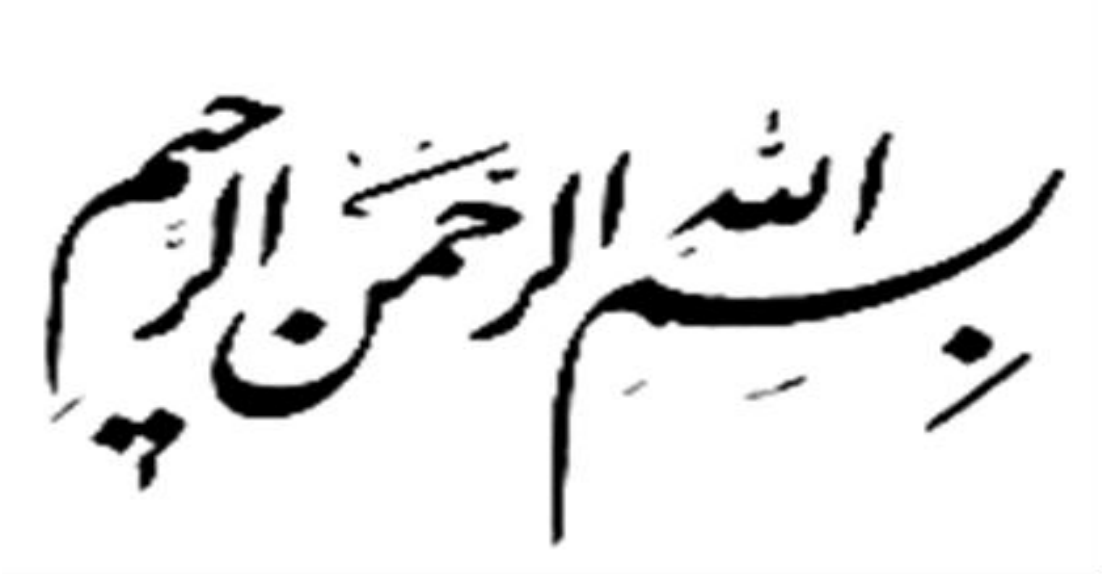

بسم الله الرحمن الرحيم

بسمه تعالی

دانشکده علوم ریاضی

تأییدیه اعضای هیأت داوران حاضر در جلسه دفاع از پایان‌نامه کارشناسی‌ارشد

اعضای هیأت داوران نسخهٔ نهایی پایان‌نامه خانم **فریبا فیاضی** رشته **ریاضی‌محض** به شماره دانشجویی **۸۹۵۲۰۵۱۲۱۹** تحت عنوان: «**تجزیه و تحلیل دنباله DNA بر اساس نظریه گراف**» را در تاریخ ۱۳۹۱/۱۰/۲۴ از نظر فرم و محتوا بررسی نموده و آن را برای اخذ درجهٔ کارشناسی‌ارشد مورد تأیید قرار دادند.

اعضای هیأت داوران	نام و نام‌خانوادگی	رتبه علمی	امضاء
۱- استاد راهنما	دکتر علی ایرانمنش	استاد	
۲- استاد ناظر داخلی	دکتر یوسف جمالی	استادیار	
۳- استاد ناظر داخلی	دکتر خسرو تاجبخش	استادیار	
۴- استاد ناظر خارجی	دکتر مجید تقدیر	استادیار	
۵- نماینده تحصیلات تکمیلی	دکتر خسرو تاجبخش	استادیار	

آیین نامه چاپ پایان‌نامه (رساله)های دانشجویان دانشگاه تربیت مدرس

نظر به اینکه چاپ و انتشار پایان نامه (رساله)های تحصیلی دانشجویان دانشگاه تربیت مدرس، مبین بخشی از فعالیتهای علمی - پژوهشی دانشگاه است بنابراین به منظور آگاهی و رعایت حقوق دانشگاه،دانش آموختگان این دانشگاه نسبت به رعایت موارد ذیل متعهد می‌شوند:

ماده ۱: در صورت اقدام به چاپ پایان نامه (رساله)ی خود، مراتب را قبلاً به طور کتبی به «دفتر نشر آثارعلمی» دانشگاه اطلاع دهد.

ماده ۲: در صفحه سوم کتاب (پس از برگ شناسنامه) عبارت ذیل را چاپ کند:

«کتاب حاضر، حاصل پایان نامه کارشناسی ارشد/ رساله دکتری نگارنده در رشته ریاضی محض
است که در سال ۱۳۹۱ در دانشکده علوم ریاضی
دانشگاه تربیت مدرس به راهنمایی سرکار خانم/جناب آقای دکتر علی ایرانمنش
از آن دفاع شده است.»

ماده ۳: به منظور جبران بخشی از هزینه‌های انتشارات دانشگاه، تعداد یک درصد شمارگان کتاب (در هر نوبت‌چاپ) را به «دفتر نشر آثارعلمی» دانشگاه اهدا کند. دانشگاه می‌تواند مازاد نیاز خود را به نفع مرکز نشر درمعرض فروش قرار دهد.

ماده ۴: در صورت عدم رعایت ماده ۳، ۵۰٪ بهای شمارگان چاپ شده را به عنوان خسارت به دانشگاه تربیت‌مدرس، تأدیه کند.

ماده ۵: دانشجو تعهد و قبول می کند در صورت خودداری از پرداخت بهای خسارت، دانشگاه می‌تواند خسارت مذکور را از طریق مراجع قضایی مطالبه و وصول کند؛ به علاوه به دانشگاه حق می‌دهد به منظور استیفای حقوق خود، از طریق دادگاه، معادل وجه مذکور در ماده ۴ را از محل توقیف کتابهای عرضه شده نگارنده برای فروش، تامین نماید.

ماده ۶: اینجانب فریبا فیاضی دانشجوی رشته ریاضی محض مقطع کارشناسی ارشد

تعهد فوق وضمانت اجرایی آن را قبول کرده، به آن ملتزم می شوم.

نام و نام خانوادگی: فریبا فیاضی

تاریخ و امضا: ۱۳۹۱/۱۱/۱۱

آیین‌نامه حق مالکیت مادی و معنوی در مورد نتایج پژوهشهای علمی دانشگاه تربیت مدرس

مقدمه: با عنایت به سیاست‌های پژوهشی و فناوری دانشگاه در راستای تحقق عدالت و کرامت انسانها که لازمه شکوفایی علمی و فنی است و رعایت حقوق مادی و معنوی دانشگاه و پژوهشگران، لازم است اعضای هیأت علمی، دانشجویان، دانش‌آموختگان و دیگر همکاران طرح، در مورد نتایج پژوهشهای علمی که تحت عناوین پایان‌نامه، رساله و طرحهای تحقیقاتی با هماهنگی دانشگاه انجام شده است، موارد زیر را رعایت نمایند:

ماده ۱- حق نشر و تکثیر پایان نامه/ رساله و درآمدهای حاصل از آنها متعلق به دانشگاه می باشد ولی حقوق معنوی پدید آورندگان محفوظ خواهد بود.

ماده ۲- انتشار مقاله یا مقالات مستخرج از پایان‌نامه/ رساله به صورت چاپ در نشریات علمی و یا ارائه در مجامع علمی باید به نام دانشگاه بوده و با تایید استاد راهنمای اصلی، یکی از اساتید راهنما، مشاور و یا دانشجو مسئول مکاتبات مقاله باشد. ولی مسئولیت علمی مقاله مستخرج از پایان نامه و رساله به عهده اساتید راهنما و دانشجو می باشد.

تبصره: در مقالاتی که پس از دانش‌آموختگی بصورت ترکیبی از اطلاعات جدید و نتایج حاصل از پایان‌نامه/ رساله نیز منتشر می‌شود نیز باید نام دانشگاه درج شود.

ماده ۳- انتشار کتاب، نرم افزار و یا آثار ویژه (اثری هنری مانند فیلم، عکس، نقاشی و نمایشنامه) حاصل از نتایج پایان‌نامه/ رساله و تمامی طرحهای تحقیقاتی کلیه واحدهای دانشگاه اعم از دانشکده ها، مراکز تحقیقاتی، پژوهشکده ها، پارک علم و فناوری و دیگر واحدها باید با مجوز کتبی صادره از معاونت پژوهشی دانشگاه و براساس آئین نامه های مصوب انجام شود.

ماده ۴- ثبت اختراع و تدوین دانش فنی و یا ارائه یافته ها در جشنواره‌های ملی، منطقه‌ای و بین‌المللی که حاصل نتایج مستخرج از پایان‌نامه/ رساله و تمامی طرح‌های تحقیقاتی دانشگاه باید با هماهنگی استاد راهنما یا مجری طرح از طریق معاونت پژوهشی دانشگاه انجام گیرد.

ماده ۵- این آیین‌نامه در ۵ ماده و یک تبصره در تاریخ ۸۷/۴/۱ در شورای پژوهشی و در تاریخ ۸۷/۴/۲۳ در هیأت رئیسه دانشگاه به تایید رسید و در جلسه مورخ ۸۷/۷/۱۵ شورای دانشگاه به تصویب رسیده و از تاریخ تصویب در شورای دانشگاه لازم‌الاجرا است.

«اینجانب فریبا فیاضی دانشجوی رشته ریاضی محض ورودی سال تحصیلی. ۹۱ مقطع کارشناسی ارشد دانشکده علوم ریاضی متعهد می شوم کِلیه نکات مندرج در آئین نامه حق مالکیت مادی و معنوی در مورد نتایج پژوهش های علمی دانشگاه تربیت مدرس را در انتشار یافته های علمی مستخرج از پایان نامه / رساله تحصیلی خود رعایت نمایم. در صورت تخلف از مفاد آئین نامه فوق الاشعار به دانشگاه وکالت و نمایندگی می دهم که از طرف اینجانب نسبت به لغو امتیاز اختراع بنام بنده و یا هر گونه امتیاز دیگر و تغییر آن به نام دانشگاه اقدام نماید. ضمناً نسبت به جبران فوری ضرر و زیان حاصله بر اساس برآورد دانشگاه اقدام خواهم نمود و بدینوسیله حق هر گونه اعتراض را از خود سلب نمودم»

امضا:.....

تاریخ : ۱۳۹۱/۱۱/۱۱

دانشکده : علوم ریاضی

پایان نامه کارشناسی ارشد

عنوان پایان نامه:

تجزیه و تحلیل دنباله DNA براساس نظریه گراف

نام دانشجو:

فریبا فیاضی

استاد راهنما:

دکتر علی ایرانمنش

دی ماه ۱۳۹۱

تقدیم به پدر و مادرم که به من چگونه زیستن را آموختند،

تقدیم به آنان که دعای خیرشان بدرقه ی راهم بود،

تقدیم به خواهر و برادر عزیزم که دوستشان دارم،

تقدیم به نیکی عزیز، امید بخش جانم.

تقدیر و تشکر

سپاس و حمد ایزد منان را که ما را از فیض بزرگ آموختن و علم بهره مند ساخت که می فرمایند:

" نون والقلم و ما یسطرون" یعنی قسم به قدرت قلم و آنچه می نگارد. زیرا با آموختن هر علمی در راستای خود پیوسته ناشناخته ها، شناخته و مجهولات کم کم جای خود را به معلومات می دهند بالاخص در علم ریاضی که خدمات این علوم در جهان چه شگفتیها آفرید و مسائل پیچیده را چه ساده و آسان گردانید.

توفیق من در رسیدن به این مرحله و کسب این درجه تمامی را مرهون و مدیون اساتید ارجمند بالاخص استاد عزیزم جناب آقای دکتر علی ایرانمنش می باشم که در اینجا قادر به توصیف زحمات بی دریغ و شایسته ایشان نیستم. براستی آب دریا را اگر نتوان کشید هم به قدر تشنگی باید چشید. باشد تا دانشجویان دیگر ی نیز از قطرات دریای عظیم علمی ایشان برخوردار گردند.

فریبا فیاضی

دی ماه

چکیده:

تعیین شباهت دنباله ها یکی از مراحل مهم در مطالعات فیلوژنتیک (شجره های وراثتی) محاسباتی است. همانطور که می دانیم، در طول تاریخ تکامل، نه تنها جهش DNA، بلکه بازآرایی نوکلئوتیدها در دنباله DNA برای هر فرد رخ می دهد. همین امر، زیست شناسان را درگیر محاسباتی برای توصیف ریاضی و تجزیه و تحلیل تشابه دنباله ها می کند. در این پایان نامه دو روش گرافی برای تجزیه و تحلیل دنباله های DNA معرفی می کنیم.

یکی از روش ها توسط زینکگین در سال ۲۰۱۱ معرفی شده است. وی برای هر دنباله DNA، یک گراف جهت دار و وزن دار معرفی نمود. همچنین با استفاده از گراف جهت دار ماتریس مجاورت و بردار نماینده تعریف نمود. این روش روی مجموعه ای از ۰,۹ kb-mt دنباله DNA از دوازده گونه مختلف پستانداران تست شده است.

روش دوم توسط ناتارجان در سال ۲۰۱۰ معرفی گردید. وی هر دنباله DNA را به گراف خطی تبدیل نمود و شاخص اتصال را برای توصیف عددی معرفی کرده است و این روش نیز بر روی ۲۳ ژن متفاوت آزمایش شده است.

در ادامه، به معرفی دو روش جدید برای توصیف عددی دنباله های DNA می پردازیم.

در روش اول از دنباله مشخصه تعریف شده توسط هی و ونگ استفاده کرده و ضریب همبستگی را روی آن تعریف می کنیم. این روش را روی ۸ گونه مختلف آزمایش می کنیم.

در روش دوم با استفاده از نمایش گرافی دو بعدی معرفی شده توسط لئو، توزیع نرمال دو متغیره را روی دنباله های DNA تعریف می کنیم. همچنین این روش را نیز روی ۱۱ گونه متفاوت آزمایش میکنیم.

واژه های کلیدی: دنباله DNA، گراف وزن دار، بردار نماینده، اندیس اتصال، ضریب همبستگی و توزیع نرمال دو جمله ای.

فهرست مطالب

عنوان **صفحه**

فهرست شکل ها.....ت

فهرست جدول ها.....ث

مقدمه.....۱

فصل ۱- مفاهیم و تعاریف اولیه.....۴

۱-۱- اسید نوکلوئیک.....۵

۱-۱-۱- ساختمان رشته ای DNA ۷

۱-۲- مفاهیم اولیه نظریه گراف.....۱۰

۱-۲-۱ نمایش هندسی گراف ها.....۱۱

۱-۳- تجزیه و تحلیل مولکول DNA.....۱۳

فصل ۲- تجزیه و تحلیل دنباله DNA براساس نظریه گراف.....۱۸

۲-۱- ساخت بردار نماینده برای دنباله DNA.....۱۹

۲-۲- ساخت گراف وزن دار.....۲۱

۲-۳- ساخت ماتریس مجاورت.....۲۲

۲-۴- معرفی سه فاصله اندازه گیری.....۲۳

۲-۵- برنامه کاربردی و تجربی داده ها و توصیف آنها.....۲۴

۲-۶- تجربیات گذشته برای گونه های ذکر شده بر پایه روش های متفاوت.....۲۶

۲-۷- انتخاب پارامتر α ۲۸
۲-۸- ماتریس تشابه بر پایه روش DET ۲۹
۲-۹- ساخت درخت دندروگرام ۳۱
۲-۱۰- تست شبیه سازی برای روش DET ۳۳
۲-۱۰-۱- ساخت مجموعه داده های شبیه سازی شده ۳۴
۲-۱۱- تبدیل دنباله DNA به گراف خطی ۳۶
۲-۱۲- شاخص های اتصال برای گراف خطیDNA ۳۸
۲-۱۳- کاربرد شاخص اتصال $h_{\chi^{pKa}}$ برای مقایسه دنباله DNA ۳۹
فصل ۳- توصیف عددی دنباله DNA براساس ضریب همبستگی ۴۴
۳-۱- ساخت دنباله مشخصه ۴۵
۳-۲- ساخت بردار نماینده ۴۶
۳-۳- ضریب همبستگی ۴۷
۳-۴- کاربرد و نتیجه گیری ۴۸
فصل ۴- توصیف عددی دنباله DNA براساس توزیع نرمال دو جمله ای ۵۱
۴-۱- نمایش گرافی ۲ بعدی از دنباله DNA ۵۲
۴-۲- توزیع نرمال ۵۶
۴-۳- توصیف عددی دنباله های DNA براساس توزیع نرمال دو جمله ای ۵۷
نتیجه گیری و پیشنهادات ۶۰
منابع ۶۱

فهرست شکل ها

عنوان **صفحه**

شکل ۱-۱- a) آدنین b) گوانین C) سیتوزین d) تیمین..... ۷

شکل ۱-۲- مارپیچ DNA..... ۸

شکل ۱-۳- پیوند بین بازهای مکمل..... ۸

شکل ۱-۴- گراف جهت دار و گراف بدون جهت..... ۱۲

شکل ۱-۵- دو نمایش از گراف پترسن..... ۱۲

شکل ۱-۶- دو نمایش از گراف کامل..... ۱۲

شکل ۲-۱- گراف وزن دار برای دنباله مورد نظر..... ۲۰

شکل ۲-۲- گراف ادغام شده شکل ۲-۱..... ۲۱

شکل ۲-۳- درخت تکامل نژادی ۱۲ گونه ۱ براساس روش های مختلف..... ۲۷

شکل ۲-۴- درخت فیزیولوژیکی ۱۲ گونه..... ۳۲

شکل ۲-۵- آ) مقادیر pk_a برای پایه ها ب) تبدیل دنباله DNA به گراف خطی وزن دار..... ۳۷

شکل ۲-۶- تجزیه و تحلیل فیلوژنتیکی از بیست و سه توالی های ژنومی میتوکندری DNA..... ۴۳

شکل ۴-۱- گراف پیرامیدین- پورین..... ۵۳

فهرست جداول

عنوان **صفحه**

جدول ۲-۱- اطلاعات ۱۲ گونه مختلف..... ۲۵

جدول ۲-۲ مقادیر مختلف L_0 را به ازای α های مختلف..... ۲۹

جدول ۲-۳- ماتریس بالا مثلثی براساس d_1..... ۲۹

جدول ۲-۴- ماتریس بالا مثلثی براساس d_2..... ۳۰

جدول ۲-۵- جدول بالا مثلثی بر اساس d_3..... ۳۱

جدول ۲-۶- مقادیر *suc-rate* برای روش *DET*..... ۳۵

جدول ۲-۷- اطلاعات DNA ۲۳ گونه مختلف..... ۴۰

جدول ۲-۸- اندیس اتصال از مرتبه ۰ تا ۱۰ برای گرافهای خطی گونه های مورد مطالعه..... ۴۱

جدول ۳-۱- اولین اکسون ژن بتا گلوبین برای هشت گونه متفاوت..... ۴۹

جدول ۳-۲- ماتریس تشابه هشت گونه مورد نظر براساس ضریب همبستگی..... ۴۹

جدول ۴-۱- دنباله *DNA* اولین اکسون ژن بتا گلوبین یازده گونه متفاوت..... ۵۴

جدول ۴-۲- مختصات $(\dot{x},\dot{y})$ با n متفاوت برای گونه های ذکرشده در (جدول ۴-۱)..... ۵۴

جدول ۴-۳- ماتریس تشابه گونه های ذکر شده براساس توزیع نرمال دو متغیره برای
$\rho = 0, s = 30$,t= ۳۱..... ۵۶

جدول۴-۴- ماتریس تشابه گونه های ذکر شده براساس توزیع نرمال دو متغیره برای
$\rho = 0, s = 50, t = 31$..... ۵۶

جدول۴-۵- ماتریس تشابه گونه های ذکر شده براساس توزیع نرمال دو متغیره برای
$\rho = 0, s = 30, t = 50$..... ۵۷

جدول۴-۶ – ماتریس تشابه گونه های ذکر شده براساس توزیع نرمال دو متغیره برای $\rho = \frac{1}{2}, s = 50, t = 31$ ۵۷

جدول ۴-۷- ماتریس تشابه گونه های ذکر شده براساس توزیع نرمال دو متغیره برای $\rho = -\frac{1}{2}, s = 50, t = 31$ ۵۷

مقدمه

تعیین توالی دقیق رشته های DNA نقش اساسی در بیولوژی، پزشکی و کشاورزی نوین دارد. تعیین توالی سنتی رشته DNA بطور مستقیم با تعیین نوکلئوتید به نوکلئوتید در طول رشته DNA صورت می گیرد. در این روش که به تعیین توالی مستقیم مرسوم است، تشخیص درست هر نوکلئوتید به تشخیص صحیح چندین نوکلئوتید در مراحل قبل وابسته است. هرچند روش مستقیم تعیین توالی نوکلئوتیدها کاربردی است، ولی روند تعیین توالی بصورتی است که بوجود آمدن یک خطا، خطاهای دیگر را در پی خواهد داشت. دقت بسیار بالا در انجام عملیات، در کم شدن مقدار خطاها موثر است ولی با این حال فقط تعیین توالی رشته هایی کوتاه در این روش امکان پذیر است.

پژوهشگران بسیاری از تخصص های مختلف دست به دست هم دادند تا بستری مناسب برای تجزیه و تحلیل دنباله های DNA مهیا شود.

یکی از این روش های موثر، تجزیه و تحلیل دنباله ها براساس نظریه گراف می باشد. در روش های گرافی معمولا از روش اقلیدسی استفاده می شود. در سال ۱۹۸۳، هاموری و راسکین[1] اولین نمایش گرافیکی را ارائه دادند[۱]. ناندی و جئو[2] نیز روش گرافیکی در فضای ۲ بعدی را ارائه نمودند[۲]. در سال های بعد، روش گرافیکی به ابعاد بالاتر نیز تعمیم یافت و از نظر بعد به پنج دسته طبقه بندی گردید، اعم از نمایش گرافی در فضای ۲بعدی تا ۶بعدی[۳-۷]. اخیرا روش های گرافی کاربرد فراوانی دارند زیرا علاوه بر ارائه خصوصیات بصری، اطلاعات عددی نیز می توان از آنها استخراج نمود که در تجزیه و تحلیل دنباله ها کمک شایانی می نمایند.

1 Humori and Ruskin

2 Nundy and Guo

در این مطالعه به تجزیه و تحلیل دنباله های DNA براساس نظریه گراف می پردازیم. در فصل اول تعاریف و مفاهیمی برای درک بهتر فصل های بعدی آورده شده است. در فصل دوم این مطالعه دو روش گرافی ارئه می گردد که بصورت زیر می باشند:

در مرجع [۸]، روشی جدید معرفی شده است که بر پایه نظریه گراف است. برای هر دنباله DNA، یک گراف جهتدار و وزن دار، بردار نماینده و ماتریس مجاورت بدست آورده است. سپس سه فاصله اندازه گیری برای بردارهای نماینده معرفی شده است که دسترسی به تجزیه و تحلیل دنباله DNA را آسان تر می کند. به عنوان کاربرد روش، آن را روی مجموعه ای ازmt 9 kb/. (۰,۹کیلو باز میتوکندری) دنباله DNA از دوازده گونه مختلف پستانداران آزمایش شده است و نتیجه آن بر پایه سه فاصله اندازه گیری بدست آمده می باشد که نتایج یکسانی دارند و با نتایج قبلی نیز مطابقت دارد. همچنین برای مشاهده قدرت و تحمل زیر دنباله های تراز شده، روی یکی از داده های مصنوعی آزمایش شده است.

همچنین در مرجع [۹]،هر دنباله DNA را به یک گراف خطی تبدیل و اندیس اتصال را بر روی آن تعریف کرده است و به کمک آن به تجزیه و تحلیل دنباله می پردازد. در انتها روش بر روی ۲۳ ژن مختلف آزمایش می شود که نشان دهنده درستی و موثر بودن روش است.

در فصل سوم این مطالعه، روش جدیدی براساس ضریب همبستگی ارائه می دهیم. هی و ونگ[1] روشی برای منظم سازی پایه های دنباله DNA براساس دسته بندی نوکلئوتیدها ارائه کردند و آن را دنباله مشخصه نامیدند[۱۰]. سپس با استفاده از دنباله های مشخصه ماتریس های ۶×۴ را معرفی نمودند که نشان دهنده دنباله های DNA می باشد. در اینجا، با استفاده ار این ماتریس بردار نماینده را تعریف می نماییم و سپس ضریب همبستگی بین دو دنباله بردار نماینده متناظر با دنباله DNA تعریف می شود. در انتها برای نشان دادن موثر بودن روش، آن را بر روی ۸ گونه متفاوت آزمایش می کنیم و خواهیم دید که نتایج بدست آمده با نتایج قبلی همخوانی دارد.

1 He and wang

در فصل چهارم این پایان نامه نیز روش جدید دیگری بر اساس توزیع نرمال دو متغیره ارائه داده ایم. در مرجع [۱۱]، لئو[1] نمایش گرافیکی در دستگاه مختصات ارائه نمود که در آن پیرامیدین (C,T) در ناحیه اول و پورین (A,G) در ناحیه چهارم قرار می گیرند. در اینجا با بکارگیری این نمایش گرافیکی توزیع نرمال را معرفی خواهیم کرد. سپس توزیع نرمال دو جمله ای بین دو دنباله DNA متفاوت تعریف می شود که به کمک آن می توان به مقایسه دنباله های DNA پرداخت. در انتها این روش را بر روی اولین اکسون از ژن بتا گلوبین ۱۱ گونه متفاوت آزمایش می کنیم.

مطالب فصل سوم و چهارم، نتایج جدیدی است که توسط نگارنده صورت گرفته است و حاصل آن مقالات [۱۲] و [۱۳] می باشند.

فصل اول

تعاریف و مفاهیم اولیه

۱-۱- اسید نوکلوئیک[1]

سال ها پیش از آنکه دانشمندان سعی کنند تا با استفاده از قوانین فیزیکی و شیمیایی پدیده های زیست شناختی را تبیین کنند، زیست شناسان با مشاهده گیاهان و جانوران، قلمرو دانش خود را گسترش می دادند. در سال ۱۸۶۹، فردریش فیشر[۲] ، در بیمارستان ماده ای را از محل عفونت که غنی از گلبول های سفید بود، استخراج کرد. میشر این ماده را نوکلوئین نامید. البته بعدها پس از شناخت ساختار شیمیایی این ماده به آن نام اسید نوکلوئیک نسبت داد. او در هنگام مطالعه روی گلبولهای سفید، هسته این سلولها را استخراج کرد و سپس روی آن محلول قلیایی ریخت. حاصل این آزمایش رسوب لزجی بود که بررسیهای شیمیایی آن نشان داد که ترکیبی از کربن، هیدروژن، اکسیژن و نیتروژن و درصد بالایی فسفر است. فیشر این ماده را نوکلوئین نامید. زمانی که ماهیت اسیدی این ماده مشخص گردید نام آن به اسید نوکلوئیک تغییر یافت.

اسید نوکلوئیک یکی از ماکرومولکول های زیستی است که وظیفه ذخیره اطلاعات ژنتیکی را در سلول بر عهده دارد. جایگاه اسیدهای نوکلوئیک در هسته و سیتوپلاسم سلول است و از واحدهایی به نام نوکلئوتید ساخته شده‌اند. اسیدهای نوکلوئیک بسپارهایی (پلیمرهایی) با زنجیر طولانی و وزن مولکولی بالا متشکل ازنوکلوئیک ها هستند. هرنوکلئوتید از قسمت های زیر تشکیل شده است.

1Nucleic Acid

۲Friedrich Miesher

- یک مولکول اسید فسفریک

- یک مولکول قند ۵ کربنی

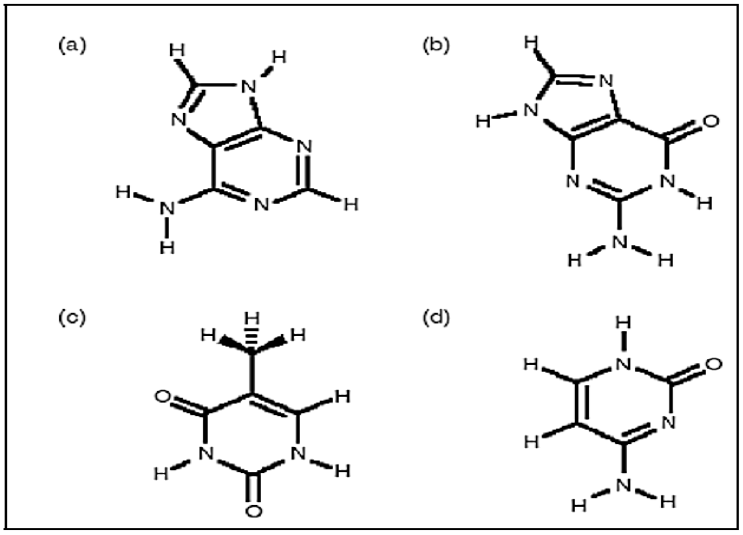
-یک مولکول باز نیتروژن دار

دو نوع اسید نوکلئیک وجود دارد. دزوکسی ریبونوکلئیک اسید(DNA) و ریبونوکلئیک اسید(RNA). اختلاف اساسی بین این دو مولکول قندی است که مورد استفاده قرار داده‌اند. DNA حاوی دزوکسی ریبوز و RNA حاوی ریبوز است.

با وجودی که شناخت اسید های نوکلوئیک به یک قرن پیش برمی گردد ولی اطلاعات اساسی در مورد ساختمان آنها فقط طی چند دهه اخیر کشف شده است[۱۴].

۱-۱-۱- ساختمان رشته ای DNA

در هسته هر سلول مولکولهایی قرار دارند که اساسی ترین اطلاعات حیات را در خود ذخیره کرده اند. این مولکول ها، دئوکسی ریبو نوکلوئیک اسید (DNA) نام دارند. DNA از چهار نوع مولکول ساخته شده است، که به آنها " نوکلئوتید " می گوییم. این چهار نوکلئوتید عبارتند از: آدنین (A)، گوانین (G)، سیتوزین (C) و تیمین (T) (شکل ۱-۱). این مولکولها خود به دو زیر گروه تقسیم می شوند: آدنین و گوانین در گروه پورین[۱] قرار می گیرند و سیتوزین و تیمین در گروه پیرامیدین[۲] جای داده می شوند.



شکل ۱-۱: a) آدنین b) گوانین C) سیتوزین d) تیمین

هنگام سنتز (ساخته شدن) مولکول DNA، نوکلئوتیدها به اسیدهای نوکلوئیک تبدیل می شوند که بعدا به هم متصل می شوند و رشته های DNA را می سازند .در نهایت یک مارپیچ دوتایی ساخته می شود. به دلیل اینکه در رشته های DNA باز A همواره در مقابل T و باز C در مقابل G قرار دارد این دو رشته را مکمل می نامند.

1 purine

2 Pyrimidine