面向复杂模型 java编程在生命起源模拟中的应用

概要：

1.Template-dependent RNA synthetase ribozyme 模板依赖的RNA合成核酶，可以催化自身的复制，即RNA复制酶；被认为可能是RNA世界最先出现的，但是1.长序列 2.缺少令人信服的拥有自我偏爱特性的机制(self-favoring feature)

而且最近有研究表明，非酶催化的模板依赖的RNA合成也是有效的

2.我们假设nucleotide synthetase ribozyme 核苷酸合成核酶是第一个出现的

3.RNA复制酶可能是后出现的，可能在protocells原初生命体之后

Introduction：

生命从非生命的条件下最初诞生，但我们无法充分定义primary life是什么样的

现代生命，的相互依赖性：DNA复制依赖于蛋白质催化，蛋白质合成依赖于DNA为基础的编码(DNA-based coding)

RNA世界的假设，在最初的生命形式里，genetic and catalytic functions 遗传和催化两个功能都由RNA实现

RNA世界的假设提供了早期生命的分子特征模型，但其细节仍待商榷，比如它是不是cellular的，还是只是naked RNA分子？

Cellular细胞形态对最初形成生命来说太复杂了，所以假设细胞形态是通过裸分子达尔文进化演化而来的比较合理

能自我复制的裸RNA分子可能从无生命环境出现，并有着进化到细胞形态的潜力；这种裸RNA分子是什么，identify鉴定它就是我们说的the problem of the RNA world

我们合理假设裸RNA分子诞生的nonlife background是一个核苷酸池

传统观点认为RNA复制酶是第一个诞生的，因为非酶催化的RNA复制似乎不足以有效地支持复制任意RNA，没有复制，达尔文进化不可能发生

传统观点有两大缺陷：

1.RNA复制酶太长了（160-200nt），在最初的生命诞生有些苦难，并且因为太长，很难通过自催化来自我复制

2.不在细胞环境中，复制酶需要足够self-favoring（来克服parasite寄生虫的问题），来确保它能在核苷酸池中繁荣；

它们可能通过’tags’标签机制或依赖于limited dispersal(perhaps on mineral surfaces) 在矿物表面的有限的扩散，识别彼此；但这些机制不是很可靠；

于是我们假设可以分子内复制，但还是太长了

最近的研究显示，在分子伴侣的帮助下，非酶催化的模板依赖的RNA合成可以更高效，并且这些分子伴侣被支持存在于早期的地球。

所以所有的RNA都可以通过非酶催化自复制。

基于非酶催化假设，核苷酸池中的所有RNA都要竞争核苷酸来复制自己。

那么如果有种RNA可以催化核苷酸的合成，那这种RNA的附近就会有相对浓度更高的核苷酸。

核苷酸的合成可能是很多反应的关键过程，比如很多种核酶的合成

这种RNA我们叫Nucleotide Synthetase Ribozyme核苷酸合成核酶，这种酶也许是第一个出现的，可以提高限速步骤。

当NR的数量达到一定规模之后，达尔文进化就可能出现，可能促进其他核酶的出现，比如说一些促进膜的形成的核酶，比如两亲分子合成核酶Amphiphilic molecule synthetase ribozyme（AMR）。它们一起合作，protocell原初生命体就可能出现。在原初生命体里，AMR可以利用NR催化出来的核苷酸，NR在膜包围中，可以比在裸露状态下更高效地催化核苷酸合成。RNA复制酶可能在细胞形态中出现。

用蒙特卡洛模拟，探索NR繁荣的possibilities和conditions

COMPUTER SIMULATION

模拟方法：蒙特卡洛（类似于cellular automata approach）

系统背景：

二维表面，NxN个格子，raw materials, nucleotides, RNA molecules在其中。

同个格子的材料可以互相作用，也可以窜到临近格去。

拥有特殊结构域的RNA可以起到NR的作用。

系统中的每种行为都设定了发生的概率

这种概率的设定本应该基于现实条件的，但是由于我们对远古时代的地球环境不甚了解，所以更有效的方法是基于基本关系先假设一组大概的概率值，然后再定量大概假设，通过模拟测试，逐步找到能够成功使NR繁荣的那组概率值。然后就可以通过控制变量研究特定的变量对NR繁荣的影响。

变量：

N[10-30]：grid的边长

TM[10^4,10^5]：Total materials 系统中的总材料数（最开始等于核苷酸前体的总量）

LRD：Length of the nucleotide synthetase ribozyme-characteristic domain (in nucleotides) [6, 16] NR的结构域的长度

PNF：Probability of nucleotide formation (not catalyzed by ribozyme) [10^-4, 10^-3] 非酶催化核苷酸前体形成核苷酸

PNFR：Probability of nucleotide formation under the catalysis of the nucleotide synthetase ribozyme [0.1, 0.9] 在酶的催化下形成核苷酸

PND：Probability of nucleotide decay [10^-3, 0.01] 核苷酸降解

PRL：Probability of the random ligation of nucleotides and oligonucleotides [10^-6, 10^-5] 核苷酸/寡核苷酸的随机链接

PBB：Probability of breaking a phosphodiester bond [10^-6, 10^-5]磷酸二酯键断裂

PAT：Probability of attracting a substrate by a template [0.1, 0.9] 通过模板吸引到底物的概率

PFP：Probability of the false base-pairing [0.01, 0.1]碱基配对的错配率

PLT：Probability of nonenzymatic template-directed ligation [0.1, 0.9]非酶催化的模板配对连接概率

PSP：Probability of the separation of a base pair [0.1, 0.9] 碱基配对之后分离的概率

PMV：Probability of the movement of raw material to an adjacent cell. [0.01, 0.1] 原材料移到临近格的概率

PMN：核苷酸移到临近格的概率Probability of the movement of nucleotide, is assumed to be half of the PMV.

RNA移动到临近格的概率：Probability of the movement of an RNA unit is assumed to be PMN/m^1/3, where m is the mass of the unit,relative to a nucleotide.

PNFR >> PNF 有NR时的核苷酸形成概率 远大于 没有的时候

PLT >> PRL 非酶催化的模板配对链接概率 远大于 核苷酸或寡核苷酸在实验室条件下的随机连接

PND >) PNF 核苷酸降解概率 略大于 无酶催化下核苷酸的形成概率

PBB ≈ PRL 磷酸二酯键断裂概率 约等于 核苷酸或寡核苷酸在实验室条件下的随机连接

PAT ≈ PSP 通过模板吸引到底物的概率 约等于 配对之后分离的概率

PFP 与 replicating fidelity复制精度有关 Fidelity = 1/(3\*PFP+1) 所以PFP不能太小

这只是个大概的范围

系统中会产生的行为：

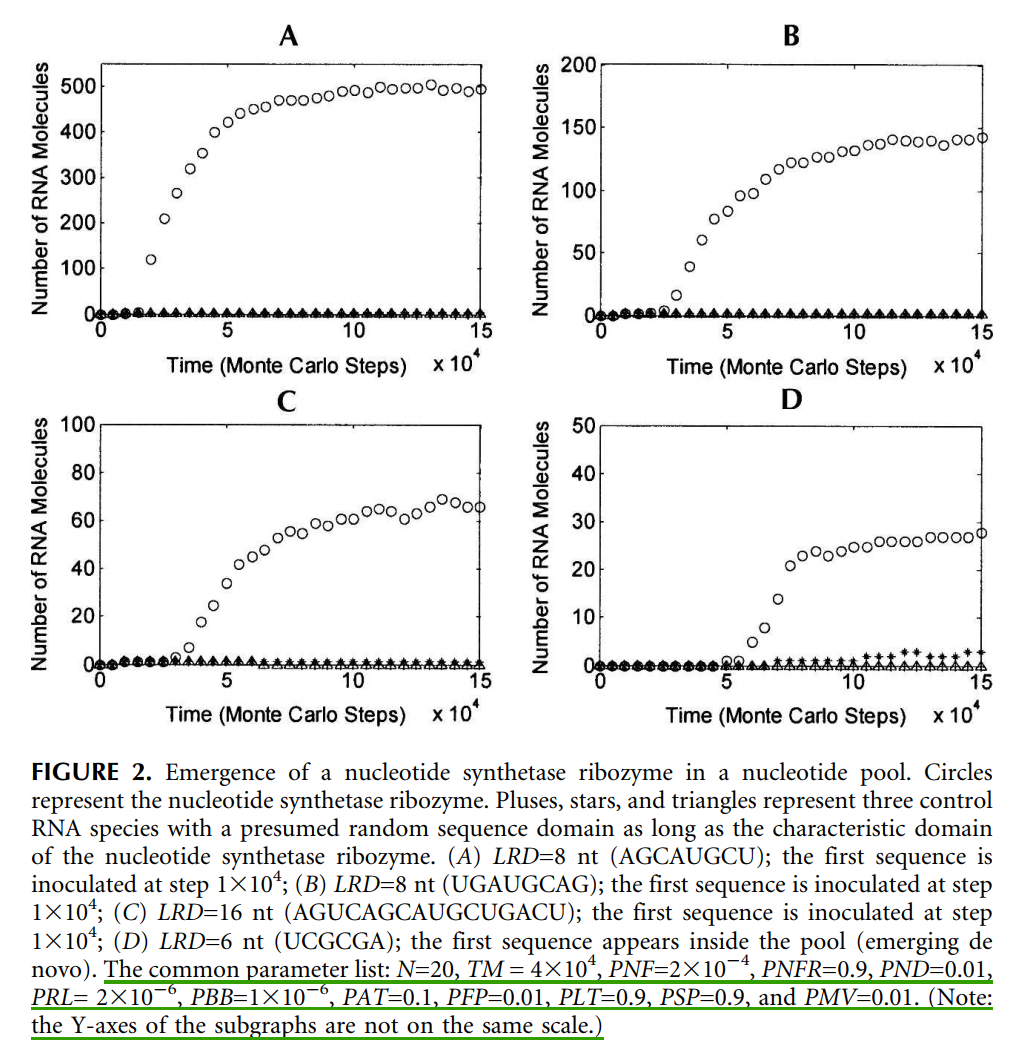
Raw material 会形成核苷酸 在没有酶的帮助下低概率形成 有NR的时候大概率形成

RNA分子通过核苷酸的随机连接形成

RNA分子通过非酶催化的模板定向连接进行复制

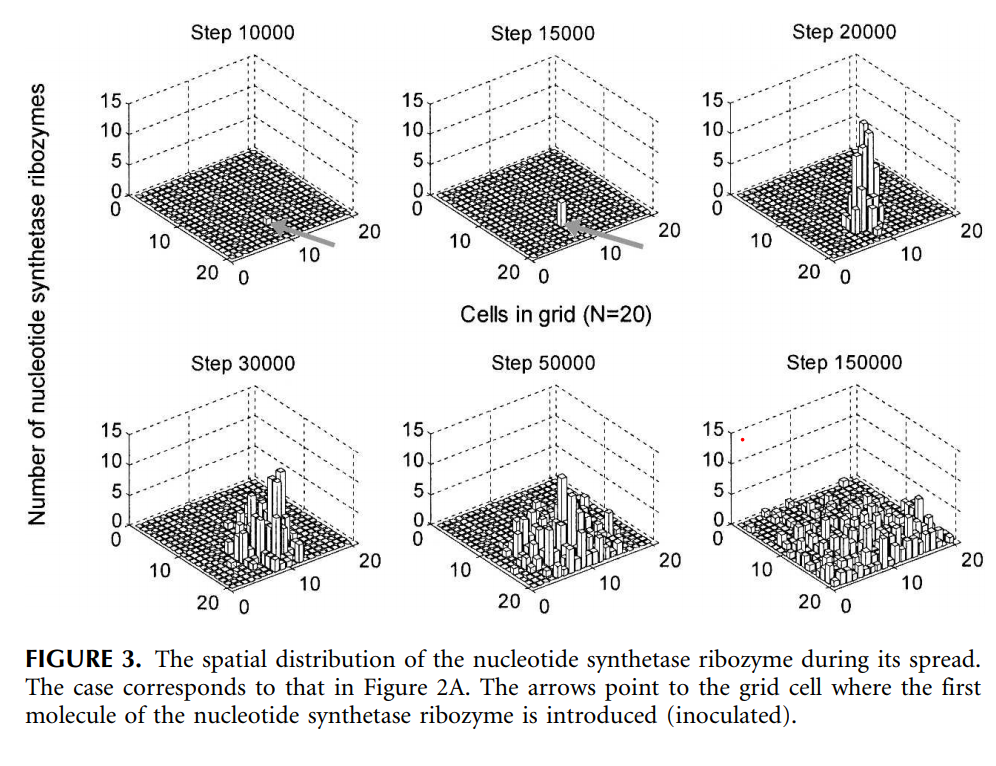
RESULTS

NR可以在系统中繁荣，结构域可以回文（更容易繁荣because each turn of template copying produces a copy of the ribozyme at once?），也可以不回文（更难形成），长结构域的传播效率低，如果核酶最开始没被注入到系统中，也可以从头合成

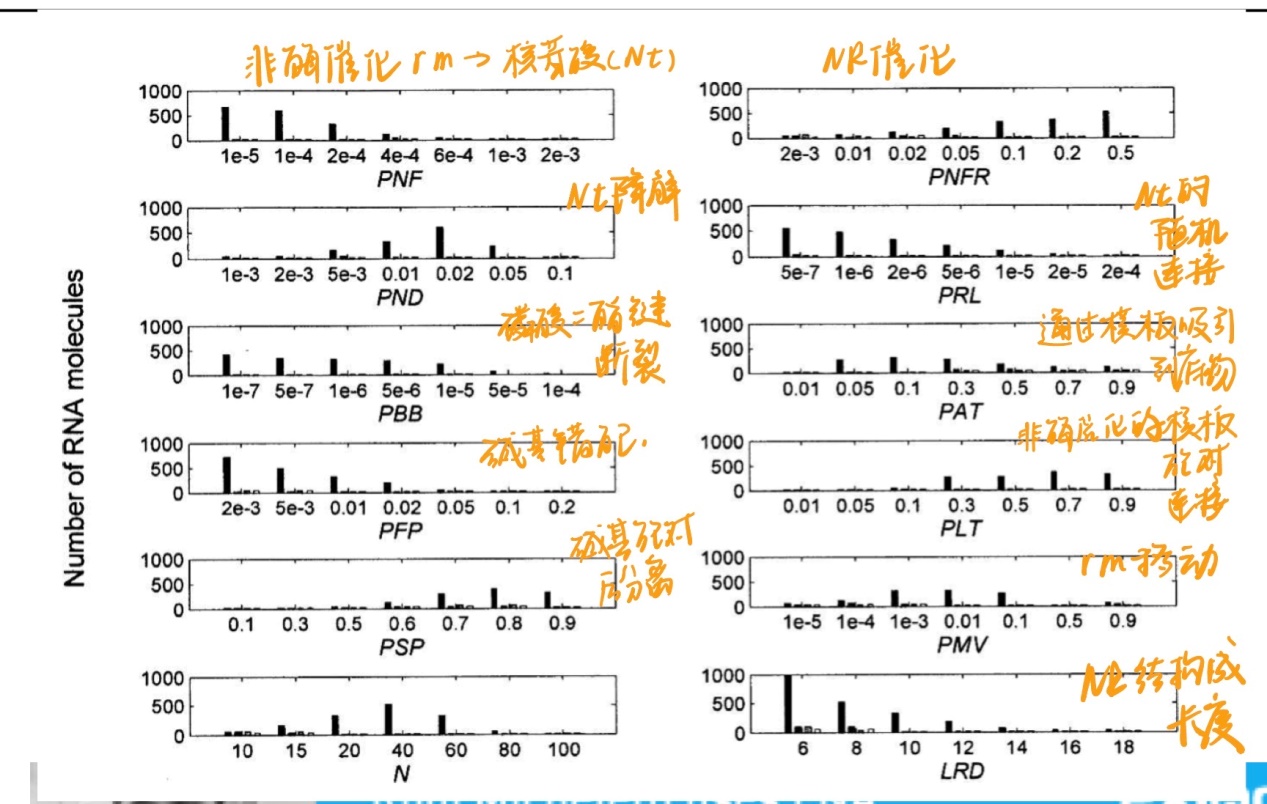


A：回文序列；B：非回文序列；C：很长的序列；D：不注入，从头合成

（Figure3）NR的空间分布：从第一个NR分子开始向周围扩散



（Figure4）dynamics of the spread of NR 各个变量对NR值的影响



第一个bar是NR，剩下三bar是对照组

Discussion

Analysis of the results

NR的传播满强大的，关键事件是第一个NR的出现，尽管能通过随机连接从头合成NR，但效率很低

随机连接的效率提升会导致spread的效率变低，也许是因为random ligation效率提升会导致更多random RNA species的出现，竞争有限的核苷酸

Limiting factor: resources的缺失

NR需要出现appearance并传播spread 但似乎利于出现的环境不利于传播，利于传播的环境不利于出现；

因此在我们的模型里，NR可能是在rocky bottom（有利于出现）中出现，并被释放到pool中（有利于传播）

RNA的移动：不能太活跃，这样会不利于利用它自己催化出来核苷酸，也不能太不活跃了，这样会不利于它利用临近格的RM；

一个好的解决方法是protocell的出现，在这里核苷酸的扩散是被闲置的，而cell可以自己移动到临近格去捕捉别处的RM；这样就会导致进化倾向：cell membrane，也许是两亲分子合成酶