دانشکده مهندسی برق و کامپیوتر

گروه مهندسی کامپیوتر

## عنوان

# ارائه یک روش هوشمند برای تحلیل و پیش‌بینی ژن‌های رمزکننده اسید فسفاتازها و کاربرد در یک گیاه خاص

پایان‌نامه کارشناسی ارشد

استاد راهنما

دکتر علی احمدی

استاد مشاور

دکتر محمد علی ملبوبی

نگارش

صدف ایرانپور طاری

شهریور ۱۳۹۲

# تقدیم به پدر و مادرم

دو بی‌کران بی‌همتا، دو زلال اندیش و دو سرو قامتی که گوهر وجودشان، نسیم کلامشان و باران محبتشان را همواره بی‌هیچ منت و ادعا مرهمی نموده اند بر خستگی‌هایم. آنان که مهر و قهرشان گرمی عشق دارد. آنان که راستی قامتم در شکستگی قامتشان تجلی یافت و قفنوس جوانیشان به پای روشنایی حیات من سوخت. در برابر وجود گرامیشان زانوی ادب بر زمین می نهم و با دلی مملو از عشق و محبت بر دستان پر مهرشان بوسه می زنم.

**تشکر و قدردانی**

ازاستاد ارجمند جناب آقای دکتر احمدی، به پاس حمایت‌ها و راهنمایی‌های ارزنده ایشان که در انتخـاب مسیر زندگی همواره یاری‌گر بنده بوده و نیز اساسی‌ترین نقش را در انجام این تحقیق داشته‌اند.

## چکیده

یکی از اهداف اصلی علم نوظهور بیوانفورماتیک، درک فرایند تنظیم ژن در موجودات مختلف است که در چند سال اخیر مساله مهندسی معکوس در مدل‌سازی و شناسایی شبکه‌های تنظیم‌کننده ژنی با استفاده از داده‌های بیان ژن محور بسیاری از پژوهش‌ها قرار گرفته است. از آن‌جا که فرآیند تنظیم ژن‌ها در سطح سلولی یک روند دینامیک و پویا است، بر این اساس شبکه‌های عصبی به‌دلیل قابلیت آموزش و تنظیم بر اساس داده‌های آموزشی و سیستم‌های فازی به دلیل قابلیت تفسیرپذیری، الگوریتم‌های مناسبی جهت این نوع محاسبات شناخته شده‌اند.

در پایان‌نامه پیش‌روی، روش‌هایی برای پیش‌بینی ارتباطات پیچیده میان ژن‌ها انجام شده است که مبنای شبکه‌های عصبی-فازی دارد. هم‌چنین روشی جهت استخراج روابط از روی داده‌های ریزآرایه پیشنهاد شده که ساختاری نوین پویا و همراه با بهینه‌سازی دارد. روش پیشنهادی، ژن‌هایی که بیشترین تاثیر بر هم دارند را به عنوان ژن‌های تنظیم‌کننده می‌یابد و نوع روابط آن‌ها شامل اثر سدکنندگی، فعال‌کنندگی و یا خنثی، مشخص می‌کند و در نهایت شبکه تنظیم ژنی بر اساس نتایج به‌دست آمده ترسیم می‌شود. اگرچه در ابتدا قصد بر آموزش شبکه با مجموعه داده‌های مربوط به اسید فسفاتازهای یک گیاه خاص بود، اما به دلیل عدم دسترسی به مجموعه داده نام‌برده، به عنوان جایگزین، از مجموعه داده‌های استاندارد ریزآرایه مربوط به ۱۲ ژن شاخص موثر در مدت سیکل جوانه زدن نوعی مخمر[1] استفاده شد. این داده‌ها شامل نمونه‌های جمع‌آوری شده در نقاط زمانی مختلف از سیکل سلولی است که جهت نمایش خروجی استخراج شده و در نتیجه تعیین بازدهی با مقایسه روابط بین ژنی به‌دست آمده تجربی، استفاده شده است. مجموعه داده‌ها به ترتیب شامل Alpha با ۱۸ نقطه زمانی و cdc15 با ۲۴ نقطه زمانی است. فعل و انفعالاتی که از میان مجموعه‌ای از ژن‌های مشخص، با نقش تنظیمی در طول چرخه سلول، حاصل شده است با نتایج پژوهشی بیولوژیکی پیشین، تحت معتبرسازی قرار گرفته‌اند. نتایج شبیه‌سازی نشان‌دهنده آن است که با اجرای روش پیشنهادی، ۱۵٪ از تعداد قوانین استخراج شده جهت بخش‌بندی فضای ورودی- خروجی کاسته شده و این منجر به کاهش شدید محاسبات می‌شود، در حالی که میزان مجموع مربعات خطای الگوریتم نیز در مقایسه با نزدیکترین روش از لحاظ الگوریتم، کاهش یافته است.



**کلید واژه:** بیوانفورماتیک، بیان ژن، شبکه تنظیم ژن، شبکه‌های عصبی- فازی، بهینه‌سازی تکاملی.


[1] Saccharomyces Cerevisiae

# فهرست مطالب


**۱ مقدمه ..... ۱**

۱-۱- مقدمه ..... ۳

**۲ تنظیم ژن وتحلیل داده‌های بیان ژن ..... ۷**

۲-۱- مقدمه ..... ۹

۲-۲- ژن ..... ۹

۲-۳- بیان ژن ..... ۱۰

۲-۴- رونویسی ..... ۱۲

۲-۵- ترجمه ..... ۱۳

۲-۱- تنظیم ژن ..... ۱۳

۲-۲- ریزآرایه‌ها ..... ۱۴

۲-۲-۱- روال عملکرد ریزآرایه‌ها ..... ۱۶

۲-۳- کاربردهای ریزآرایه‌های DNA ..... ۲۰

۲-۴- تحلیل داده‌های ریزآرایه‌های بیان ژن ..... ۲۲

۲-۵- نتیجه گیری ..... ۲۴

**۳ روش‌های مدل‌سازی شبکه‌های ژنی ..... ۲۶**

۳-۱- مقدمه ..... ۲۸

۳-۲- معادلات دیفرانسیل ..... ۲۸

۳-۳- شبکه‌های بولی ..... ۳۰

۳-۴- شبکه‌های بولی تصادفی ..... ۳۰

۳-۵- شبکه‌های بیزی ..... ۳۲

۳-۶- شبکه‌های بیزی دینامیک ..... ۳۶

۳-۷- روش‌های مبتنی بر تئوری اطلاعات ..... ۳۷

۳-۸- روش‌های خوشه‌بندی ..... ۳۸

۳-۹- مدل‌های فضای حالت ..... ۴۰

۳-۱۰- برنامه نویسی ژنتیکی ..... ۴۱

۳-۱۱- شبکه عصبی-فازی ..... ۴۳

۳-۱۱-۱- توضیح یک شبکه عصبی-فازی در بیان ژن(روش منشائی) ..... ۴۴

۳-۱۱-۱-۱- مرحله اول، پیش‌بینی با شبکه عصبی-فازی ..... ۴۴

۳-۱۱-۱-۲- مرحله دوم، آموزش وزن قوانین ..... ۴۷

۳-۱۱-۱-۳- مرحله سوم، تعیین نوع روابط بین ژن‌ها ..... ۴۸

۳-۱۱-۲- نقد روش منشائی، کاربرد شبکه عصبی-فازی در بیان ژن ..... ۵۰

۳-۱۲- نتیجه‌گیری ..... ۵۱
**۴ روش پیشنهادی ..... ۵۲**
۴-۱- مقدمه ..... ۵۴
۴-۲- خصوصیات الگوریتم پیشنهادی و ماهیت رفتار آن ..... ۵۴
۴-۳- یادگیری ساختار ..... ۵۷
۴-۴- مراحل شبکه فازی- عصبی ..... ۶۴
۴-۵- ساخت قوانین ..... ۶۷
۴-۶- انتخاب تنظیم‌کننده‌های هر ژن ..... ۶۸
۴-۷- آموزش و اصلاح ساختار ..... ۶۹
۴-۸- آموزش پارامترهای قوانین ..... ۷۲
۴-۹- محاسبه نوع تاثیر ژن/ژن‌های ورودی بر ژن خروجی ..... ۷۳
۴-۱۰- نتیجه‌گیری ..... ۷۵
**۵ پیاده‌سازی و ارزیابی ..... ۷۶**
۵-۱- مقدمه ..... ۷۸
۵-۲- مجموعه داده مخمر ..... ۷۸
۵-۳- فاز اصلاح ساختار ..... ۸۱
۵-۴- مرحله کاهش و آموزش قوانین ..... ۸۶
۵-۵- نتیجه‌گیری ..... ۹۱
**۶ نتیجه‌گیری و پیشنهادات ..... ۹۲**
۶-۱- نتیجه‌گیری ..... ۹۴
۶-۲- پیشنهادات ..... ۹۴
**۷ مراجع ..... ۹۶**

# فهرست شکل‌ها


شکل۲-۱ ساختار کلی سلول ..... ۱۰
شکل ۲-۲ رونویسی از روی DNA و ترجمه از روی RNA و تولید یک پروتئین ..... ۱۱
شکل ۲-۳ فرآیند رونویسی از روی RNA به وسیله RNA POLYMERASE ..... ۱۲
شکل ۲-۴ بازتابش نور فلورانس در زیر آرایه‌ها. ..... ۱۸
شکل ۲-۵ رویه تولید داده‌های ریزآرایه‌های بیان ژن ..... ۱۹
شکل ۲-۶ ماتریس بیان ژن ..... ۲۰
شکل ۳-۱ بلوک سازندهی یک شبکه بولی تصادفی ..... ۳۲
شکل ۳-۲ شبکه بیزی با توزیع‌های احتمال شرطی و استقلال‌های شرطی ..... ۳۴
شکل ۳-۳ شبکه بیزی نشان‌دهنده تاثیر دو متغیر بر روی متغیر دیگر. ..... ۳۵
شکل ۳-۴ الف: زنجیره مارکوف با نمایش شبکه بیزی دینامیک ب: مدل مارکوف پنهان با نمایش شبکه بیزی دینامیک .... ۳۷
شکل ۳-۵ مقایسه روشهای بیزی، تئوری احتمال و معادلات دیفرانسیل ..... ۴۰
شکل ۳-۶ شمای کلی شبکه عصبی-فازی استفاده شده در مقاله منشائی ..... ۴۵
شکل ۴-۱ الگوریتم پیشنهادی پروژه ..... ۵۶
شکل ۴-۲ نمونه‌ای از توابع عضویت با سه مجموعه فازی بیان یک ژن ..... ۵۷
شکل ۴-۳ استخراج قوانین فازی ..... ۵۸
شکل ۴-۴ انواع بخش‌بندی فضای ورودی دو بعدی ..... ۶۰
شکل ۴-۵ نحوه حرکت راه‌حل‌ها در الگوریتم BPSO ..... ۷۱
شکل ۵-۱ روابط موجود در چرخه سلولی مخمر ..... ۸۰
شکل۵-۲ بلوک قوانین فازی اولیه ..... ۸۱
شکل۵-۳ بلوک‌های مربوط به فاز اصلاح ساختار ..... ۸۲
شکل۵-۴ بلوک‌های مرحله کاهش و آموزش قوانین ..... ۸۶
شکل ۵-۲ روابط استخراج شده توسط الگوریتم از داده‌های CDC-15 ..... ۸۹

# فهرست جداول

جدول۵-۱ اسامی ژن‌های مجموعه‌داده CDC15 ..... ۷۹
جدول ۵-۲ تعداد قوانین استخراج شده قبل از اصلاح ساختار ..... ۸۱
جدول ۵-۳ قوانین استخراج شده بعد از اصلاح ساختار ..... ۸۳
جدول ۵-۴ مجموع قوانین فازی پس از اصلاح..... ۸۳
جدول ۵-۵ مجموع مربعات خطای شبکه قبل از اصلاح ..... ۸۴
جدول ۵-۶ مجموع مربعات خطای شبکه پس از اصلاح ..... ۸۵
جدول ۵-۷ مجموع مربعات خطا پس از اتمام تمامی مراحل الگوریتم بر حسب تنظیم‌کننده‌ها ..... ۸۷
جدول ۵-۸ مقایسه مجموع مربعات خطای هر ژن خروجی الگوریتم پیشنهادی و مرجع[۸۶] ..... ۸۷
جدول ۵-۹ مجموعه‌های استخراجی روابط موجود میان ژن‌ها..... ۸۸
جدول ۵-۱۰ مقایسه تعداد روابط استخراجی الگوریتم پیشنهادی و سایر روش‌ها ..... ۸۹
جدول ۵-۱۱ مقایسه روابط استخراجی الگوریتم پیشنهادی و سایر روشها ..... ۹۰

فصل اول

# مقدمه

## ۱-۱- مقدمه

در سال‌های اخیر، افزایش امکانات تحقیقاتی و آزمایشگاهی منجر به سهولت استخراج داده‌های بیولوژیک شده است. از آن‌جایی که پردازش، تحلیل و بررسی این داده‌ها بسیار زمان‌بر، پرهزینه و گاهی اوقات غیرعملی است، نیاز به استفاده از روش‌های مهندسی و محاسباتی به‌جای استفاده از روش‌های زمان‌بر آزمایشگاهی به‌صورت روز افزون احساس می‌شود.

بیوانفورماتیک دانش استفاده از علوم کامپیوتر و آمار و احتمالات در شاخه زیست شناسی مولکولی است. در چند دهه اخیر، با پیشرفت در زیست‌شناسی مولکولی وبا بهره از تجهیزات مورد نیاز تحقیق، با استفاده از کامپیوتر، نرم‌افزارهای متعدد و بانک‌های اطلاعاتی بیولوژیکی، تحقیقات وسیعی در خصوص پروتئین‌ها و ژن‌ها به‌عمل آمده است[۱]. هم‌چنین در این زمینه، تعیین توالی ژنوم بسیاری از گونه‌های موجودات پیشرفت قابل توجهی کرده است، تا جایی که پروژه‌های تعیین توالی ژنوم‌ها از پروژه‌های بسیار رایج به‌حساب می‌آیند. امروزه توالی ژنوم بسیاری از موجودات ساده مانند باکتری‌ها (از دسته پرکاریوت‌ها)، تا موجودات بسیار پیشرفته چون یوکاریوت‌ها[1] شناسایی شده است. در این راستا پروژه شناسایی ژنوم انسان در سال ۱۹۹۰ آغاز شد و در سال ۲۰۰۴ پایان یافت و اکنون اطلاعات کامل مربوط به توالی هر ۴۶ کروموزوم انسان موجود است[۲]. با استفاده از اطلاعات به‌دست آمده، تلاش‌های پژوهشی زیادی نیز در این رشته صورت گرفته است، از جمله: تنظیم توالی، کشف ژن، گردآوری ژنوم، تنظیم ساختار پروتئینی، پیش‌گویی ساختار پروتئینی، پیش‌بینی بیان ژن و تعاملات پروتئین – پروتئین.

تنظیم بیان ژن یا تنظیم ژن، به کنترل مقدار و زمان تغییرات در تولید محصول یک ژن اشاره می‌کند. محصولات ژن‌های مختلف می‌توانند به‌صورت مستقل یا به صورت ترکیبی، رونویسی از روی یک ژن را فعال یا متوقف کنند. شبکه‌های تنظیم ژنی این ارتباطات را در قالب یک شبکه مدل می‌کنند. در حال حاضر درک شبکه‌های تنظیم‌کننده ژنی و فهم فرایندهای تنظیمی در یک سلول در سطح ژن، یک هدف مهم دربیولوژی محاسباتی است.

[1] همه جانداران، پرو کاریوت یا یوکاریوت هستند. سلول‌های پروکاریوتی از سلول‌های یوکاریوتی ساده‌تر هستند و هسته‌ای که با غشایی از دیگر بخش‌های سلول جدا شده باشد، ندارند. باکتری‌ها پروکاریوت هستند. سلول‌های یوکاریوتی بزرگ‌ترند و سازمان‌بندی درونی پیچیده‌تری دارند.

مدل‌سازی شبکه‌های ژنی می تواند در زمینه های مختلف مانند اکتشافات داروهای جدید، کاهش اثرات جانبی روش‌های درمانی، شناسایی کامل‌تر بیماری‌های ژنتیکی، انتخاب کاندیداهای ژن درمانی و مقایسه و بررسی الگوهای بیان ژن‌های با عملکرد ناشناخته و به‌دست آوردن ایده‌هایی در مورد عملکرد آن‌ها کاربرد داشته باشد. با توجه به اهمیت زیاد مکانیسم تنظیم ژن در انسان و احتمال ایجاد بیماری در صورت وجود اشکال در این فرایند و نیز با در نظر گرفتن عدم توانایی روش‌های سنتی در شناسایی کامل و جامع ارتباطات تنظیمی بین ژن‌ها، نیاز به استفاده از شبیه‌سازی و مدل‌سازی کامپیوتری در این زمینه به صورت روز افزون احساس می‌شود.

گرچه پژوهش‌های بسیاری برای درک مکانیزم‌های تنظیم ژن انجام شده است، اما هنوز دامنه این پژوهش‌ها از مرز موجودات ساده نگذشته است. اگر بتوان رفتار ژن‌ها را در سلول های موجودات پیچیده‌تر به‌طور کلی شناسایی و تحلیل کرد، آن‌گاه می‌توان امیدوار بود که شبیه‌سازی این شبکه‌ها در کامپیوتر، استفاده‌های علمی بسیاری داشته باشد. برای مثال می‌توان با استفاده از این سیستم‌ها برای بیماری‌های ژنتیکی یا بیماری‌هایی که در اثر بهم خوردن تعادل فعالیت ژن‌ها در سلول ها به‌وجود می‌آیند. برای مثال برخی از سرطان‌ها، داروهای مناسب طراحی نمود و بررسی تاثیر آن‌ها را بر روی وضعیت سلول و در نتیجه موجود زنده انجام داد[۳].

در سال‌های اخیر، تکنولوژی جدیدی به‌نام ریزآرایه[1]های DNA، امکان مطالعه هم‌زمان بیان هزاران ژن را در یک آزمایش برای محققان فراهم می‌کند. مدل‌سازی شبکه‌های ژنی از داده‌های تجربی ریزآرایه، تشریح اعمال سلولی در سطح مولکولی را ساده‌تر می‌سازد. داده‌های سری زمانی ریزآرایه به صورت یک ماتریس عددی با هزاران سطر (ژن) و چند ده ستون (نقطه زمانی) هستند. این داده‌ها مشکلات خاصی دارند که تحلیل آن‌ها را مشکل می‌کند. برخی از این مشکلات شامل وجود نویز و تعداد نمونه‌های کم در مقایسه با تعداد پارامترهای مجهول است.

تنظیم ژنی، یک مساله غیرخطی است و مدل‌های خطی، قادر به اخذ تعاملات غیرخطی بین ژن‌ها نیستند، بنابراین لازم است که مدل ارائه شده قادر به در نظر گرفتن جنبه غیرخطی روابط بین ژن ها نیز باشد. هم‌چنین روش‌ای مطرح شده باید قادر باشد که نویز موجود در داده‌ها و هم‌چنین غیر یقینی بودن سیستم تنظیم ژن‌ها را در نظر گرفته و علاوه بر آن قابل تعمیم به شبکه های بزرگ ژنتیکی باشد. روش‌های مختلفی برای مدل‌سازی شبکه‌های تنظیم‌کننده ژنتیکی با استفاده از داده-

[1] MicroArray

های سری زمانی ریزآرایه پیشنهاد شده‌اند که از جمله آن‌ها می‌توان به شبکه‌های بولی، شبکه‌های بولی احتمالاتی، مدل‌های خطی، معادلات دیفرانسیل، شبکه‌های بیزی، شبکه‌های بیزی دینامیک و یا روش‌های مبتنی برخوشه‌بندی اشاره کرد.

هدف این پایان‌نامه آن است که یک مدل برای استنتاج شبکه‌های تنظیم ژنی با استفاده از داده-های سری زمانی بیان ژن معرفی گردد، به قسمی که قادر باشد تمام موارد ذکر شده را در نظرگرفته شود. به این منظور از روش‌هایی برای پیش‌بینی ارتباطات پیچیده میان ژن‌ها استفاده شده است که مبنای شبکه‌های عصبی-فازی دارد. هم‌چنین روشی جهت استخراج روابط از روی داده‌های روش ریزآرایه پیشنهاد شده که ساختاری نوین پویا و همراه با بهینه‌سازی دارد. روش پیشنهادی، ژن‌هایی که بیشترین تاثیر بر هم دارند را به عنوان ژن‌های تنظیم‌کننده می‌یابد و نوع روابط آن‌ها شامل اثر سدکنندگی، فعال‌کنندگی و یا خنثی، مشخص می‌کند اگرچه در ابتدا قصد بر آموزش شبکه با مجموعه داده‌های مربوط به اسید فسفاتازهای یک گیاه خاص بود، اما به دلیل عدم دسترسی به مجموعه داده نام‌برده، به عنوان جایگزین، از مجموعه داده‌های استاندارد ریزآرایه مربوط به ۱۲ ژن شاخص موثر در مدت سیکل جوانه زدن نوعی مخمر[1] استفاده شد. این داده‌ها شامل نمونه‌های جمع-آوری شده در نقاط زمانی مختلف از سیکل سلولی است که جهت نمایش خروجی استخراج شده و در نتیجه تعیین بازدهی با مقایسه روابط بین ژنی به‌دست آمده تجربی، استفاده شده است.در نهایت شبکه تنظیم ژنی بر اساس نتایج به‌دست آمده ترسیم و در نهایت با مقایسه نتایج با روش‌های پیشین، کارایی روش تائید می‌شود.

در این پایان‌نامه ابتدا در فصل دوم مفاهیم اولیه ژن، بیان ژن و تنظیم بیان ژن را شرح داده سپس در فصل سوم به مرور برخی روش‌های موجود در این زمینه می‌پردازیم. در فصل چهارم روش پیشنهادی برای استنتاج این شبکه ها و در فصل پنجم نتایج به‌دست آمده را شرح خواهیم داد. در فصل آخر نیز جمع‌بندی و سپس پیشنهاداتی برای ادامه کار مطرح خواهد شد.

[1] Saccharomyces Cerevisiae

فصل دوم

# تنظیم ژن وتحلیل داده‌های بیان ژن

## ۲-۱- مقدمه

در این فصل به معرفی دانش نو و تازه متولد بیوانفورماتیک و ویژگی‌ها و مفاهیم آن پرداخته و جایگاه این پایان‌نامه در مباحث بیوانفورماتیکی روشن خواهد شد. مطالبی که در این فصل بیان خواهد شد شامل آشنایی کوتاهی با علم بیوانفورماتیک و مفاهیم پایه‌ای در ارتباط با پایان‌نامه و شبکه‌های تنظیم ژن است.

در ارتباط با استنتاج شبکه‌های تنظیم ژنی، داشتن اطلاعات حول فرآیند تنظیم ژن و شبکه‌های تنظیم ژنی لازم است، بنابراین در این فصل کلیاتی از این مقوله ارائه خواهد شد. هدف آن است که چنان‌چه سخنی از ژن، بیان ژن، استخراج مقادیر آن توسط ریزآرایه و روش‌های آن به میان آمد، ابهامی وجود نداشته باشد. هر چند سعی بر آن بود که به‌صورت معقولی، فهم صورت مساله و درک مدل پیشنهادی، به مفاهیم زیستی وابستگی خاص تخصصی پیدا نکند، اما درک نسبی از مفاهیم زیستی مرتبط، کمک به دریافت مختصات بحث و فهم بهتر موضوع خواهد کرد.

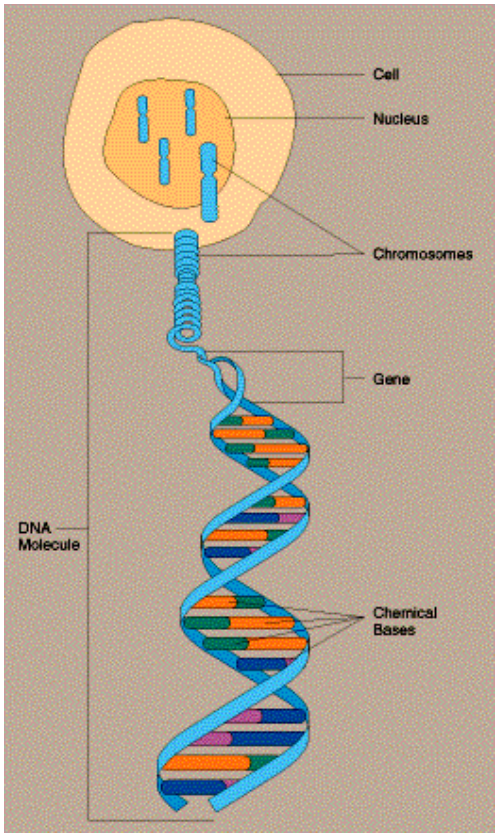
## ۲-۲- ژن

ژن‌ها واحد وراثت هستند. آرایش ژنتیکی یک موجود زنده (ترکیب ژن‌های آن)، تعیین‌کننده مشخصات آن، مانند رنگ چشم‌های یک جانور یا بوی گل یک گیاه، است. بیش‌تر ژن‌ها اطلاعات مربوط به ساخت پروتئین ها را در بر دارند و در توالی‌های مولکول DNA ذخیره می‌شوند.

همه موجودات زنده به جز ویروس‌ها از سلول تشکیل شده‌اند. به عنوان مثال مخمر یک موجود تک سلولی است، در حالی که در بدن انسان تریلیون‌ها سلول وجود دارد. در هسته سلول‌ها کروموزوم و در کروموزوم DNA وجود دارد(شکل۲-۱). DNA از دو بخش کدکننده و غیرکدکننده تشکیل شده که بخش کدکننده آن‌ها را ژن می‌نامند، بنابراین ژن‌ها قسمت‌های رشته‌های DNA[1] یا RNA[2] هستند که ویژگی‌های خاصی از یک موجود زنده را کد می‌کنند. برای این‌که یک ژن بتواند اثر خود را نمایان سازد باید ابتدا به پروتئین ترجمه شود. پروتئین‌ها مولکول‌های بزرگی هستند، که اساس هر

[1] Deoxyribo Nucleic Acid
[2] Ribo Nucleic Acid

ارگانیزمی را تشکیل می‌دهند. ترجمه ژن‌ها به واسطه ماکرومولکول‌های دیگری به‌نام RNA انجام می‌شود. همه سلول‌ها در یک ارگانیزم ژن‌های مشابهی دارند، اما این ژن‌ها در زمان‌های مختلف و در شرایط مختلف، بیان متفاوتی دارند.



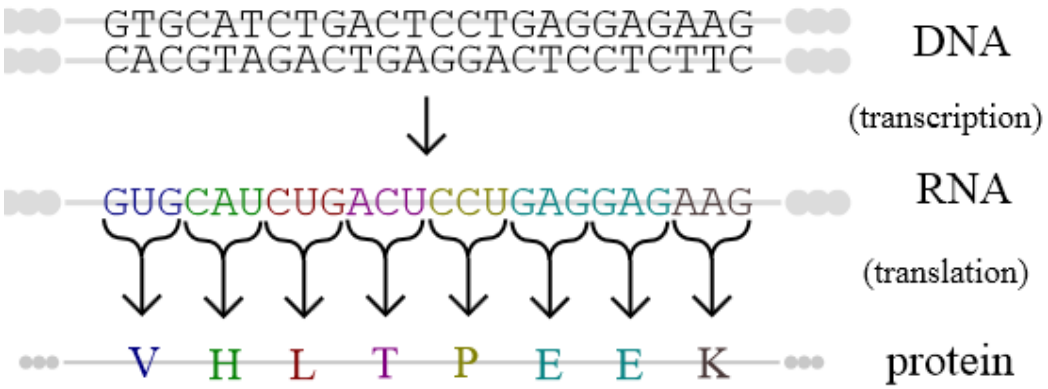

شکل۲-۱ ساختار کلی سلول [۵]

## ۲-۳- بیان ژن

بیان ژن[۱] فرایندی است که طی آن با استفاده از اطلاعات موجود در ژن، فراورده‌های ژنی عامل[۲] تولید می شوند. محصول ژن‌ها بیشتر پروتئین ها هستند و از محصولات غیر پروتئینی می‌توان به

[1] Gene Expression
[2] Functional Gene Products

rRNA[1]، tRNA[2] و snRNA[3] اشاره کرد. مراحل مختلفی را می‌توان برای فرایند بیان ژن در نظر گرفت که شامل رونویسی[4]، اتصال RNA، ترجمه[5] و تغییرات پس از ترجمه یک پروتئین است. تنظیم ژن به سلول این امکان را می‌دهد تا بتواند ساختار وکاربرد خود را کنترل کند و این مسئله پایه‌ای است برای تفاوت‌های سلولی، دگرگونی (تکامل) و مهارت تطبیق موجودات زنده با شرایط جدید. تنظیم ژن هم‌چنان می‌تواند به عنوان یکی از زیر لایه‌های تکامل در نظر گرفته شود، زیرا کنترل زمان‌بندی و مقدار ژن می‌تواند تاثیرات مهمی در عملکرد ژن‌ها درون یک سلول یا کل موجود پرسلولی داشته باشد. در علم ژنتیک، بیان ژن یکی از مهم‌ترین مسائل بنیادی است که کمک می‌کند تا ژنوتیپ[6] به‌صورت فنوتیپ[7] ظاهر شود. در واقع کدهای ژنتیکی که در رشته‌های DNA ذخیره شده‌اند به‌وسیله بیان ژن تفسیر می‌شوند و نحوه بیان باعث به‌وجود آمدن فنوتیپ در موجود زنده خواهد شد. طی این فرایند، از اطلاعات موجود در ژن‌ها برای تولید پروتئین ها استفاده می‌شود. این فرآیند شامل رونویسی از ژن‌ها به مولکول‌های mRNA و سپس تولید پروتئین‌ها با استفاده از اطلاعات موجود در mRNA ها (فرایند ترجمه) است. این مراحل درشکل ۲-۲ نشان داده شده‌اند.




شکل ۲-۲ رونویسی از روی DNA و ترجمه از روی RNA و تولید یک پروتئین[۴]

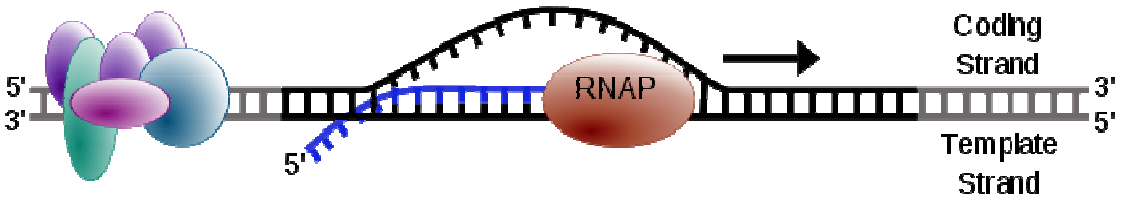
[1] Ribosomal RNA
[2] Transfer RNA
[3] Small Nuclear RNA
[4] Transcription
[5] Translation
[6] Genotype
[7] Phenotype

همان‌طور که در شکل۲-۲ نیز مشهود است DNA از رشته‌های تکرار شونده متشکل از واحدهای سازنده‌ای از جنس نکلئوتید ساخته شده و هر رشته از طریق بازهای آلی در هر دو رشته به یکدیگر متصل می‌شوند. این اتصال بین دو باز آلی نوکلئوتیدهای دو طرف رشته است. به این بازهای متصل به هم باز مکمل گفته می‌شود. بازهای آلی به چهار شکل سیتوزین C، گوآنین G، تیمین T و آدنین A وجود دارند که از این میان، باز آدنین مکمل تیمین، و باز گوآنین مکمل سیتوزین است. در فرآیند بیان ژن RNA با کپی برداری از روی قسمتی از رشته DNA ساخته می‌شود(در هسته) و با ترجمه از روی RNA نیز پروتئین مورد نظر تولید می‌شود(در سیتوپلاسم).

با این‌که در یک موجود خاص، تمام سلول‌ها مجموعه ژن‌های مشابهی دارند، این الگوی بیان ژن‌ها در یک سلول است که ساختار و عملکرد آن‌را تعیین می‌کند. ژن‌ها در سلول‌های مختلف و در شرایط محیطی مختلف و زمان‌های مختلف بیان متفاوتی دارند. به همین دلیل مقدار پروتئین های تولید شده نیز متفاوت است. عاملی که ساختار، وظیفه و عملکرد دو نوع سلول را از یکدیگر متمایز می‌کند، مجموعه ژن‌های فعال در هر کدام است. ژن فعال، ژنی است که برای رونویسی آماده است. وضعیت تنظیمی یک ژن را می‌توان با اندازه‌گیری سطح بیان mRNA آن اندازه‌گیری نمود[۵].

## ۲-۴- رونویسی

هر مولکول DNA شامل دو رشته است که هر کدام دو سر ’3 و’5 دارند (شکل۲-۳). مناطق کد شده شامل اطلاعات ژنتیکی هستند و قسمت‌هایی که اطلاعات ژنتیکی ندارند در تولید RNA ها کمک می‌کنند. تولید یک کپی RNA از روی DNA را فرایند رونویسی می‌نامیم، این فرایند به‌وسیله RNA Polymerase انجام می‌شود که در هر لحظه یک نوکلئوتید RNA را به رشته RNA در حال تولید، اضافه می‌کند.




شکل ۲-۳ فرآیند رونویسی از روی RNA به‌وسیلهRNA polymerase[۶]

رشته RNA با رشته DNA اصلی یکسان است، تنها با این تفاوت که در آن به جای تیمین[1]( T ) از یوراسیل[2]( U) استفاده شده است. عمل رونویسی در پروکاریوت[3] ها بوسیله یک پلیمراز انجام می شود اما دریوکاریوت ها این عمل بوسیله سه پلیمر مختلف صورت می‌گیرد که پلیمراز اول برای رونویسی rRNA ها است و پلیمراز دوم برای رو نویسی mRNA،snRNA و... استفاده می‌شود و پلیمراز سوم در ترجمه tRNA،rRNA کاربرد دارد. عمل رونویسی پس از رسیدن به یک توالی مشخص خاتمه می‌یابد[۷].

## ۲-۵- ترجمه

برای برخی از ژن‌ها، محصول نهایی همان RNA ای است که در مراحل قبل تولید شده، اما برای بسیاری مانندmRNA ها این آخرین مرحله نیست. mRNA ها به همراه خود کد یک یا چند توالی پروتئین را حمل می‌کنند. در هر توالی mRNA اطلاعات ژنتیکی به‌وسیله کدهای سه تایی به نام کدون (codon) که پشت سر هم قرار دارند ذخیره شده است. در فضای سیتوپلاسم برای هر کدام از این کدون ها یکtRNA با یک سه تایی مکمل وجود دارد که همراه خود آمینواسید مشخصی را حمل می‌کند. سپس این آمینو اسیدها به‌وسیله یک ریبوزوم و هم‌چنین ترتیب کدون‌ها بر روی mRNA به هم متصل می‌شوند و توالی پروتئینی به‌دست می‌آید. از روی هر رشته mRNA می‌توان تعداد زیادی توالی پروتئین ساخت[۷].

## ۲-۱- تنظیم ژن

سلول برای انجام وظایف مختلف و در زمان‌ها و شرایط مختلف، به پروتئین‌های متفاوتی نیاز دارد. برای مثال سلول‌ها پس از قرار گرفتن در یک شرایط جدید مانند کمبود مواد غذایی، استرس و یا شوک گرمایی و یا در حین چرخه‌های حیاتی مثل رشد، با فعال کردن یک برنامه رونویسی جدید و تولید پروتئین های مورد نیاز عکس العمل نشان می‌دهد. تنظیم بیان ژن یا تنظیم ژن[4] به کنترل

[1] Thymine
[2] Urasil
[3] Prokaryote
[4] Gene Regulation