

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه قم

دانشکده علوم پایه

پایان نامه دوره کارشناسی ارشد ریاضی کاربردی

عنوان:

شکل‌های ساختارهای شبه گره RNA

استاد راهنما:

دکتر غلام حسن شیردل

استاد مشاور:

دکتر علی اصغر فروغی

نگارنده:

محمد رضا محقق

زمستان ۱۳۸۹

تقدیم به:

مادرم فرشته خالصم و پدرم که فروغ مهرشان صیقل روحم است و توصیفشان را سینه ای به وسعت دریا باید تا بتوان اندکی از مهربانی بسیارشان را در کلام وصف آورد.

تشکر و قدردانی

حمد و سپاس خدای را که توفیق کسب دانش و معرفت را به ما عطا فرمود. در اینجا بر خود لازم می دانم از تمامی اساتید بزرگوار، به ویژه اساتید دوره کارشناسی ارشد که در طول سالیان گذشته مرا در تحصیل علم و معرفت و فضائل اخلاقی یاری نموده اند تقدیر و تشکر نمایم.

از استاد گرامی و بزرگوار جناب آقای دکتر شیردل که راهنمایی اینجانب را در انجام تحقیق، پژوهش و نگارش این پایان نامه تقبل نموده اند نهایت تشکر و سپاسگزاری را دارم.

از جناب آقای دکتر فروغی به عنوان استاد مشاور که با راهنمایی خود مرا مورد لطف و عنایت خویش قرار داده اند صمیمانه تشکر و سپاسگزاری می کنم.

از اساتید گرامی و بزرگوار جناب آقای دکتر فرجامی و آقای دکتر احمدی نیا که نظارت پایان نامه مرا بر عهده داشتند، کمال تشکر را دارم.

چکیده:

در این پایان نامه شکل‌های انتزاعی k - نامتقاطع و σ - کانونی ساختارهای شبه‌گره RNA را مطالعه می‌کنیم. V_k - شکل‌ها که تعمیم شکل‌های گیگریچ^۱ بوده، به منظور دسته‌بندی ساختار k - نامتقاطع شبه‌گره RNA توسعه داده شده‌اند و شکل خاصی از ساختارهای دوم RNA می‌باشند. با استفاده از یک تقریب جدید توابع مولد V_k - شکل‌ها را محاسبه می‌کنیم که برای پیش‌بینی ساختارهای RNA بسیار مفید هستند و اخیراً توجه بسیاری از دانشمندان را به خود جلب نموده‌اند. سپس دو نوع خاص از V_k - شکل یعنی IV_k^1 - شکل و IV_k^5 - شکل را بررسی می‌کنیم.

واژه‌های کلیدی: ساختار دوم RNA، شبه‌گره، ساختار k - نامتقاطع RNA، شکل.

فهرست مطالب

عنوان..... صفحه

مقدمه ۱

فصل اول:

مفاهیم پایه ای

۱-۱ ژن ۵

۲-۱ اصل بنیادی در زیست شناسی ۵

۳-۱ اسید نوکلئیک ۷

۴-۱ DNA ۸

۵-۱ RNA ۹

۶-۱ ساختمان RNA ۱۰

۷-۱ گراف ۱۰

۸-۱ متغیرهای مختلط ۱۲

۹-۱ سری تیلور ۱۴

۱۰-۱ سری لوران ۱۴

۱۱-۱ همگرایی مطلق و یکنواخت سری های توانی ۱۵

۱۲-۱ توابع مولد ۱۷

۱۳-۱ سری های متناهی مشتق پذیر ۱۸

۱۴-۱ نمونه ابر بحرانی ۲۲

۱۵-۱ کلاس های ترکیباتی ۲۳

فصل دوم:

شمارش ساختارهای RNA

۱-۲ مقدمه ۲۷

۲-۲ دسته بندی ساختارهای RNA ۲۷

۳-۲ از تابلو به گشت ۳۴

۴-۲ شمارش گشت های تصادفی ۴۰

فصل سوم:

شکل V_k

۱-۳ ساختار هسته ۴۸

۲-۳ شکل ها ۵۰

۳-۳ توابع مولد ۶۱

۴-۳ توابع مولد ساختارهای هسته ۶۱

۵-۳ توابع مولد ساختارهای k - نامتقاطع و σ - کانونی ۶۸

فصل چهارم:

شکل IV_k^1 و شکل IV_k^5

۱-۴ مقدمه ۷۹

۷۹.....	۲-۴ ترکیبیات IV_k^5 - شکل
۸۰.....	۳-۴ ترکیبیات IV_k^1 - شکل
۸۹.....	۴-۴ نتیجه گیری
۹۱.....	کارآینده
۹۲.....	برنامه ها
۹۵.....	منابع
۹۸.....	واژه نامه انگلیسی - فارسی

فهرست جداول

صفحه	عنوان
۴۴	جدول ۱-۲ تعداد ساختارهای k - نامتقاطع
۴۵	جدول ۲-۲ میزان رشد ساختارهای k - نامتقاطع
۵۹	جدول ۳-۲ رشد نمایی از V_k - شکل‌ها
۸۷	جدول ۱-۴ میزان رشد نمایی از IV_k^1 - شکل‌ها
۸۹	جدول ۲-۴ میزان رشد نمایی از ساختارهای دلخواه k - نامتقاطع و ۲-کانونی
۹۰	جدول ۳-۴ میزان رشد نمایی از ساختارهای دلخواه k - نامتقاطع و ۳-کانونی

فهرست شکل‌ها

صفحه	عنوان
۶	شکل ۱-۱ اصل بنیادی در زیست‌شناسی
۷	شکل ۲-۱ فرایند رونویسی
۹	شکل ۳-۱. ساختمان <i>DNA</i> و <i>RNA</i>
۱۱	شکل ۴-۱ اصل بنیادی در زیست‌شناسی
۲۸	شکل ۱-۲ دیاگرام‌های نامتقاطع
۲۸	شکل ۲-۲ دیاگرام k -نامتقاطع
۲۹	شکل ۳-۲ تطابق ۳-نامتقاطع
۳۰	شکل ۴-۲ عناصر ساختاری در <i>RNA</i>
۳۱	شکل ۵-۲ یک ساختار ۳-نامتقاطع ۶-کانونی
۳۳	شکل ۶-۲ ساختار دوم <i>RNA</i>
۳۳	شکل ۷-۲ ساختار شبه‌گره
۳۴	شکل ۸-۲ تابلو یانگ
۳۵	شکل ۹-۲ جدول نوسانی
۳۶	شکل ۱۰-۲ جدول یانگ استاندارد
۳۶	شکل ۱۱-۲ جدول یانگ استاندارد
۳۷	شکل ۱۲-۲ نمایش یک گشت
۴۰	شکل ۱۳-۲ تناظر بین دیاگرام‌های k -نامتقاطع و گشت
۴۲	شکل ۱۴-۲ اصل انعکاس
۴۸	شکل ۱-۳ ساختارهای هسته‌ای
۴۹	شکل ۲-۳ نحوه بدست آوردن ساختار هسته‌ای
۵۰	شکل ۳-۳ تولید V_k -شکل
۵۳	شکل ۴-۳ ۱-قوس‌ها
۶۲	شکل ۵-۳ از شکل‌ها به ساختارها

۶۴	شکل ۳-۶ نگاشت هسته ای
۶۴	شکل ۳-۷ شرح مرحله ۱
۶۶	شکل ۳-۸ شرح مرحله ۲
۶۹	شکل ۳-۹ نگاشت شکل
۷۰	شکل ۳-۱۰ انبساط شکل‌ها
۷۲	شکل ۳-۱۱ مرحله ۲
۸۱	شکل ۴-۱ نگاشت شکل

فهرست علائم

$H \rtimes K$	حاصل ضرب نیمه مستقیم H در K
$\det A$	دترمینان ماتریس A
$[n]$	مجموعه $\{1, 2, \dots, n\}$
$f(n) \square g(n)$	f و g به طور مجانبی هم ارزند
$I_\alpha(z)$	تابع بسل نوع مرتبه اول و از مرتبه α
$[z^n]f(z)$	ضریب z^n در بسط مک لورن تابع $f(z)$
$\exp(x)$	e^x

مقدمه:

موضوع این پایان نامه در حوزه ریاضیات زیستی قرار دارد که بر مدل سازی طبیعی فرآیند های زیستی با استفاده از ابزارها و روش های ریاضی مبتنی است. کاربرد ریاضیات در زیست شناسی تاریخ طولانی دارد اما اخیراً به دلایل زیر سروصدایی را در بین علاقمندان خود ایجاد کرده است:

- ۱- پیشرفت های اخیر ابزارهای ریاضی نظیر تئوری های بی نظمی، که به فهم مکانیزم های پیچیده و غیرخطی در زیست شناسی کمک زیادی می کند.
 - ۲- افزایش توان کامپیوتری که محاسبات و شبیه سازی ها را برای انجام آنچه که در گذشته ممکن نبوده، امکان پذیر می سازد.
 - ۳- افزایش گرایش به آزمایشات مجازی (فضای کامپیوتری) به سبب عواقب و مشکلاتی که ممکن است تحقیقات روی انسان یا حیوان داشته باشد.
- به علت تنوع دانش های تخصصی، تحقیقات ریاضیات زیستی اغلب با همکاری بین ریاضی دانان، فیزیک دانان، زیست شناسان، پزشکان، جانورشناسان، شیمی دانان و... انجام می گیرد.
- در زیر فهرست تعدادی از حوزه های تحقیقاتی ریاضیات زیستی آورده شده است:

۱- مدل سازی نورون ها و کارسینوژن ها

۲- مکانیک بافت های زیستی

۳- مدل سازی سرطان ها و شبیه سازی آن ها

۴- مدل سازی چرخه سلولی

۵- مدل سازی ریاضی محل التیام زخم

۶- شمارش ساختارهای RNA

پیش بینی ساختار RNA از اهمیت بسیار زیادی در علم بیوانفورماتیک برخوردار است و شمارش ساختارهای مختلف RNA، نقش کلیدی در الگوریتم های پیش بینی دارد.

چندین روش برای پیش بینی ساختارهای RNA با شرایط گوناگون موجود است که بسیاری از آنها مدل های آماری هستند که در مسائل ترکیباتی ریشه دارند و بسیار ساده تر از مدل های بر پایه انرژی آزاد شده می باشند. به همین دلیل ساختارهای ترکیباتی RNA در زیست شناسی بیشترین توجه را در ۳۰ سال اخیر به خود جلب کرده است.

اکثر روش های پیش بینی محاسباتی بر اساس مینیمم انرژی آزاد شده، تعداد چین خوردگی های زیر بهینه را محاسبه می کنند و باید ساختار اصلی را در میان تمام این ساختارهای دوم ممکن شناسایی کنیم. به همین دلیل بیشترین تلاش برای بدست آوردن یک پیش بینی خوب برای ساختار های دوم RNA صورت گرفته است. با استفاده از شکل های انتزاعی هر کلاس از ساختارهای دوم مشابه با یک شکل نشان داده می شود و ساختار اصلی را می توان در میان این اشکال یافت.

در فصل اول این پایان نامه، مفاهیم پایه ای زیست شناسی و ریاضی مورد نیاز برای فصل های بعدی را ارائه می کنیم. در فصل دوم ساختارهای مختلف RNA را تعریف کرده و سپس به بررسی و شمارش ساختارهای

k -نامتقاطع و σ -کانونی RNA می پردازیم. در فصل سوم، V_K -شکل و ساختار هسته را تعریف کرده و توابع مولد آنها را بدست می آوریم. در فصل چهارم ترکیبیات V_k^1 -شکل و V_k^5 -شکل ها را مورد بررسی قرار می دهیم.

فصل اول

مفاهیم پایه ای

در این فصل ابتدا مفاهیم زیست شناسی مانند ژن، نوکلئوتید، RNA و DNA را معرفی می کنیم که برای مطالعه بیشتر در این موارد به منبع [۱] می توانید مراجعه کنید. سپس مفاهیم ریاضی مورد نیاز برای فصل های بعدی مانند گراف، متغیرهای مختلط و ... را ارائه می کنیم. برای مطالعه بیشتر در مورد گراف به منبع [۳] و برای متغیرهای مختلط به منبع [۲] مراجعه شود.

۱.۱ ژن:

هر موجود زنده برای تشکیل و پایداری حیات خود به اطلاعات زیستی نیاز دارد. هرگونه از جانداران، دارای اطلاعات ژنتیکی یکتایی هستند که ژنوم نامیده می شود.

تعریف ۱.۱.۱: ژن واحد وراثتی است که صفتی خاص یا نشان ویژه را پدید می آورد.

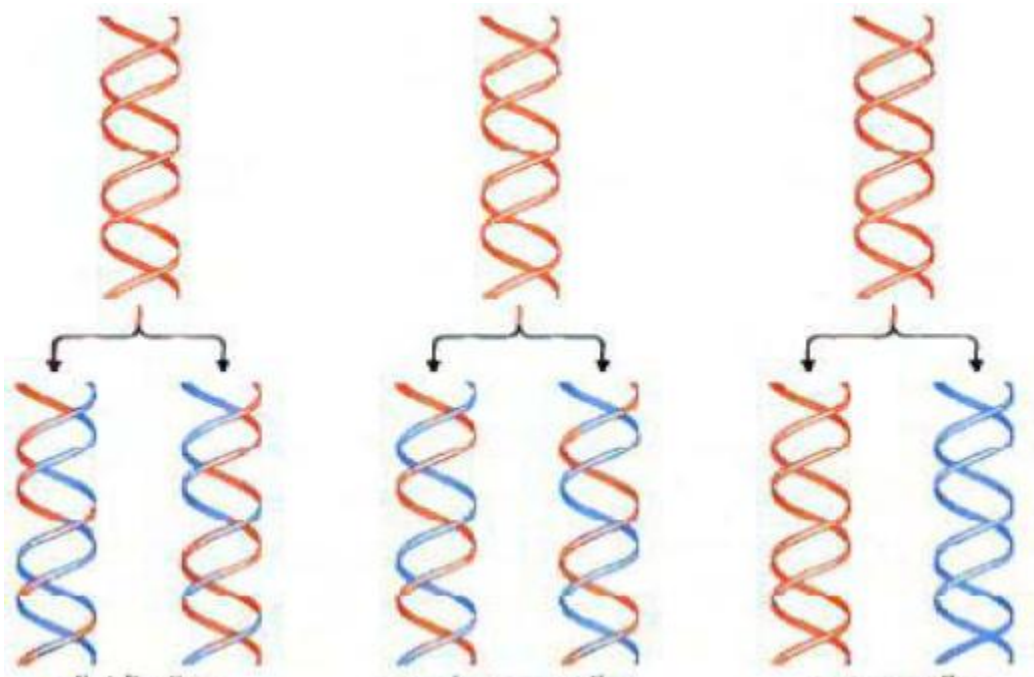
به بیان دیگر، ژن واحد وراثتی است که اطلاعات ژنتیکی را از نسلی به نسل بعد منتقل می کند.

ژن بخشی از مولکول های RNA و DNA محسوب می شود که دارای کارکرد ویژه ای است.

۲.۱ اصل بنیادی در زیست شناسی مولکولی:

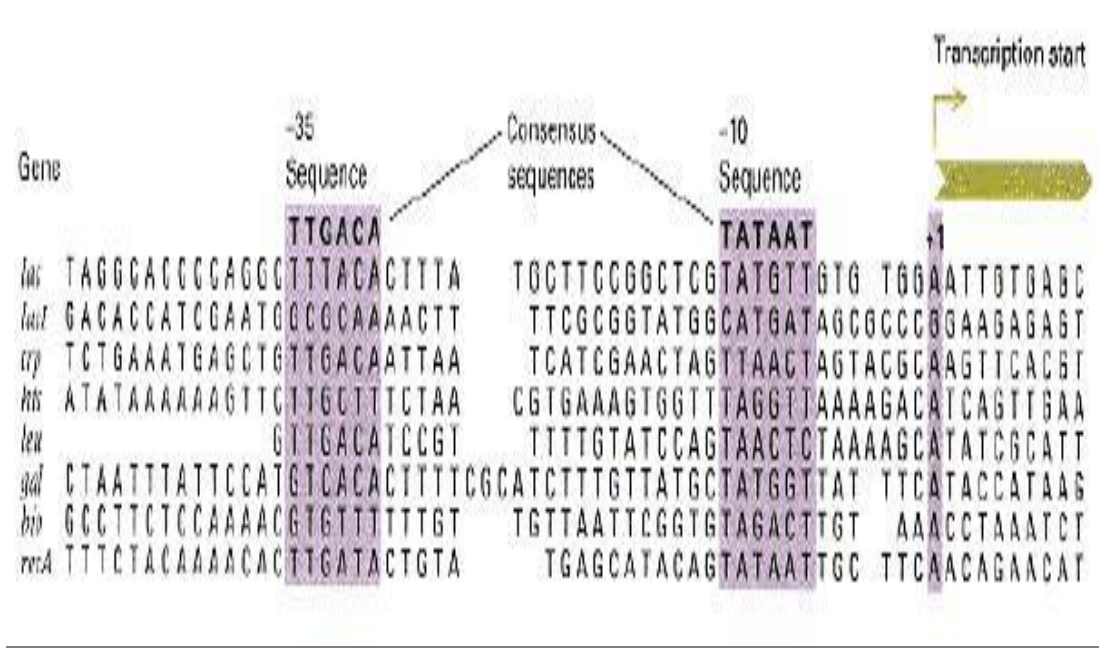
اصل بنیادی در زیست شناسی نشانگر مسیرهای عمومی انتقال اطلاعات وراثتی از راه همانند سازی DNA، رونویسی RNA از DNA و ترجمه به پروتئین می باشد.

در فرایند همانند سازی از DNA والد، مولکول DNA ی دختری ساخته می شود که بطور کامل همانند DNA ی والد است. در این مرحله پروتئین ها و آنزیم های متعددی نقش دارند که مهم ترین آنها، آنزیم DNA پلیمراز هستند.



شکل ۱.۱: اصل بنیادی در زیست شناسی

در فرآیند رونویسی یا نسخه برداری از روی ژن های ساختاری، مولکول های RNA ساخته می شود. در رونویسی برخلاف مرحله همانند سازی تنها قسمت هایی از ژنوم به RNA نسخه برداری می شوند. رونویسی را آنزیم های RNA پلیمراز وابسته به DNA با کمک چند پروتئین کمکی دیگر انجام می دهند. در اصل بنیادی به جریان دوسویه اطلاعات میان DNA و RNA توجه شده است یعنی طی فرآیندی به نام رونویسی معکوس، RNA به DNA رونویسی می شود.



شکل ۲.۱: فرآیند رونویسی

۳.۱ اسید نوکلئیک:

ماکرو ملکول هایی هستند که از چندین زیر واحد توسط اطلاعات مختلف تشکیل شده اند و دارای پراکندگی های بسیار زیادی می باشند. بطوریکه که تمام خواص موجود در یک جاندار را از نسلی به نسل دیگر انتقال می دهند.

همه جانداران اعم از باکتری ها، گیاهان و حیوانات دارای دو نوع اسید نوکلئیک خاص خود (DNA, RNA) هستند. در بین این جانداران فقط ویروس ها هستند که صرفاً RNA یا DNA دارند.

در سال ۱۸۹۶، فردریک میشر^۱ در هسته سلول آمیزه ای از ترکیبات را کشف کرد که نوکلئوین نامید. ترکیب اصلی نوکلئوین DNA است. در پایان قرن نوزدهم ساختار عمومی DNA و RNA شناسایی شد. اسید نوکلئیک از واحدهایی به نام نوکلئوتید تشکیل شده اند.

نوکلئوتیدهای تشکیل دهنده اسید نوکلئیک ها آدنین (A)، گوانین (G)، یوراسیل (U)، سیتوزین (C) و تیمین (T) می باشند.

نوکلئوتیدها حلقه های مسطحی دارند که اندازه آنها بین ۳ تا ۴ آنگستروم (هر آنگستروم 10^{-10} متر است) می باشد. وقتی مارپیچ دوتایی تشکیل می شود، مولکولهای A با مولکول های T رشته مقابل و مولکول G با مولکول C رشته روبه رو پیوند هیدروژنی برقرار می کند.

قبل از تعریف RNA لازم است مفاهیم زیر را در نظر بگیریم.

۴.۱ DNA (دوکسی ریبو نوکلئیک):

بصورت ترکیب دو زنجیره ای از نوکلئوتیدهای سیتوزین (C)، تیمین (T)، آدنین (A) و گوانین (G) تشکیل شده که این نوکلئوتیدها توسط قانون جفت گیری واتسن – کریک^۱ با هم بصورت G-C و A-T پیوند برقرار می کنند.

این زنجیره ها بصورت مارپیچ در هم فرو رفته اند و توسط پیوندهای هیدروژنی این زنجیره ها باهم اتصال دارند.

DNA در داخل هسته تحت تاثیر عوامل مختلفی فعال یا غیر فعال می شود. اگر DNA به هر دلیلی فعال شود می تواند از روی خودش، همانند خودش را بسازد (همانند سازی) و می تواند توسط پیام های خاص خود RNA را هم بسازد.

^۱Watson-Crick