Quase um século se passou para que aquela substancia gosmenta isolada na cozinha de um frio castelo alemão por Friedrich Meisher em 1989 fosse reconhecida como a determinante das características dos seres vivos. A maioria dos cientistas na época apostava nas proteínas presentes na cromatina, pois combinações de 20 blocos estruturais poderiam explicar melhor a diversidade biológica, mais do que as quatro formas químicas do DNA: adenina, timina, citosina e guanina.

Os experimentos realizados em plena guerra mundial por Oswald Avery, Colin McLeod e Maclyn McCarthy, proporcionaram fortes evidencias de que a informação genética estaria no DNA a partir da comprovação do princípio transformante (descrito por Griffths em 1929) em bactérias causadoras de pneumonia. Contudo, ainda restavam dúvidas até que Martha Chase, Alfred Hershey e um liquidificador, deram provas definitivas sobre DNA como responsável pela perpetuação da informação genética.

Tal descoberta deu início a uma espécie de competição no âmbito cientifico para desvendar como os componentes dessa molécula estavam dispostos. Nesse cenário, pesquisadores por toda Europa e nos Estados Unidos entraram nessa empreitada, entre eles James Watson e Francis Crick. O primeiro realizava seu doutoramento no laboratório de Cavendish pertencente ao departamento de Física da Universidade de Cambridge, estudando difração de raios-X em polipetídeos sob a orientação de Max Perutz. Watson por sua parte, era formado em zoologia nos Estados Unidos e após um período em Copenhagen, mudou-se para Cavendish para realizar um pós-doutoramento em cristalografia de globinas sob a supervisão de John Kendrew. Entretanto, a pesar de terem projetos diferentes, ambos rapidamente compartilharam suas inquietudes em relação à estrutura do DNA e começaram recopilar evidências, e montar um quebra cabeças a partir da construção de modelos e dedução lógica:

1. 1929 – Phoebus Levene, já tinha descrito que o DNA como um polímero de nucleotídeos, cada um formado por um grupo fosfato, um açúcar desoxirribose, e uma base nitrogenada que podia ser de quatro tipos: adenina, timina, citosina e guanina.
2. 1947 – Erwin Chargaff estudando DNA de diferentes organismos, tinha chegado à conclusão de que as proporções de cada nucleotídeo se mantinham constantes entre células do mesmo organismo e entre organismos da mesma espécie, mas que entre espécies diferentes essas proporções variavam. Também, e mais importante, que as proporções de adenina sempre correspondiam a com as de timina, e as de citosina com as de guanina, mas, A+T não se correspondia com C+G, caracterizando o que agora chamamos de regras de Chargaff.
3. 1952 – Alexander Todd, um professor de química orgânica da Universidade de Cambridge, percebeu que as ligações fosfodiester entre nucleotídeos adjacentes sempre aconteciam no sentido 5´-3´.
4. Nesse mesmo ano Rosalyn Franklin e Raymond Gosling do King´s College obtiveram a foto 49 a partir de difração de raios-x em cristais não-hidratados de DNA, na qual ficava evidente que essa molécula era uma hélice com um diâmetro de 20ºA.
5. Também em 1952, Jerry Donahue, quem foi realizar um pós-doutoramento de 6 meses em Cavendish e dividiu a sala com Watson e Crick, observou que a ligação N-glicosídica deveria ser perpendicular ao eixo formado pelos fosfatos e açúcares.
6. Linus Pauling, quem era uma dos melhores competidores da corrida, publicou em seguida um artigo na PNAS onde sugeria um modelo para o DNA como uma hélice tripla com os fosfatos no interior e as bases se projetando para fora. A pesar de frustrante, a nossa dupla dinâmica rapidamente percebeu que nesse modelo os fosfatos não estariam ionizados e consequentemente não constituía um ácido.
7. Por último, em 1953, Rosalind Franklin e seu ajudante conseguiram a famosa foto 51 a partir de cristais hidratados, onde era facilmente perceptível que o DNA era uma dupla hélice, que cada volta tinha 34 Aº e que os fosfatos estariam na parte externa.
8. Se o DNA tinha 20 Aº de diâmetro, e era uma dupla hélice, ao parear duas pirimidinas, o diâmetro seria menor, o oposto aconteceria se fossem pareadas duas pirimidinas. Desse modo somente poderiam pariar uma pirimidina com uma purina, e isso respeitava as regras de Chargaff.
9. Por último, faltava descobrir como as duas hélices eram ligadas uma com a outra, e a resposta veio a partir das ligações de hidrogênio previamente descritas por Linus Pauling em 1950. O rearranjo então daria duas pontes de hidrogênio entre adenina e timina e três pontes entre guanina e citosina.

Assim, em 23 de abril de 1953, Watson e Crick publicaram na Nature uma das descobertas mais importantes da biologia, a estrutura do DNA. O modelo proposto foi fundamental em três pontos:

1. Ele explicava como a disposição dos nucleotídeos poderia especificar a sequência de aminoácidos em uma proteína;
2. Trocas de nucleotídeos poderiam explicar a ocorrência de mutações;
3. Sugeria um modelo de cópia que explicaria a perpetuação da informação.

Logo após, a comunidade cientifica começou a especular sobre diferentes modelos de como a duplicação poderia ocorrer. Assim, apareceram três modelos mais promissores: o modelo conservativo, onde as fitas se abririam como um zíper e cada uma serviria de molde para a cópia da fita complementar, tendo duas cadeias filhas, cada uma composta por uma fita “velha” e uma “nova”; um modelo conservativo, onde a fitas serviriam de molde mas a cadeia “velha se manteria com tal” e a cadeia filha seria formada por duas fitas novas; e o terceiro e menos provável modelo dispersivo, onde após a duplicação ambas as fitas seriam compostas por trechos provenientes das fitas molde e as fitas novas.

Esse dilema foi elegantemente desvendado por Matt Meselson e Franklin Stahl. Eles cultivaram bactérias em meio de cultivo com uma fonte de nitrogênio pesado (N15) durantes várias gerações e depois essas bactérias “marcadas” foram transferidas para meio de cultivo com nitrogênio leve (N14) e a cada geração extraíram DNA e acompanharam seu peso através de centrifugação em gradientes de Césio. Na geração zero, todo o DNA extraído das bactérias crescidas em N15, apresentavam uma única banda pesada, ou seja mais no fundo do tubo. Já na Geração 1, onde as bactérias teriam passado por um ciclo de replicação em N14, a banda única se apresentou em uma posição superior, mais leve. Essa fase do experimento descartava o modelo conservativo pois se fosse verdade, deveria ter aparecido duas bandas no tubo de centrifugação, uma leve e uma pesada. Já na geração dois, os pesquisadores observaram que o DNA formava duas bandas, mas desta vez, uma banda correspondia a banda intermediaria e a outra a uma banda mais leve ainda, ou seja, somente formada por N14.

Ainda, foi visto que a medida que passavam as gerações, a banda intermediaria se tornava cada vez mais fina, e a banda leve proporcionalmente mais larga, descartando assim o modelo dispersivo.

Mas mesmo com a confirmação de que a replicação era semi-conservativa, como o processo aconteceria? Desde o ponto de vista químico já era sabido que a polimerização dos nucleotídeos no sentido 5´-3´era resultado do ataque nucleofílico do grupo hidroxila 3´no fosfato alfa do nucleotídeo trifosfato a ser incorporado. Essa reação acarretaria na liberação de um grupo pirofosfato, e a subsequente hidrólise do mesmo liberaria energia suficiente para fazer o processo irreversível.

Contudo, o pontapé inicial para o entendimento paulatino do mecanismo per se e seus componentes foi dado por duas descobertas fundamentais. A primeira foi o isolamento de uma enzima capaz de sintetizar DNA a partir de uma fita molde por Arthur Kornberg em 1956, a qual batizou de DNA polimerase. A segunda foi a descrição de John Cairns da Universidade Nacional da Austrália, do cromossomo bacteriano. Na sua publicação “The Bacterial Chromosome and its manner of replication as seen by autoradiography”, ele demonstrou que o cromossomo de *E. coli* não só era uma molécula circular única, senão que se replicava a partir de um único ponto (ao que denominou origem de replicação), e a mediada que duas forquilhas avançavam em sentidos opostos, aumentava o tamanho do que chamou bolha de replicação.

Hoje sabemos que a duplicação do DNA é um mecanismo que envolve grandes complexos proteicos, mas ao mesmo tempo é rápido e preciso. De forma geral, podemos dividi-lo em três etapas:

1. Iniciação: onde acontece o reconhecimento da origem de replicação, separação das fitas complementares, estabilização do complexo e começo da síntese.
2. Alongamento: onde as fitas novas são estendidas pelo replissomo a partir da cópia das fitas molde;
3. Terminação: onde a síntese é concluída, os fragmentos são ligados, e os cromossomos duplicados são separados um do outro.

Em procariotos, a duplicação do DNA começa com o reconhecimento da origem de replicação chamado OriC, uma sequência de 245 pb, que se caracteriza por apresentar dois motivos, um formado por 3 repetições de 13 bases cada (13-mer) ricas em A-T, e 4 repetições de 9 bases cada uma (9-mer), que servem de sítio de ligação para a proteína DnaA. Esta última ao reconhecer o sitio de origem se liga de forma cooperativa (até 20 unidades) à essa região de 60-80 pb e com o auxílio de ATP gera uma força contorsional levogira que favorece a desnaturação da região 13-mer adjacente permitindo o recrutamento de DnaB, uma helicase de DNA que medeia a abertura das fitas molde mediante a quebra das pontes de hidrogênio.

Existem 12 helicases em *E. coli*, todas atuam como hexâmeros mantendo uma estrutura similar em formato de anel que se liga ao DNA de fita simples. Contudo elas não são capazes de se ligar sem o auxílio de carregadores como a DnaC que mediante o uso de ATP produz uma quebra na helicase, abrindo o anel e permitindo sua montagem na fita simples.

Uma vez em posição, DnaB, recruta DnaG ou primase, um monômero de 60 kDa responsável pela síntese de um *primer* ou iniciador que fornece à DNA polimerase o grupo 3´-OH para a extensão da nova fita. Ao mesmo tempo, o DNA de fita simples que vai sendo gerado é recoberto por tetrâmeros da proteína SSB (do inglês *single strand binding*). Estas proteínas de 74 kDa, se ligam também de forma cooperativa e por interação eletroestática independente de sequência, para proteger as fitas simples de degradação.

A primase então, sintetiza um pequeno fragmento de RNA de 11 nucleotídeos a partir do reconhecimento da sequência GTC, e a DNA polimerase começa a estender a fita com o auxílio de cátions de magnésio no seu centro catalítico que facilitam o ataque nucleofílico do 3´-OH do nucleotídeo recém incorporado, no fosfato alfa do nucleotídeo entrante.

A bactéria *E. coli*, apresenta várias DNA polimerases. Nos anos 60s, Paula de Lucia e John Cairns a procura de mutantes para a DNA polimerase descrita por Kornberg, observaram após analisarem 3477 colônias em um experimento de “Brute-Force-Mutant” que todos esses mutantes possuíam níveis normais de DNA polimerase, contudo a colônia 3478, não era capaz de crescer, apesar dos níveis da polimerase também serem normais. Dessa forma, devia existir uma outra DNA polimerase responsável pela duplicação do DNA. Logo, foi visto que a DNA polimerase inicial (a qual chamaram de DNApol I e seu gene *PolA* – acrônimo de Paula) tinha um papel no reparo do DNA e de fato a responsável pela duplicação do cromossomo bacteriano era a DNA polimerase III.

De forma geral, todas as DNA polimerases compartilham uma estrutura similar que se assemelha ao formato de uma mão direita, onde o DNA a ser replicado se acomoda no sulco formado entre os “dedos e a palma”, sendo o “polegar” responsável pela atividade exonuclease. Todas as DNApol são capazes de sintetizar 1000 nt/sec, no entanto, a processividade é muito baixa, sendo incluídos somente entre 10 e 200 nt antes de dissociar. A pesar disso, existem proteínas auxiliares capazes denominadas grampos deslizantes capazes de aumentar a processividade para 50.000 nt/sec.

A DNA pol III é diferenciada em relação ás outras pois forma um grande complexo multiproteico (holoenzima) de 900kDa: duas subunidades alfa (que compõem os centros catalíticos), uma subunidade épsilon com atividade exonucleoase 3´-5´, uma subunidade theta que auxilia na montagem do complexo, uma subunidade Tau que medeia a dimerização, uma subunidade beta que tem a função do grampo deslizante, e o carregador formado pelas subunidades gamma, psi, chi, e delta.

Mas, se a síntese só acontece no sentido 5´-3´, como seria sintetizada a fita complementar? Esse questionamento foi respondido em 1968 pelos pesquisadores da Universidade de Nagoia Reiji e Tsuneko Okazaki. Eles realizaram experimentos de “pulse and chase” e viram que de fato, uma das fitas era sintetizada de forma contínua, mas a outra, era sintetizada em pequenos fragmentos de 1000-2000 nucleotídeos que logo eram reconstituídos e incorporados em cadeias maiores. Dessa forma, a cadeia 5´-3´é sintetizada de forma contínua e a fita 3´-5´sintetizada de forma descontínua. Entretanto, a pesar de essa diferença, a síntese acontece de forma simultânea nas duas fitas, cada uma sintetizada por uma das subunidades catalíticas da polimerase que a medida que avança e sintetiza a cadeia contínua, forma uma alça que permite a síntese dos fragmentos de Okazaki também no sentido 5´-3´.

Assim, a medida que a DNApol III avança (uma para cada forquilha de replicação em sentidos opostos), os *primers* são removidos pela DnaH. Posteriormente, a DNApol I sintetiza os trechos faltantes e uma enzima ligase une os gaps mediando ligações fosfodiester entre fosfatos adjacentes. Entretanto, a natureza circular o cromossomo bacteriano, implica em mecanismos que garantam o termino da duplicação, caso contrário, as forquilhas de duplicação poderiam avançar indefinidamente. Para isto, assim como existe um ponto de começo, exatamente oposto a este, existe um grupo de sequências de terminação denominadas Ter, cada uma com 23 pares de bases não-palíndromos nomeados de A a J que sinalizam o termino e recrutam a proteína de 37 KDa Tus. Embora o mecanismos exatos de termino da duplicação ainda não se conheça totalmente, acredita-se que a ligação desta proteína bloqueie o avanço da polimerase e induza sua liberação.

O sucesso da duplicação do DNA bacteriano também depende de outros fatores adicionais, as topoisomerases, necessárias para a diminuição dos superenrolamentos decorrentes do avanço das forquilhas de duplicação na dupla hélice circular. Existem dois tipos básicos, as topoisomerases de tipo I, que realizam cortes transientes em uma das fitas do DNA mediante ligações covalentes com resíduos de tirosinas na sua estrutura que permitem o giro em sentido contrário, e as tipoisomerase de tipo II que realizam cortes de fita dupla (também usando ligações covalentes com resíduos de tirosina na sua estrutura) e permitam a passagem de uma fita através do corte para remover tensões contorcionais. Estas últimas também são fundamentais na resolução dos cromossomos, uma vez que a natureza circular do DNA, acarreta na formação de um concatenado.

A duplicação do DNA em eucariotos por outro lado, apesar de ser muito parecida com proteínas análogas, é obviamente mais complexo, sendo que ainda existem muitos fatores desconhecidos.

Para começar, nos eucariotos não existe uma origem de replicação definida por sequencias específicas, e sim, muitas origens distribuídas ao longo dos cromossomos lineares que se acredita sejam reconhecidas a partir da ausência de nucleossomos.

Em *Sacharomyces cerevisie* a duplicação começa com a ligação de ORC, um hexâmetro capaz de recrutar a helicase MCM2-7 com o auxílio dos carregadores Cdc6 e Cdt1. Uma vez no local de início, os carregadores e ORC são liberados e um novo par de carregadores localiza uma segunda unidade da helicase MCM2-7 no sentido contrário. A diferença das helicase procariota, esta não precisa de DNA de fita simples para se ligar. A medida que as helicases avançam, uma proteína SSB formada por três subunidades chamada RPA, recobre o DNA de fita simples, e a primase começa a sintetizar os iniciadores.

Da mesma forma que em procariotos, nos eucariotos existem várias DNA polimerases (16 conhecidas) também com funções diversas. Contudo, para a duplicação de fato, só são necessárias 3, a DNApol alfa, DNApol delta e a DNApol épsilon (a DNA pol beta, por exemplo somente atua durante o reparo de lesões no DNA).

A DNA pol alfa, é a que faz a função de primase sintetizando pequenos primers, mas a diferença da primase procariota, esta sintetiza um primeiro trecho de RNA de 10 nucleotideos seguidos por 20-30 nucelotidos de DNA. Uma vez sintetizado o primer, acontece uma troca da polimerases num processo denominado “pol switch”, onde a DNA pol alfa, é substituída pela DNA pol delta na fita contínua e pela DNA pol Epsilon na fita descontinua. A processividade das polimerases eucariotas é garantida pelo grampo deslizando PCNA carregado pelo fator RFC, contudo, nos eucariotos, o processo é muito mais lento, chegando a adicionar somente 2000 nucleotídeos por segundo.

Esta diferença em relação ao procariotas radica no fato do DNA eucarioto não estar nú, e sim associado a histonas formando nucleossomos. Dessa forma, a velocidade da duplicação do DNA depende da disponibilidade de unidade de histonas para a incorporação. Primeiramente, as histonas após sua síntese no citoplasma, devem ser transportadas ao núcleo, e sua inserção depende de chaperonas e complexos modificadores, num *timeline* preciso que garante a desestabilização dos nucleossomos á frete da forquilha de replicação e a remontagem após sua passagem. Embora muitos desses componentes não sejam conhecidos, sabe-se que nesse processo também participam PCNA e a helicase MCM2-7. As Chaperonas CAF1 e AF1, em primaira instancia tem o papel de inibir a ligação de das histonas H3 e H4 com as outras duas, mas após a acetilação das caudas aminoterminais destas, o tetrâmero formado por duas unidades de H3 e duas de H4 são incorporadas ao DNA, após isso, a chaperona NAP1 medeia a incorporação de dois dímeros formados por uma unidade da histona H2A e uma unidade de H2B cada uma. Também, sabe-se que a incorporação das histonas é aleatória, sendo os novos nucleossomos formados por histonas recém sintetizadas e histonas recicladas. Ainda, após a montagem dos novos nucleossomos, complexo modificadores garantem que as modificações epigenéticas presentes na cromatina antes da duplicação sejam perpetuadas.

Ao mesmo tempo, a medida que as forquilhas avançam, os primers devem ser removidos, entretanto neste caso, não há uma RNAse, senão que a própria DNA pol epsilon continua sintetizando ao encontrar os iniciadores, empurrando as bases de RNA. Esses trechos “expostos” para fora da hélice são removidos posteriormente pela endonuclease FEN1 (que também participa no reparo por excisão de bases) e os trechos unidos por ligações fosfodiesteer mediadas por uma ligase.

Por outro lado, a natureza linear do s cromossomos eucariotos impõe um problema, pois na fita descontinua, a remoção dos primers leva a perda desse trecho nas extremidades dos cromossomos, e consequentemente a cada ciclo de replicação os cromossomos encurtariam e genes seriam perdidos.

Em 1985, Elizabeth Blackburn durante seu pós-doutoramento, observou que os microcromossomos de *Tetrahymana sp*. apresentavam várias repetições dos nucleotídeos TTGGGG nas suas extremidades. Num experimento realizado junto com Jack Szostak, onde os microcromossomos desse protozoário sem as terminações eram inseridos em leveduras, eles observaram que os cromossomos exógenos eram rapidamente degradados. Contudo, quando as sequencias de repetição eram incorporadas nos cromossomos a serem inseridos nas leveduras, uma vez dentro das células, esses cromossomos eram mantidos como tais, chegando a conlcução de que essas repetições telomêrica tinham a função crucial de evitar a degradação dos cromossomos eucariotos. Posteriormente, Elizaberth junto a sua pós-doutoranda Carol Greider, isolaram uma enzima que denominaram telomerase capaz de estender os telomeros e garantir a perpetuação da informação (lhes rendendo um premio Nobel dividivo com Szostak).

Hoje sabemos que a telomerase, é um complexo ribonucleico formado por dois componentes principais, o RNA TERC (com tamanho variável entre organismos) e uma transcriptase reversa denominada TERT que junto dom a disquerina, NOP10, NHP2 e GAR1 formam o componente proteico. Assim, a cada ciclo, a telomerase é capaz de acrescentar repetições no extremo da cadeia continua usando o TERC para pareamento complementar, e a transcriptase reversa para adicionar novas bases á extremidade. Isso permite com que na fita descontinua possam ser sintetizados novos *primers* e a polimerase seja capaz de estender a fita e assim evitar o encurtamento. Contudo, sabe-se que a telomerase não é expressa em todas as células do organismo, e eventualmente o encurtamento telomerico determina a “vida útil” da célula, uma característica observada muitos anos atrás por Hayflick.

Uma outra característica dos telômeros é a presença de um complexo multiproteico chamado de Shelterin que em humanos é composto por TRF1, TRF2 RAP1, TN2, TPP1 e POT1. Estas proteínas não só regulam o funcionamento da telomerase, senão que também servem de proteção antidegradação do telômero mediando a formação de uma alça denominada T-loop através da invasão de ~30 bases de DNA fita simples.

Assim, mesmo com alguns componentes ainda desconhecidos, o entendimento dos processos e os componentes subjacentes à duplicação do DNA, tem trazido não só riqueza ao conhecimento do funcionamento dos seres vivos, senão que também tem trazido grandes avanços em vários âmbitos da Biologia. Dentro destes, a evento mais revolucionário foi a invenção da reação em cadeia pela polimerase por Kary Mullis em 1983. Através dela, hoje em dia qualquer sequência de DNA conhecida pode ser amplificada em um tubo, mediante a inclusão de um molde, nucleotídeos trifosfato, magnésio, e uma DNA polimerase termoresistente isolada da bactéria *Thermus aquaticus*. Essa metodologia tem servido de alicerce para a identificação de genes, clogangem e sequenciamento, sendo fundamental para o aprimoramento de diagnósticos clínicos, e para o crescimento de diversas áreas incluindo a biotecnologia, agroindústria, ciências forenses, estudos evolutivos e muitas outras aplicações.

Ainda o conhecimento da maquinaria de replicação, tem proporcionado meios de combater doenças como o câncer, evidenciado pela produção de drogas que interferem com o metabolismo do DNA, como os inibidores de topoisomerases II, VP-16 e VM-26, inibidores de topoisomerases do tipo I com o Irinotecan, desenvolvimento de análogos de base, inibidores de telomerase, e muitos mais.