] 吴乃虎. 基因工程原理(第 2 版上册) [M]. 北京: 科学出版社, 2001.

[15] Xu Weihua. Et al. Molecular characterization of the gene encoding the precursor protein of DH2PBAN[J]. Biochim Biophys Acta, 1995, 1262(1): 83- 89.

作者简介:

饶妮妮 女, 1963 年 5 月生于四川宜宾, 教授, 博士生导师, 1983 年、1989 年在电子科技大学获工学学士、硕士学位, 1997-1998 为英国百拉德福德大学电子工程系访问学者, 在国内外发表涉及信息加密技术、移动通信以及生物医学信号处理的论文 30 余篇. 主持参与科研和教学研究项目 10 余项. 现在的研究方向: 信号/图像处理、生物信息学、移动通信.