METODE ZA DOLOČEVANJE ZAPOREDJA GENOMA:

NASLEDNJA (NEXT) IN TRETJA GENERACIJA VISOKOZMOGLJIVEGA SEKVENCIRANJA

Naslednja generacija (HTS-NG) – PCR amplifikacija za ojačitev signala

- Sekveniranje z amplifikacijo na kroglicah (Roche/454FLX)

- **Sekvenciranje s sintezo** (Illumina/Solexa Genome analyzer)

-**Sekvenciranje z ligacijo** (Applied Biosystems SOLID System)

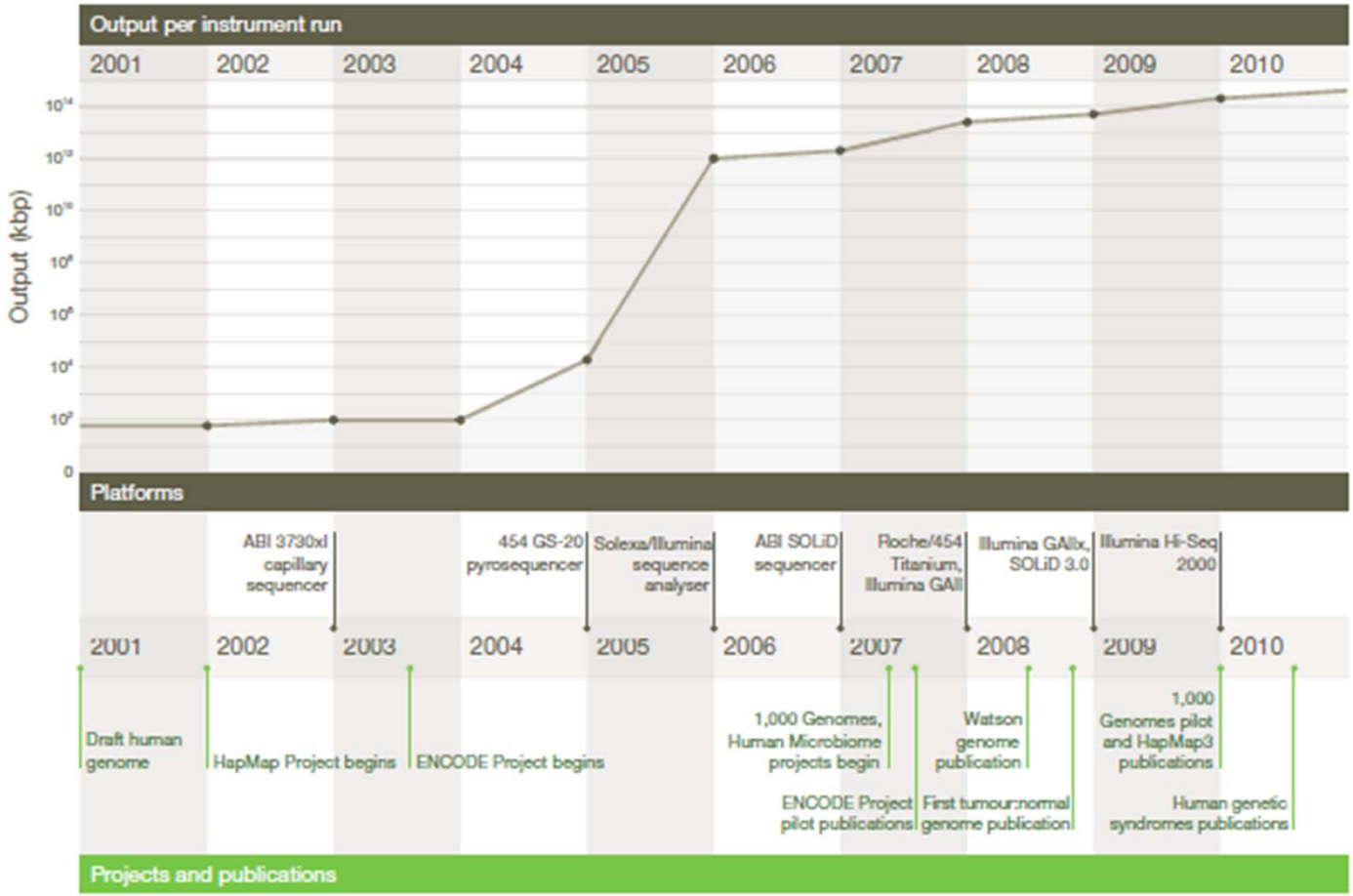
-.

Tretja generacija – možno sekvenciranje brez ampifikacije

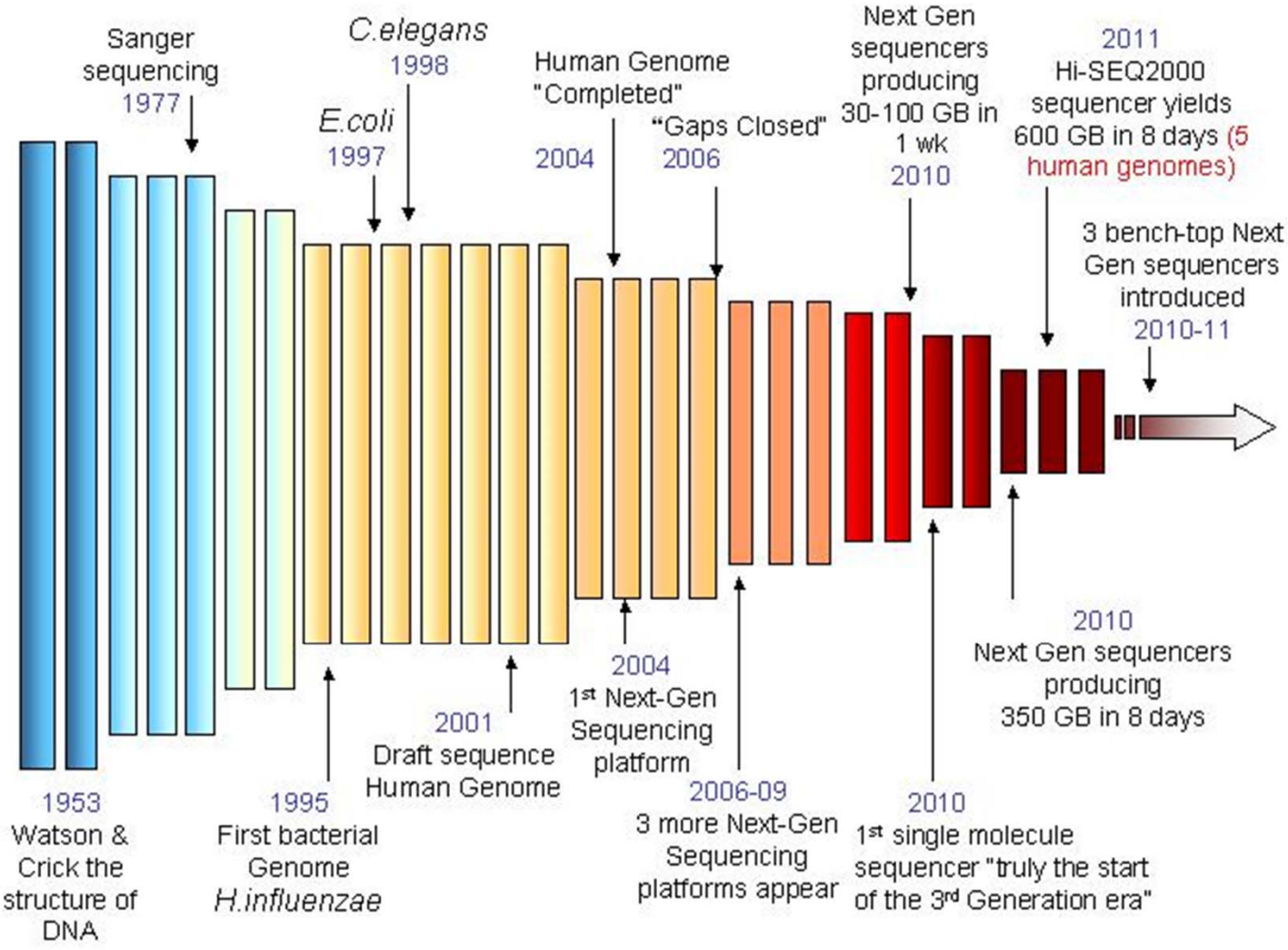
- **Heliscope** (tSMS – true single molecule sequencing, 2007)
- **SMRT** (single molecule real time sequencer)
- **RNAP** (single molecule real time sequencer)
- **Nanopore** DNA sequencer
- **Ion Torrent, Ion Proton** (ABI) – Še vedno potrebuje PCR, a se uvršča med 3. generacijo, ker za detekcijo ni potrebna optika
- ……..

A decade's perspective on DNA sequencing technology

Elaine R. Mardis[1]



Časovna skala revolucije sekvenciranja



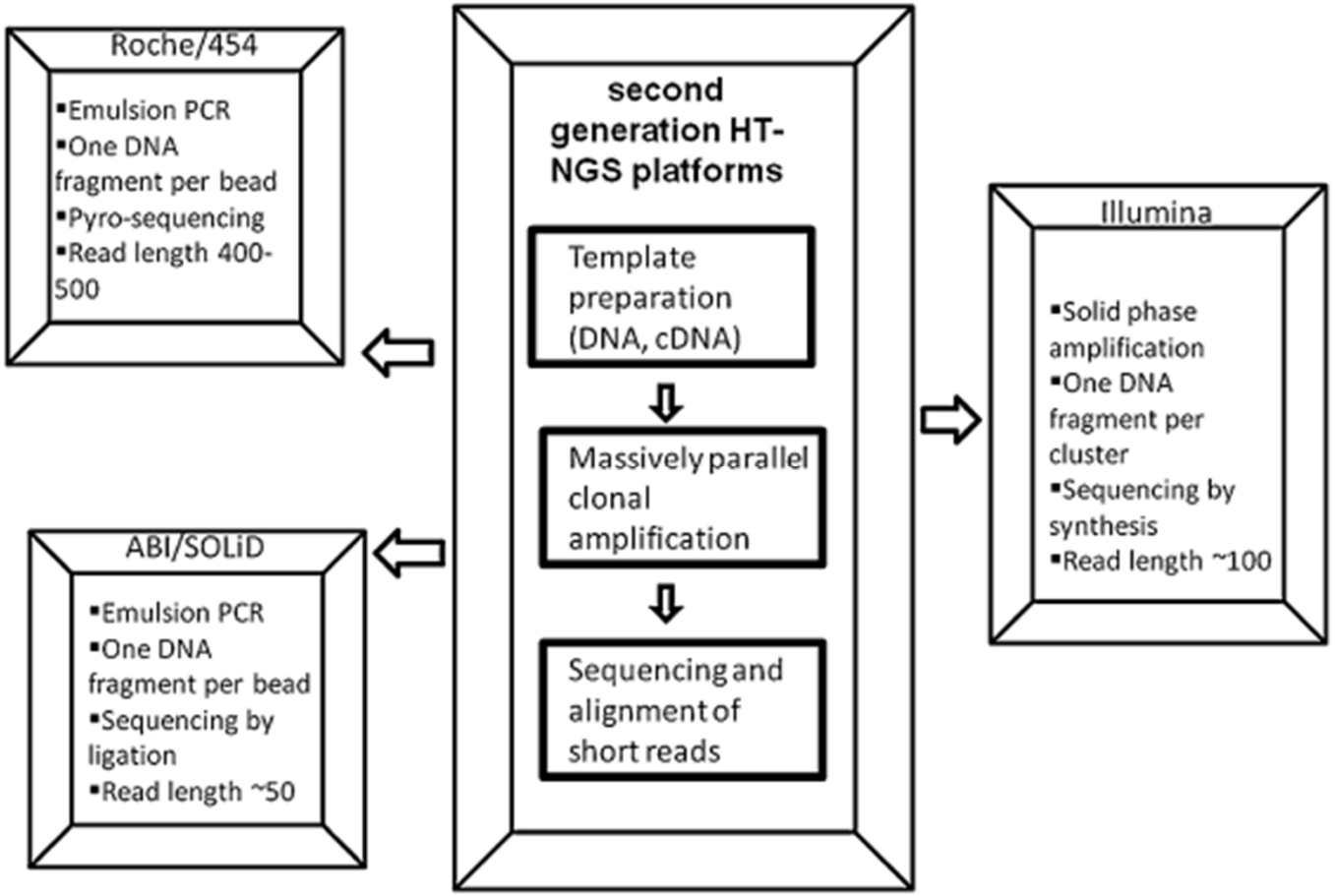
http://www.ipc.nxgenomics.org/newsletter/no11.htm

Eksplozija novih tehnologij: Muzeji ZDA že iščejo relikvije iz zgodnjih dni genomike !!

1996: “Blue Baby” prototip aparata PCR

Na desetine znanstvenih muzejev v ZDA želi ohraniti relikvije genomske revolucije (Museum genomics initiative”)

SKUPNE LASTNOSTI IN POSEBNOSTI TREH GLAVNIH PLATFORM NASLEDNJE GENERACIJA SEKVENCIRANJA



J Appl Genetics (2011) 52:413–435

Skupne lastnosti NG-HTS:

-Kompleksnost encimskih reakcij, kemije, programske in strojne opreme, optike, itd.

-Premočrtna priprava knjižnic (vzorcev) pred sekvenciranjem.

-Priprava knjižnic fragmentov DNA s prileganjem za platformo značilnih podaljškov (angl. Linker) in PCR amplifikacija.

-PCR amplifikacija enoverižnih fragmentov knjižnice in sekveniranje pomnoženih fragmentov. *(Roche – emulzijske kroglice; ABI – kloni kroglic; illumina – kloni mostičkov)*

PRIMERJAVA METOD 2. GENERACIJE SEKVENCIRANJA S SANGERJEVO METODO

Table 1: (a) Advantage and mechanism of sequencers. (b) Components and cost of sequencers. (c) Application of sequencers.

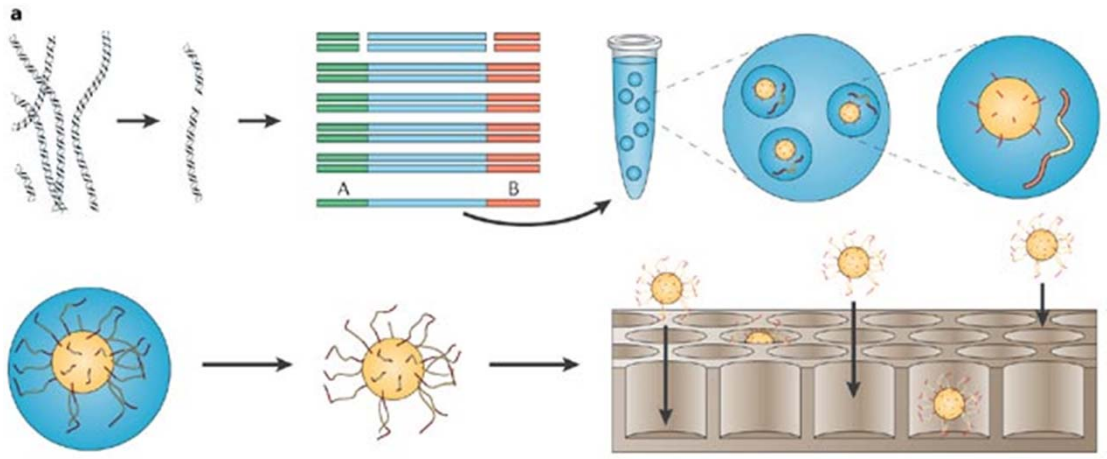
(a)

Sequencer	454 GS FLX	HiSeq 2000	SOLiDv4	Sanger 3730xl
Sequencing mechanism	Pyrosequencing	Sequencing by synthesis	Ligation and two-base coding	Dideoxy chain termination
Read length	700 bp	50SE, 50PE, 101PE	50 + 35 bp or 50 + 50 bp	400~900 bp
Accuracy	99.9%*	98%, (100PE)	99.94% *raw data	99.999%
Reads	1 M	3 G	1200~1400 M	—
Output data/run	0.7 Gb	600 Gb	120 Gb	1.9~84 Kb
Time/run	24 Hours	3~10 Days	7 Days for SE 14 Days for PE	20 Mins~3 Hours
Advantage	Read length, fast	High throughput	Accuracy	High quality, long read length
Disadvantage	Error rate with polybase more than 6, high cost, low throughput	Short read assembly	Short read assembly	High cost low throughput

(b)

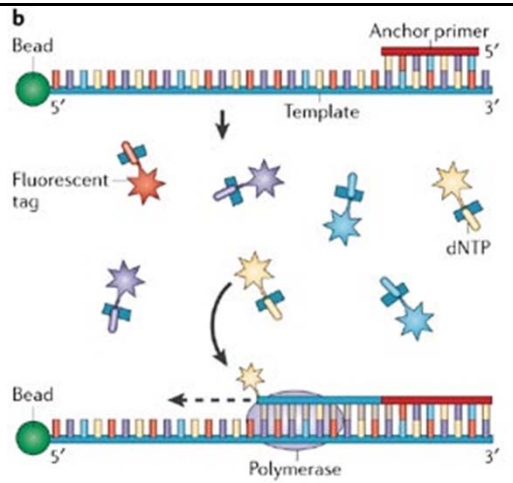
Journal of Biomedicine and Biotechnology
Volume 2012, Article ID 251364, 11 pages

Splošni koncepti priprave klonalnih mrež in sekveniranje

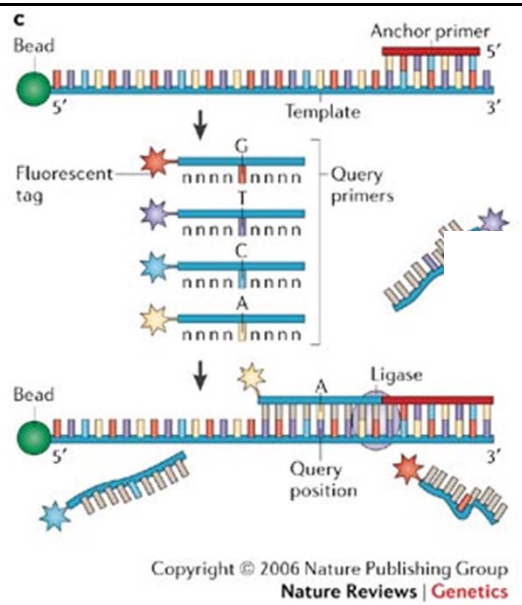


Bead chips

Sequencing by synthesis

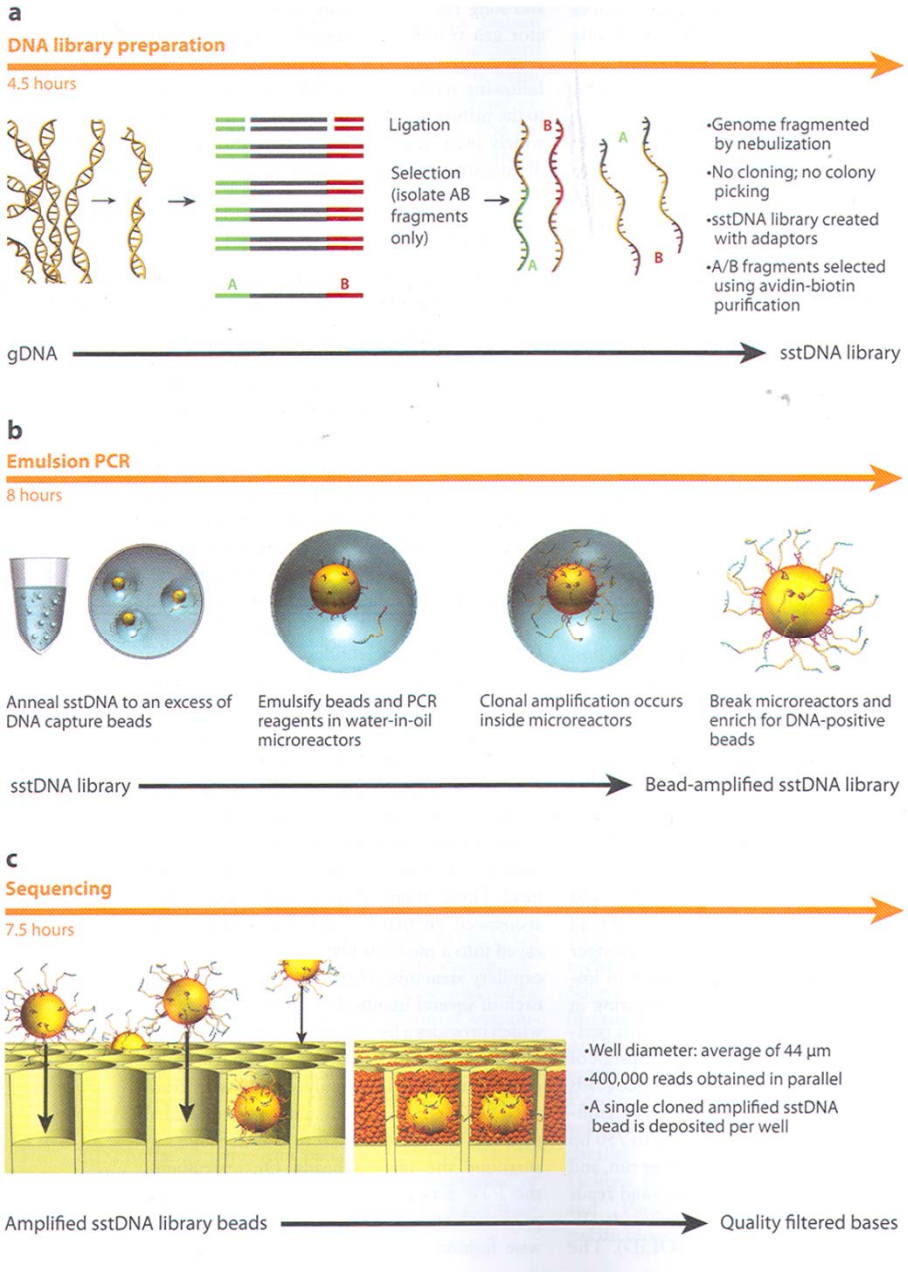


Sequencing by ligation



Fan *et al. Nature Reviews Genetics* **7**, 632–644 (August 2006) | doi:10.1038/nrg1901

Tehnogija 454



Library fragments are mixed with agarose beads with oligos complementary to adapter sequences on the library.

Each bead is associated with a single fragment.

Each fragment-bead complex is isolated into individual oil:water micelles with PCR mixture.

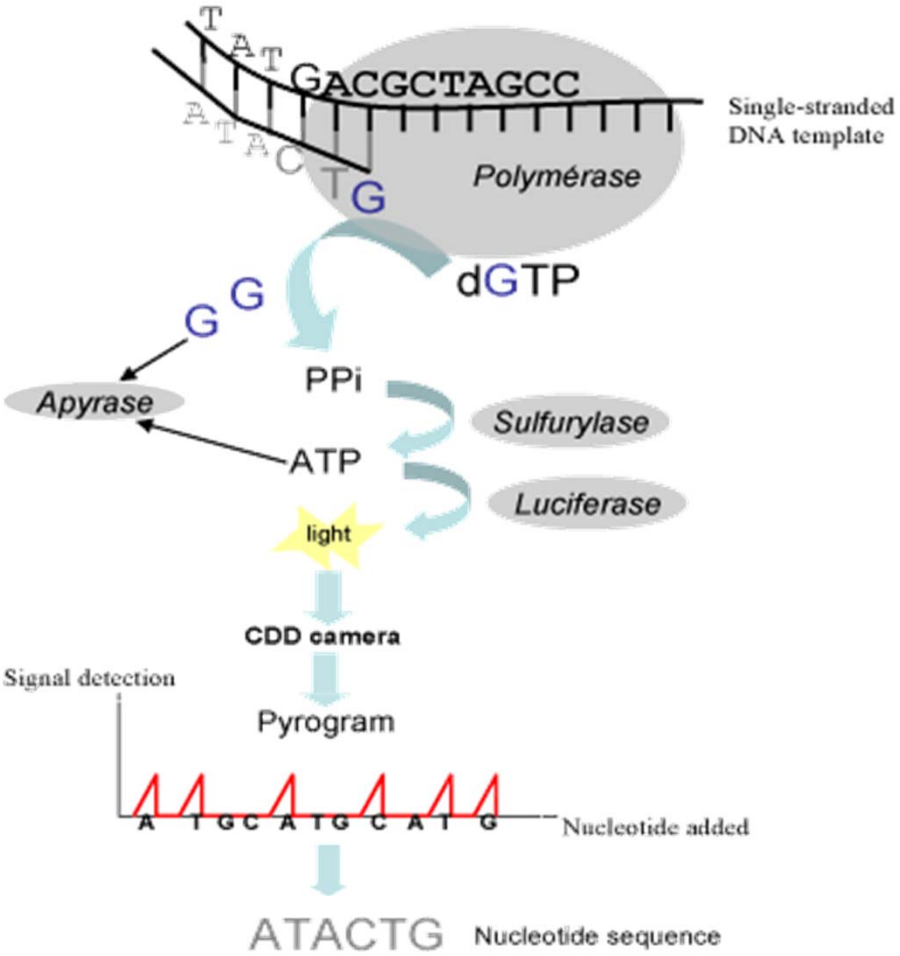
Thermal cycling of this emulsion PCR of the micelles produces amplified unique sequences on the bead surface.

“En mass” sequencing of PCR products on picotiter plates (PTP) with single beads in each picowell.

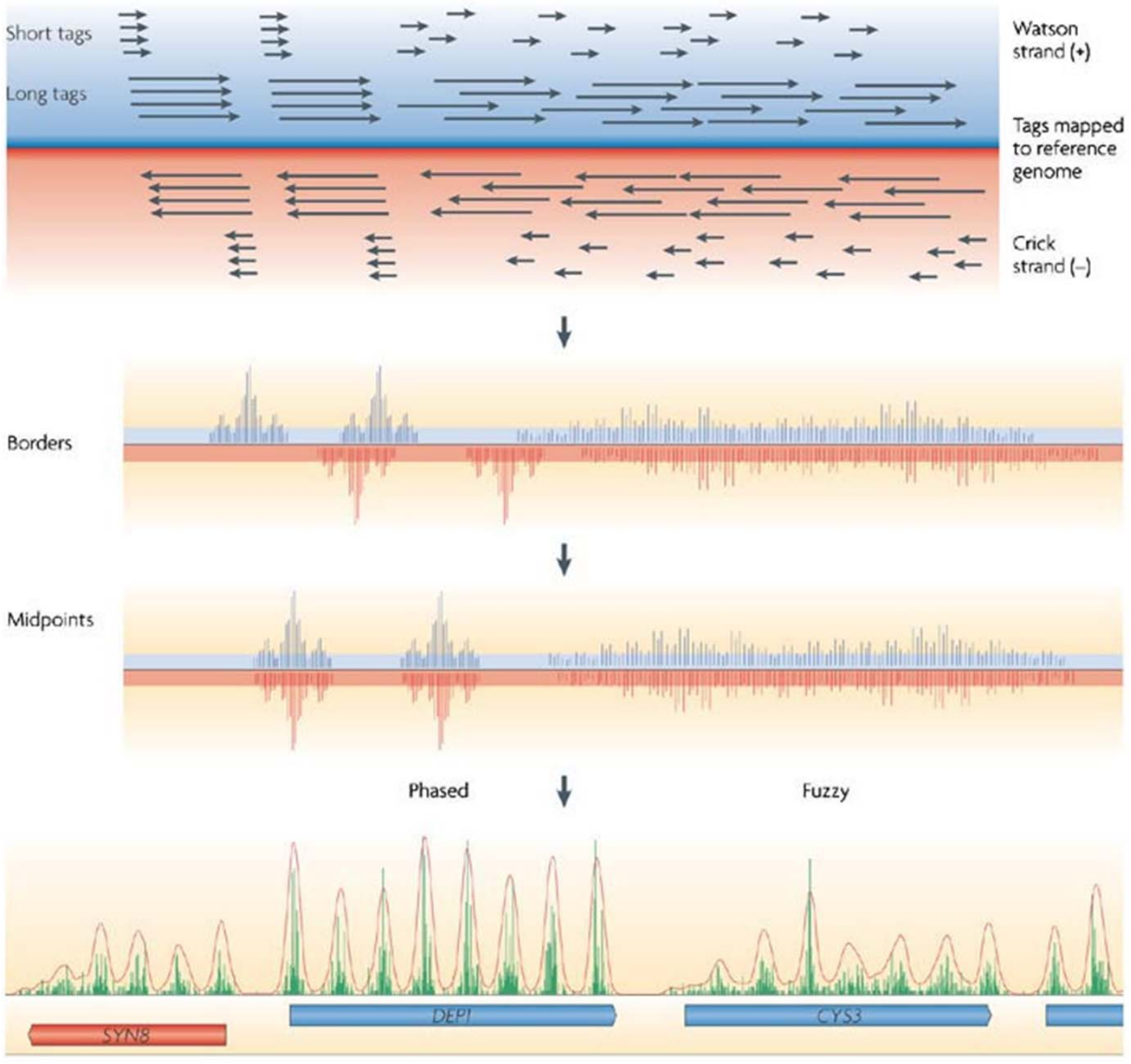
Enzyme/substrate containing beads for the pyrosequencing reaction are added to wells that act as floww cells for addition of individual pure nucleotide solutions. The CCD camera records the light emitted at each bead.

Mardis E.R. Annual Review of Genomics and Human Genetics 9: 387-403 (2008).

Tehnogija 454 - Princip pirosekvenciranja

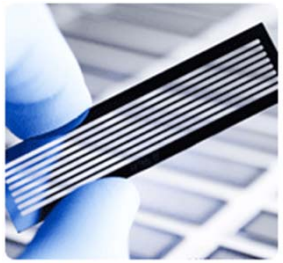


Izpis Watsonovega in Crickovega genoma po pirosekvenciranju




James Watson's genome sequenced at high speed : Nature News - Windows Internet Explorer
http://www.nature.com/news/2008/080416/full/452788b.html
File Edit View Favorites Tools Help
James Watson's genome sequenced...
comments on this story
Stories by subject
Biotechnology
Business
Cell and molecular biology
Genetics
Health and medicine
Stories by keywords
Craig Venter
Human genome sequencing
Next-generation sequencing technology
James Watson
454 LifeSciences
This article elsewhere
Blogs linking to this article
Published online 16 April 2008 | Nature | doi:10.1038/452788b
News
James Watson's genome sequenced at high speed
New-generation technology takes just four months and costs a fraction of old method.
Meredith Wadman
The first full genome to be sequenced using next-generation rapid-sequencing technology is published today (see page 872)[1], marking another milestone in the extraordinarily fastmoving field of human genome sequencing.
It took just four months, a handful of scientists and less than US$1.5 million to sequence the 6 billion base pairs of DNA pioneer James Watson. The achievement is first proof of principle that these rapid-sequencing machines can decipher large, complex genomes (see page 819)[2]. Made in this case by Connecticut-based 454 Life Sciences — a division of Roche Diagnostics — they allow many more sequencing reactions to proceed at the same time, on the same surface, than the previous generation of machines that produced the inaugural human genomes[3,4]. That change has had big pay-offs in speed, efficiency and, ultimately, cost (see Table 1).
most recent
commented
Fat cell numbers stay constant through adult life
05 May 2008
"Raft" offers route for drug treatments
02 May 2008
Climate troubles brewing for beer makers
02 May 2008
Of myths and men
02 May 2008
'Ocean deserts' are growing
01 May 2008
Related stories
Ready or not
10 April 2008
Watson's folly
25 October 2007
Common sense for our genomes
18 October 2007
Genomics: The personal side of genomics
04 October 2007
All about Craig: the first 'full' genome sequence
06 September 2007
Internet
100%
start
2007
FG2-2008
7 Internet ...
Inbox - Micro...
3:39 PM

Tehnologija illumina



is based on arrays of randomly assembled glass (silica) beads;
the beads have oligonucleotides covalently attached to the surface;
each bead has about one million oligos on its surface;
all oligos on each bead have the same sequence

Attached DNA fragments are extended and **bridge amplified** to create an ultra-high density sequencing flow cell with 80-100 million clusters, each containing ~1,000 copies of the same template. These templates are sequenced using a robust four-color DNA sequencing-by-synthesis technology that employs reversible terminators with removable fluorescent dyes. This novel approach ensures high accuracy and true base-by-base sequencing, eliminating sequence-context specific errors and enabling sequencing through homopolymers and repetitive sequences.

the beads are randomly assembled on the arrays, and the location of a particular probe is initially unknown;a process called decoding is used to find the location of each bead;

Illumina sekvenciranje s sintezo

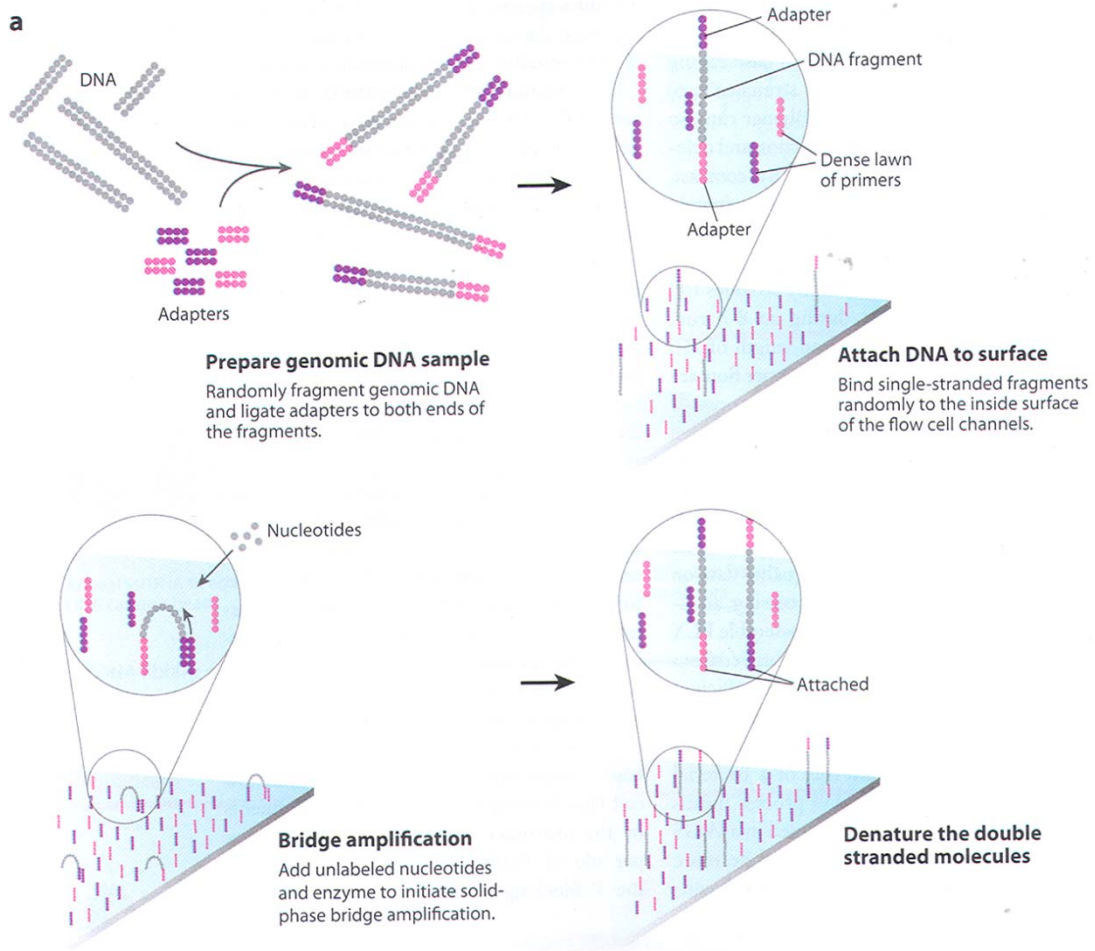
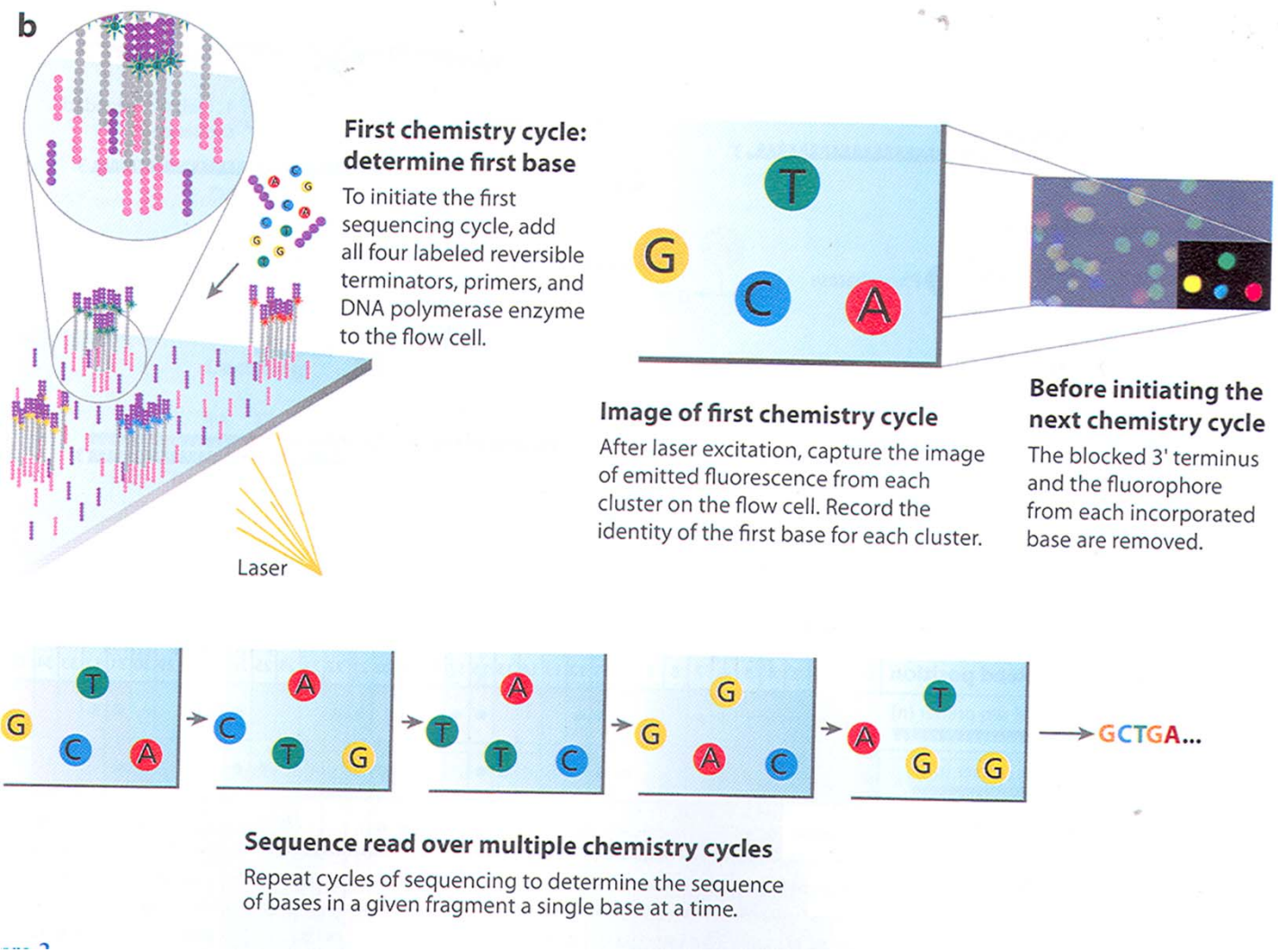


Figure 2

The Illumina sequencing-by-synthesis approach. Cluster strands created by bridge amplification are primed and all four fluorescently labeled, 3′-OH blocked nucleotides are added to the flow cell with DNA polymerase. The cluster strands are extended by one nucleotide. Following the incorporation step, the unused nucleotides and DNA polymerase molecules are washed away, a scan buffer is added to the flow cell, and the optics system scans each lane of the flow cell by imaging units called tiles. Once imaging is completed, chemicals that effect cleavage of the fluorescent labels and the 3′-OH blocking groups are added to the flow cell, which prepares the cluster strands for another round of fluorescent nucleotide incorporation.

Illumina sequencing by synthesis

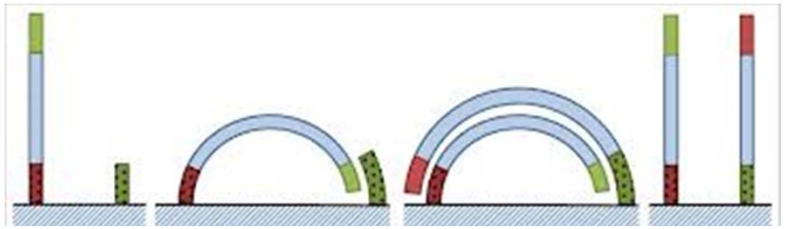


PCR z mostovi (Bridge PCR)

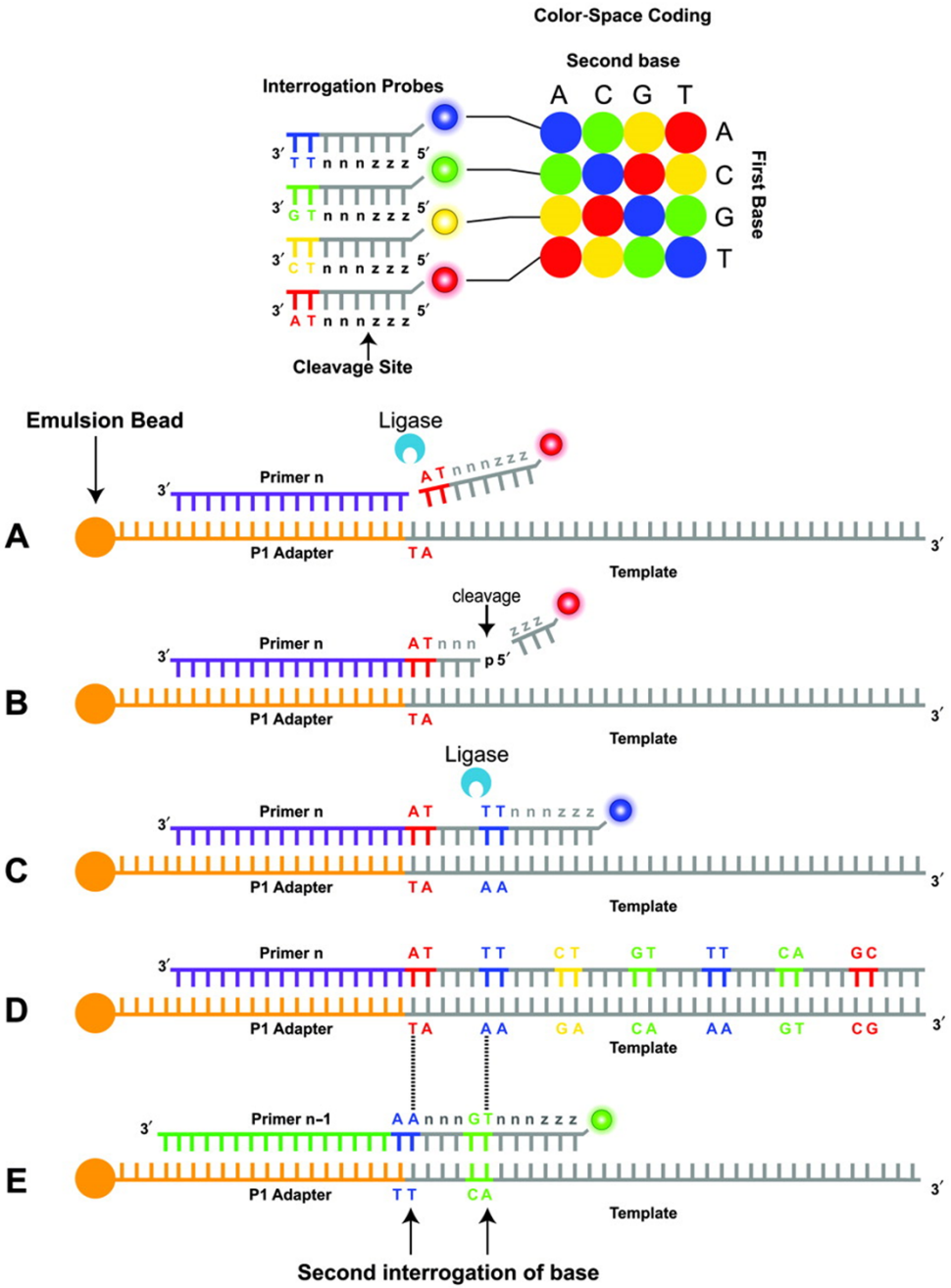
Začetni ologonukleotidi niso v raztopini ampak so naključno razporejeni in vezani na trdno podlago (silica steklo).

Molekule, ki jih želimo pomnožiti (knjižnica) morajo imeta adaptorje, ki so komplementarni oligonukelotidom na površini. Knjižnico razporedimo po površini (važna je pravilna koncentracija) in kjerkoli posamezna molekula pristane, bo našla komplementarni oligonukleotid.

V reakciji PCR molekule, ki jih pomnožujejo, niso trdno vezane na površino (vezani so le oligoukuleotidi). Po prvem ciklu PCR pa se oligonukleotid, ki je vezan na podlago, podajša na podlagi sekvence, ki jo pomnožujemo. Tako zdaj to zaporedje ostane vezano na površino z enim koncem. Drugi konec je prost, in v 2. ciklu PCR tvori most s komplementarnim oligonukelotidom na površini. Sedaj se ta oligonukleotid podaljša . Končni rezultat so skupine molekul, ki so se kopirale preko mostov med matrico in na površino vezanih "forward" in "reverse" oligonukelotidov.



ABI- Sekvenciranje z ligacijo

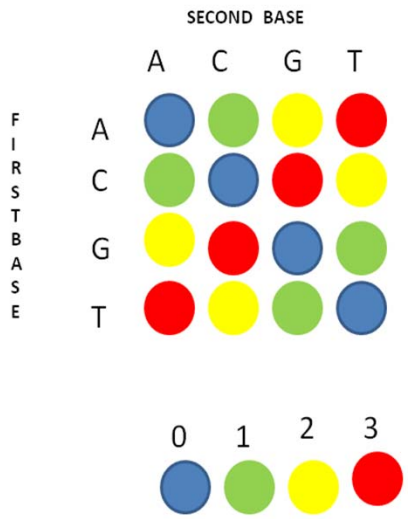


Vsaka proba je oktamer, ki ima 2 specifičnu bazi, ki se prilegata na zaporedje in 6 degeneriranih baz (nnnzzz) z eno fluorescento oznako na 5’ koncu.

Zaporedje 2 specifičnih baz da 16 2-baznih kombincaij (4 (2)).

Pri vsakem koraku sekvenciranja se na matrico prelije vseh 16 2-baznih kombinacij. Po prileganju in ligaciji detektiramo fluorescenco preden se odcepijo zadnje 3 degenerirane baze. Sledi prileganje in ligacija naslednje probe.

Mehanizem dekodiranja z barvno kodo po sekvenciranju z ligacijo



	A	C	G	T
A	0	1	2	3
C	1	0	3	2
G	2	3	0	1
T	3	2	1	0

AA	0
CA	1
GA	2
TA	3
AC	1
CC	0
GC	3
TC	2
AG	2
CG	3
GG	0
TG	1
AT	3
CT	2
GT	1
TT	0

Example:
[0 -> AA, GG, CC, TT ; 1 -> CA, AC, TG, GT; 2-> GA, TC, AG, CT; 3 -> TA, GC, CG, AT]
>44_35_267_F3
T2022021320300011100012222322112122 2

T2 -> TC (number 2 can be GA,TC,AG,CT: but only TC starts with T, so the first number is deciphered to 'TC')
0 -> CC
2 -> CT
2 -> TC
0 -> CC
2 -> CT
1 -> TG
3 -> GC
2 -> CT
0 -> TT and so on...

So, the colorspace translates into CCTCCTGCTT.....

Detektorski oligonukleotidi so 8-meri z eno fluorescentno nalepko. Skupaj je **1024 različnih detekcijskih oligonukleotidov.**

Če je prva baza znana, drug bazo lahko "izračunamo" s pomočjo dekodirne tabele.

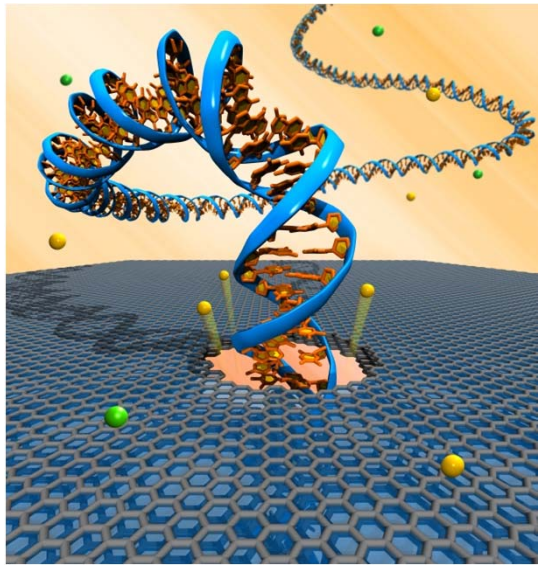
TEHNIČNE SPECIFIKACIJE IN CENA VISOKOZMOGLJIVIH APARATOV SEKVENCIRANJA 2. GENERACIJE

(b)

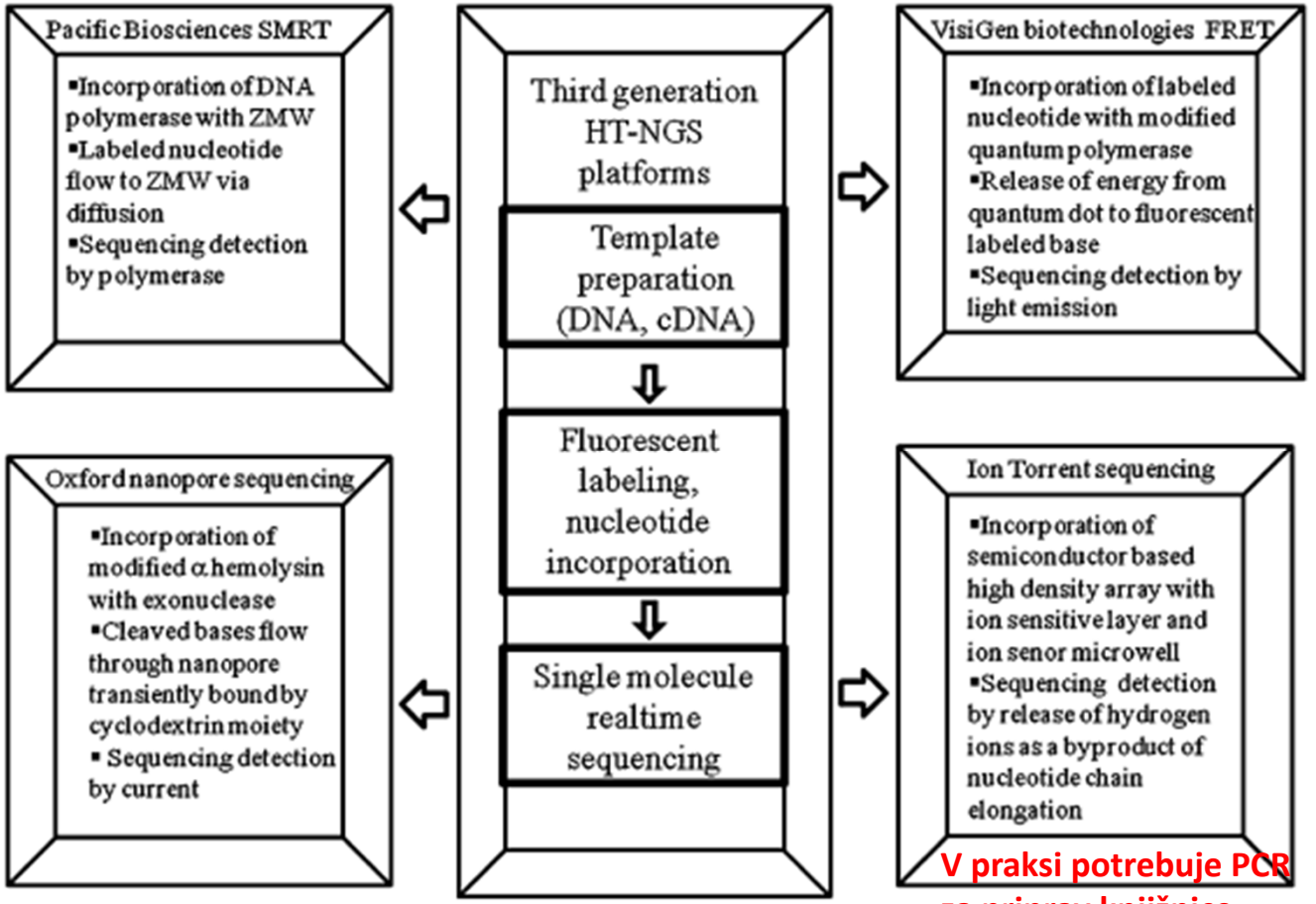
Sequencers	454 GS FLX	HiSeq 2000	SOLiDv4	3730xl
Instrument price	Instrument $500,000, $7000 per run	Instrument $690,000, $6000/(30x) human genome	Instrument $495,000, $15,000/100 Gb	Instrument $95,000, about $4 per 800 bp reaction
CPU	2* Intel Xeon X5675	2* Intel Xeon X5560	8* processor 2.0 GHz	Pentium IV 3.0 GHz
Memory	48 GB	48 GB	16 GB	1 GB
Hard disk	1.1 TB	3 TB	10 TB	280 GB
Automation in library preparation	Yes	Yes	Yes	No
Other required device	REM e system	cBot system	EZ beads system	No
Cost/million bases	$10	$0.07	$0.13	$2400

TRETJA GENERACIJA SEKVENCIRANJA

- **Heliscope** (tSMS – true single molecule sequencing, 2007)
- **SMRT** (single molecule real time sequencer)
- **RNAP** (single molecule real time sequencer)
- **Nanopore** DNA sequencer
- **IBM** sequencer
- **Ion Torrent** DNA sequencer* (potrebuje PCR)



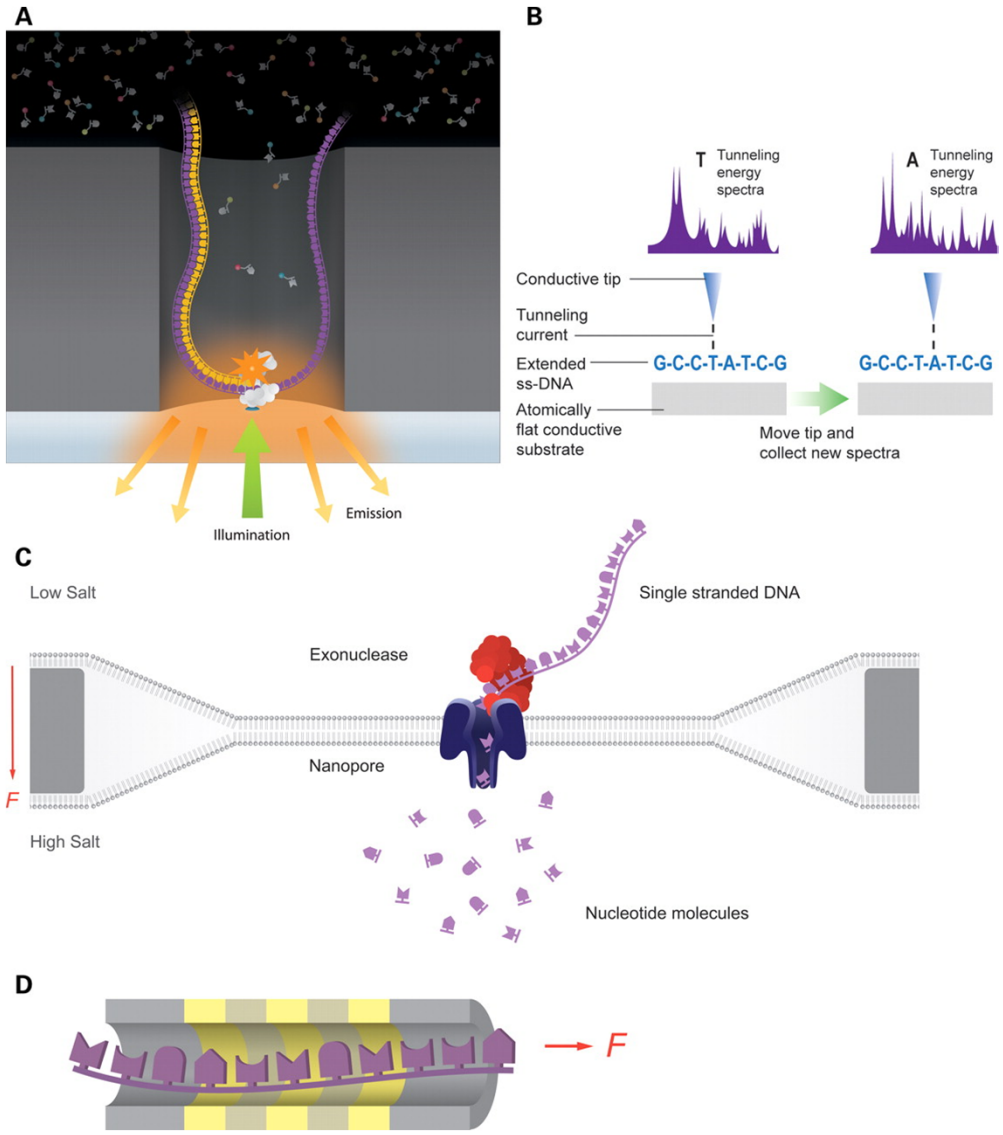
SKUPNE LASTNOSTI IN POSEBNOSTI TRETJE GENERACIJE SEKVENCIRANJA



V praksi potrebuje PCR za priprav knjižnice

J Appl Genetics (2011) 52:413–435

KAKO DELUJE 3. GENERACIJA SEKVENCIRANJA?



Schadt E E et al. Hum. Mol. Genet. 2010;19:R227-R240

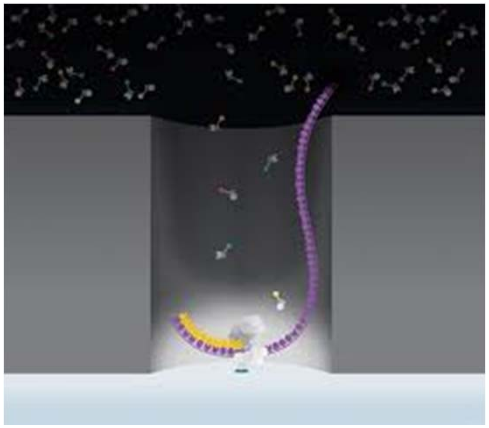
(A) Pacific Biosciences technology – direktno opazovanje DNA sinteze na eni molekuli DNA. Fiksirana DNA polimeraza, po dodajanju baz merimo fluorescenco gama-označenih fofonukleotidov.

(B) Mnoge firme razvijajo metode za detekcijo sekvenciranja prek elektronskega mikroskopa. Tehnologija Reveo – enojnoverižna DNA se raztegne in pregleda.

(C) Oxford Nanopore tehnologija- merjenje električne prevodnosti pri premiku molekule DNA skozi nanoporo.

(D) IBM tehnologija DNA tranzistorjev – meri električni tok pri premiku posameznih baz enojnoverižne DNA, ki se premikajo čez režo

SMRT (single molecule real time sequencer)



Vsaka SMRT celica ovsebuje tisoč “zero-mode waveguides” (ZMW).

Ena molekula DNA polimeraze je pripeta na dno vsakega ZMV, kar omogoča opazovanje polimerizacije posamezne molekule DNA polimeraze.

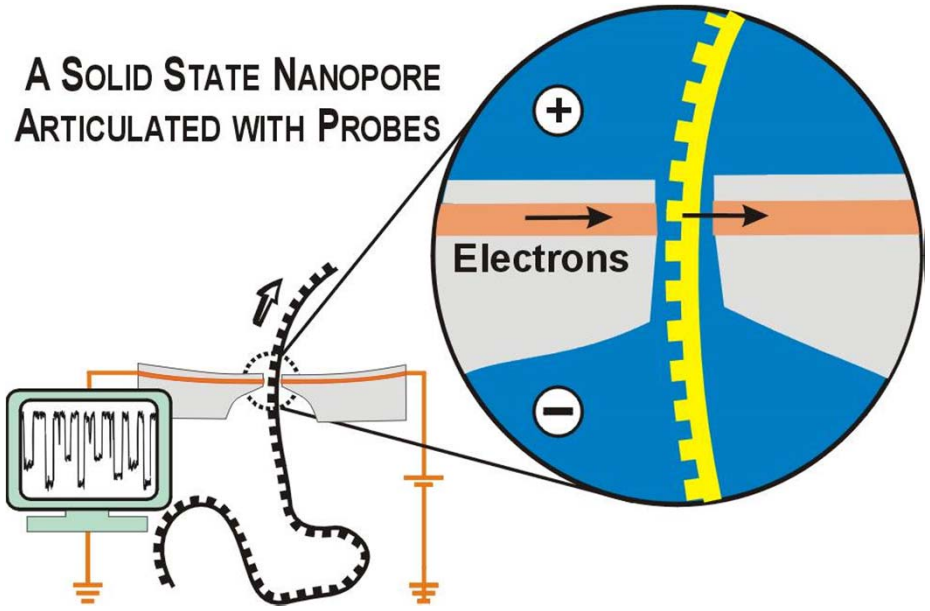
Opazujemo sintezo novonastale molekule DNA s fluorescentno označenimi dNTP.

SEKVENCIRANJE Z NANOPORAMI

Ni barvnega označevanja nukleotidov in CCD detekcije.

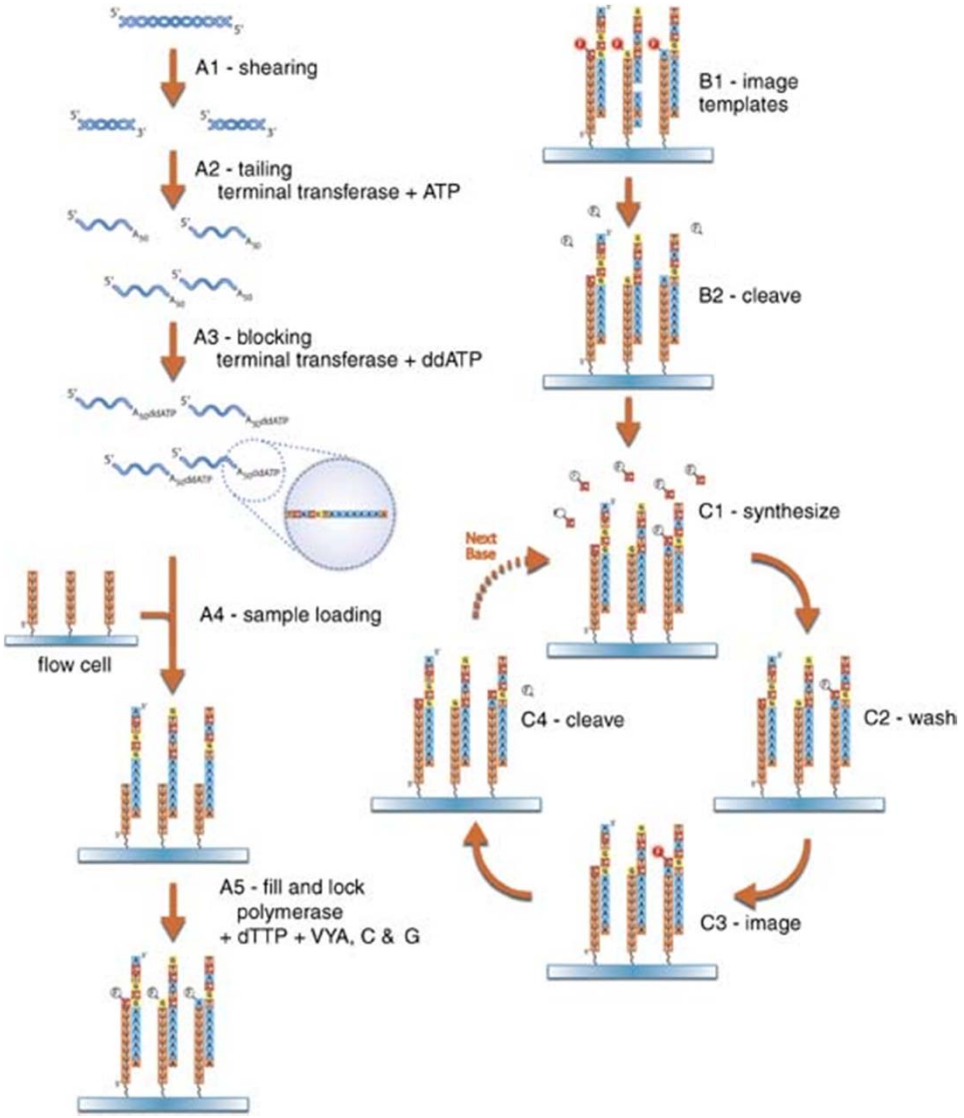
Izrablja transokacijo DNA čez nanopore in s tem povezane odčitke električnega signala. Nukleotid blokira ionski tok čez nanoporo – vsak nukleotid ima drugačno časovno periodo blokiranja!

.



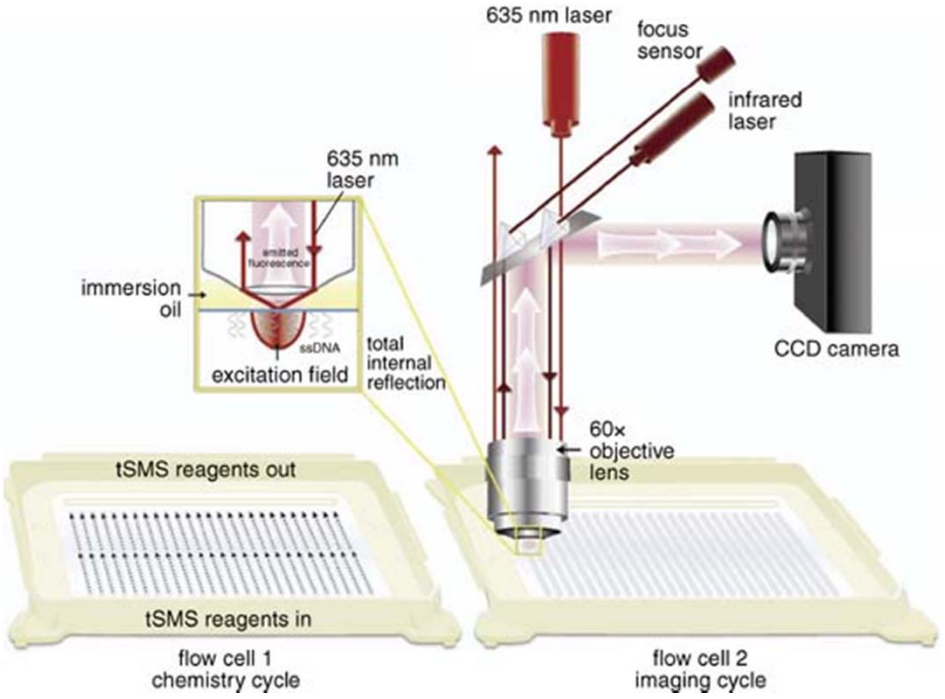
Nanopora: alfa-hemolizin, kovalentno vezan na molekulo ciklodekstrina

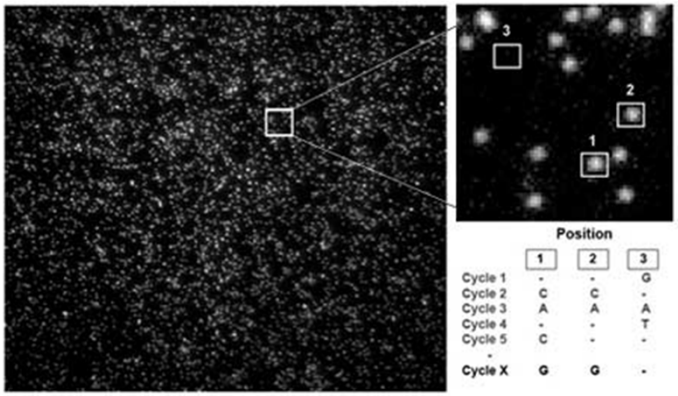
tSMS (Heliscope) – PRAVO SEKVENCIRANJE POSAMEZNIH MOLEKUL



- Priprava knjižnice z naključno razgradnjo DNA
- Dodajanje poli-A repov
- Hibridizacija z oligo-dT v pretočnih celicah
- Sekvenciranje z dodajanjem posameznih fluorescentnih nukleotidov
- Odčitavanje signala
- Odčitki dolgi 55 bp, 8 dni za 28 Gb v enem zagonu reakcije

tSMS (Heliscope) – ODČITAVANJE SIGNALA PRI SEKVENCIRANJU POSAMEZNIH MOLEKUL





Slika, ki jo podaja HeliScope molekulski sekvenator. V povečavi je prikazana molekula DNA, ki je v tem ciklu vključila nukleotid “G”.

http://www.helicosbio.com/Technology/TrueSingleMoleculeSequencing/tabid/64/Default.aspx

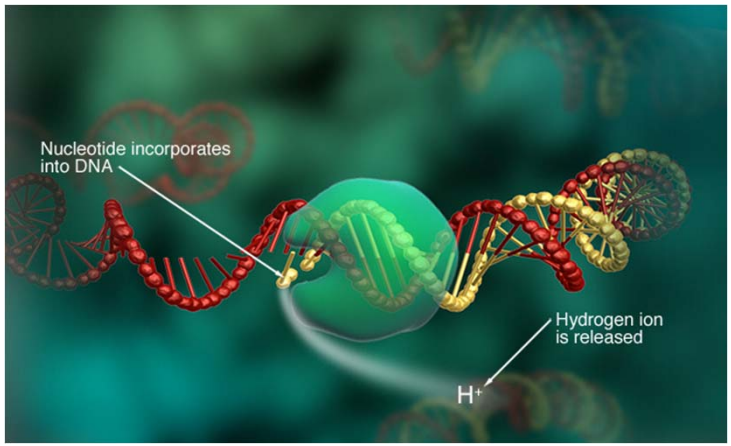
ION TORRENT TEHNOLOGIJA SEKVENCIRANJA

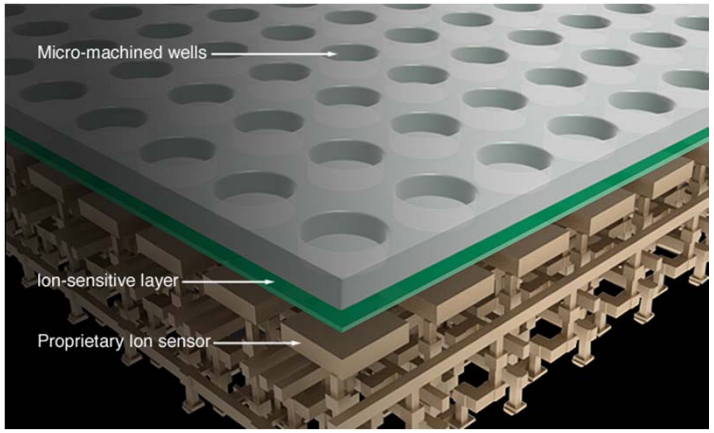
PostLight™ tehnologija sekvenciranjem na polprevodniških čipih.

Pri detekciji niso potrebni optični instrumenti.

Direktna povezava med kemično in digitalno informacijo.

MEHANIZEM VGRAJEVANJA NUKLEOTIDOV PRI SEKVENCIRARNJU Z ION TORRENTOM





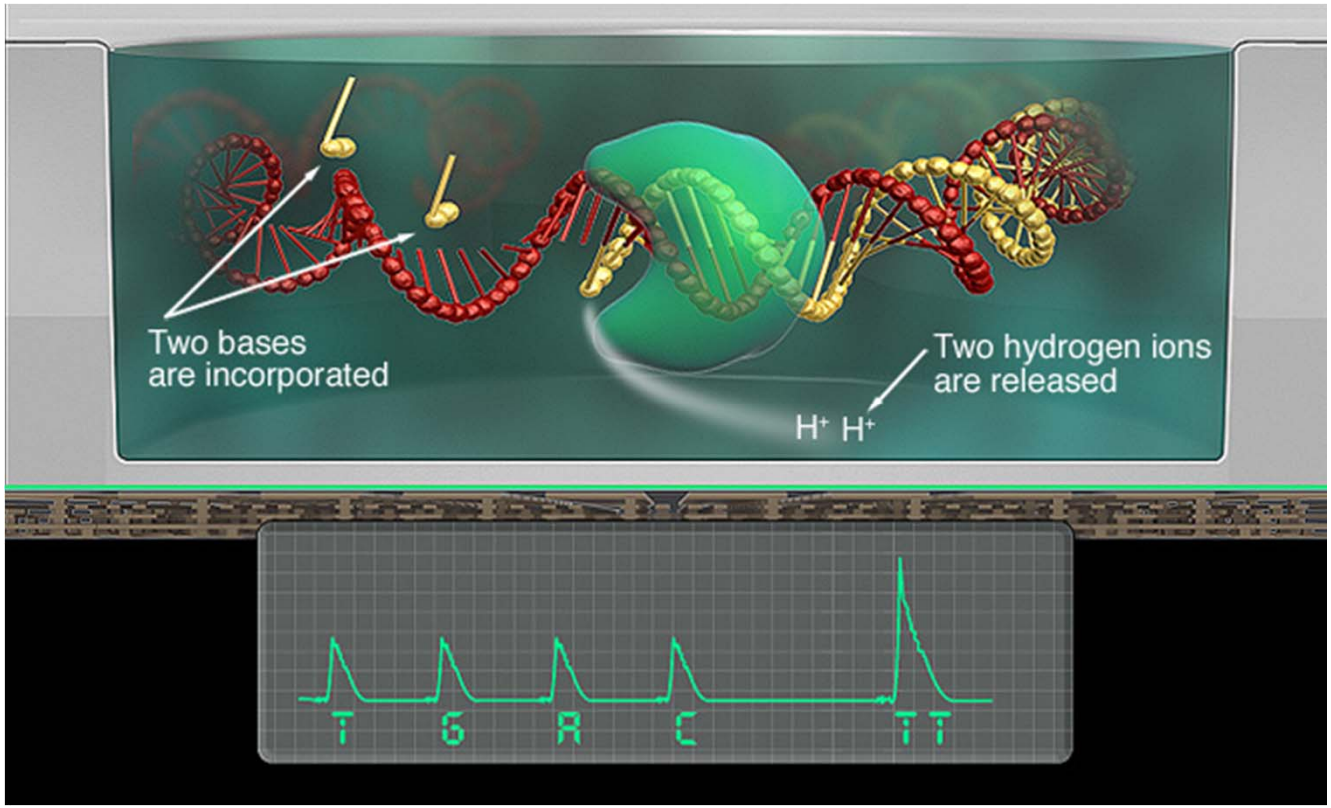
V naravi vgraditev vsakega nukleotida v verigo DNA vodi do sprostitve vodikovega protona (H+).

Ion Torrent vsebuje visoko-gostotno mrežo mikroaparatur v žepkih (luknjicah), kjer se proces vgrajevanja nukelotidov dogaja masivno in paralelno.

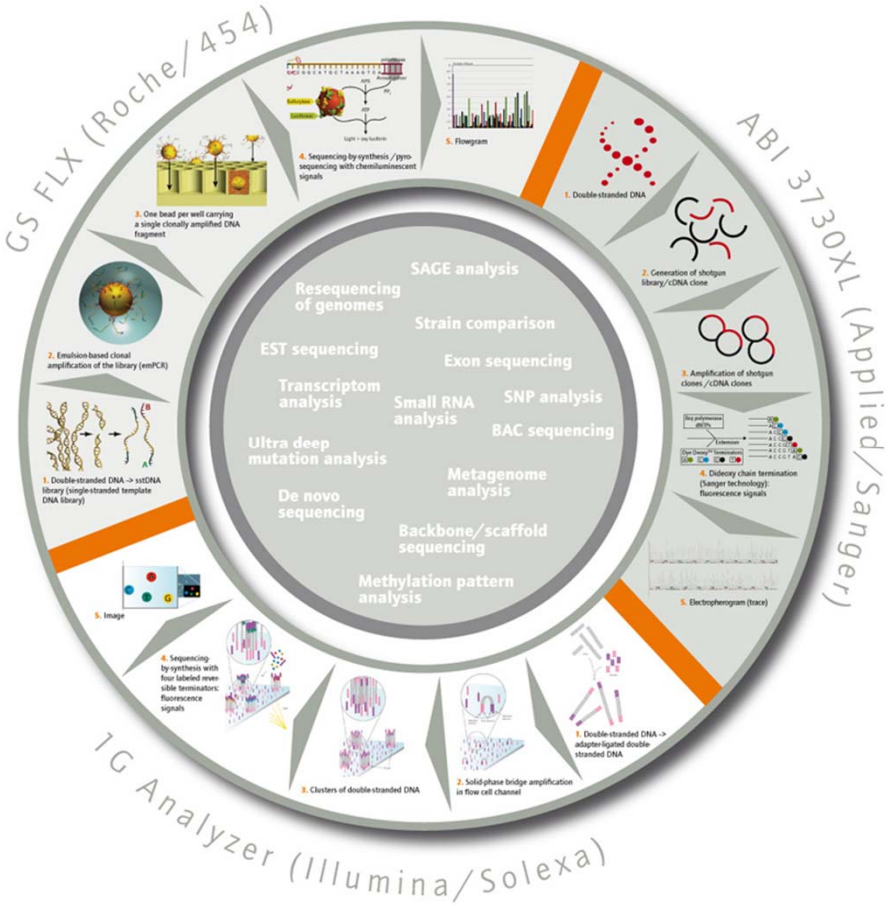
Vsak žepek vsebuje svojo molekulo DNA. Pod žepkom je za ione občutljiva plast (senzor), ki beleži pH spremembe zaradi sproščenih protonov.

Ker se nukleotidi dodajajo zaporedno, vemo, zaradi katerega nukleotida je prišlo do spremembe pH.

BELEŽENJE SIGNALA PRI VGRAJEVANJU NUKLEOTIDOV



UPORABA NASLEDNJE GENERACIJE SEKVENCIRANJA



HTS-NG od leta 2005 dalje revolucija v raziskavah genomov (tudi rastlinskih, živalskih, človeškega)

V letu 2003 je NHGRI (National Human Genome Research Institute napove 100-kratno znižanje cene na bp v 5 letih in v 10 leith 10 000-kratno znižanje cene na bp, kar bi privedlo do “1000 $ genoma”.

UPORABA RAZLIČNIH TEHNOLOGIJ 2. GENERACIJE SEKVENCIRANJA

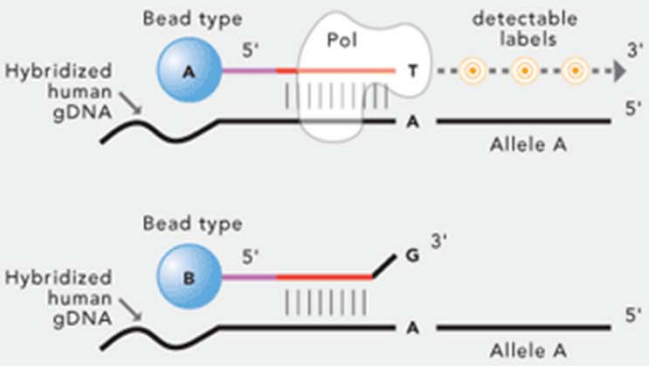
(c)

Sequencers	454 GS FLX	HiSeq 2000	SOLiDv4	3730xl
Resequencing		Yes	Yes	
De novo	Yes	Yes		Yes
Cancer	Yes	Yes	Yes	
Array	Yes	Yes	Yes	Yes
High GC sample	Yes	Yes	Yes	
Bacterial	Yes	Yes	Yes	
Large genome	Yes	Yes		
Mutation detection	Yes	Yes	Yes	Yes

Uporaba nove generacije sekvenciranja – določevanje SNP

Infinium I Assay Overview

Allele-Specific Primer Extension



Infinium II Assay Overview

Single Base Extension

The Infinium I and Infinium II Assay are two approaches used to interrogate SNPs within the genome

Infinium Assay Workflow

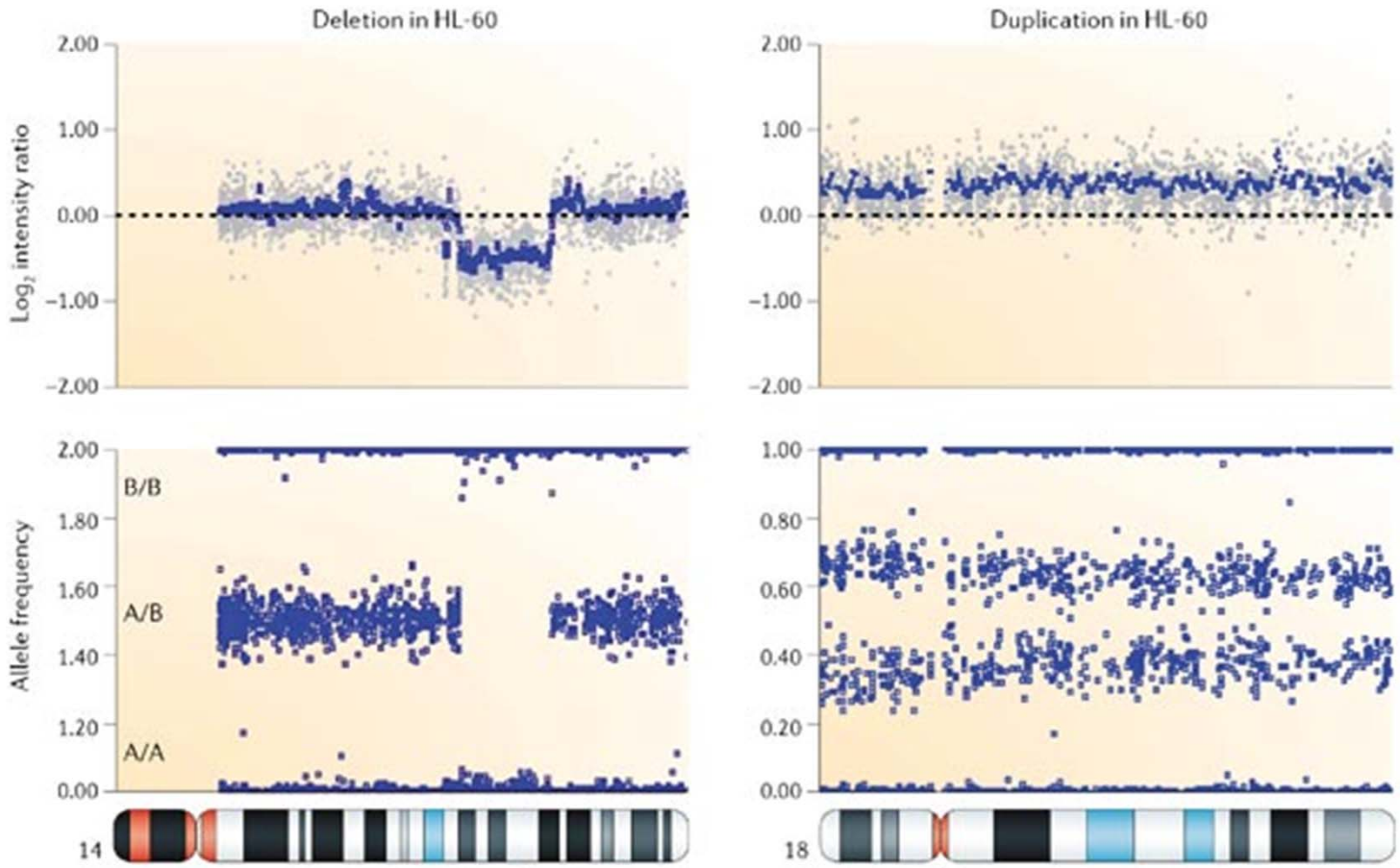
		Total time	(hands-on)
DAY 1			
STEP 1	Set up DNA amplification	20:20	(0:20)
	Total time (hands-on time):	20:20	(0:20)
DAY 2			
STEP 2	Fragment amplified DNA	1:10	(0:10)
STEP 3	Precipitate amplified DNA	2:30	(0:40)
STEP 4	Resuspend amplified DNA	1:10	(0:10)
STEP 5	Prepare BeadChip	0:30	(0:30)
STEP 6	Hybridize sample to BeadChip	17:00	(0:45)
	Total time (hands-on time):	22:20	(2:15)
DAY 3			
STEP 7	Extend and stain BeadChip	3:00	(0:35)
STEP 8	Scan BeadChip (1 scanner)	4:00	(0:20)
	Total time (hands-on time):	7:00	(0:55)
	Workflow total time:	49:40	(3:30)

Times are calculated for two technicians processing 8 HumanHap300 BeadChips.

STOP = optional stopping point

illumina

Uporaba nove generacije sekvenciranja – študije CNV (copy number variation)



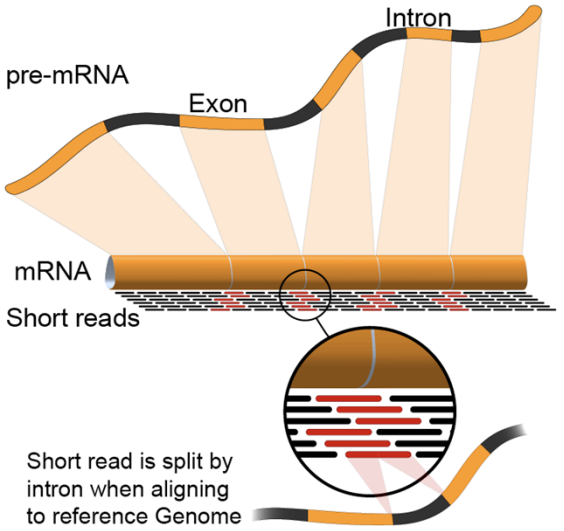
Fan *et al. Nature Reviews Genetics* **7**, 632–644 (August 2006) | doi:10.1038/nrg1901

nature REVIEWS GENETICS

POLEG DNA LAHKO Z DRUGO IN TRETJO GENERACIJO SEKVENCIRAMO TUDI RNA ALI DOLOČAMO NUKLEOTIDNA ZAPOREDJA, KI SO V STIKU S PROTEINI

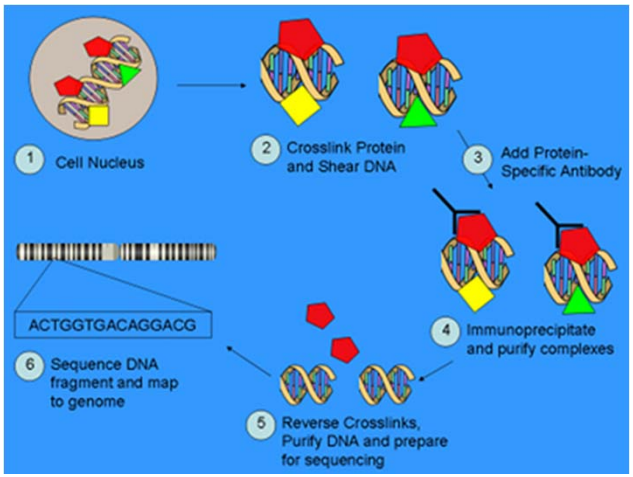
RNA-seq:

- sekvencriranje cDNA z NG-HTSs
- direktno sekvenciranje RNA s tretjo generacijo



Chip-seq:

- Kromatinska imunoprecipitacija proteinov, ki so vezani na DNA ali RNA in sekvenciranje predelov z 2. ali 3. generacijo.



http://seqanswers.com/forums/showthread.php?t=742

Povzetek

• Nova generacija visokozmogljivostnega sekvenciranje omogoča razpoznavanje zaporedij DNA na ravni celega genoma, z resolucijo posameznega baznega para.

• Iz vsakega vzorca se pripravi z adaptorji ligirana knjižnica, ki vsebuje vse v vzorcu prisotne fragmente DNA ali RNA (cDNA).

• Vse platforme bazirajo na ligaciji adaptorjev in pomnoževanju, imajo pa različne pristope sekvenciranja:

-Pirosekvenciranje (Roche-Nimblegen)

-Sekvenciranje s sintezo (Illumina-Solexa)

-Sekvenciranje z ligacijo (ABI)

• Tretja generacija sekvenciranja je še v razvoju. Temelji na metodah, ki pred sekvenciranjem ne potrebujejo pomnoževanja DNA in detekcijo ne potrebujejo ali optičnega signala ampak detektirajo električni signal.

ZMOŽNOST HITREGA DOLOČANJA NUKLEOTIDNEGA ZAPOREDJA POSAMEZNIKOVEGA GENOMA JE PRIVEDLO DO OBDOBJA INDIVIDUALNE (OSEBNE, POSAMEZNIKU PRILAGOJENE) GENOMIKE

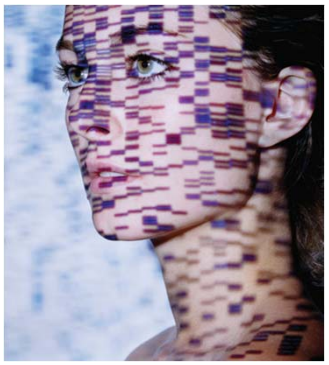
Vedno več podjetij ponuja sekveniranje predelov posameznikovega genoma za ceno nekaj 100 €.

Kako posameznik lahko uporabi informacijo o svojem genomu?

Kaj si s to informacijo na sedanji stopnji znanja lahko pomaga zdravnik?

“Manjkajoča dednost” predstavlja največjo oviro pri odkrivanju z boleznimi povezanih genov. Vzrok je prispevek velikega števila lokusov, ki imajo posamezno majhen učinek in lahko učinkujejo v različnih kombinacijah.

REALNOST INDIVIDUALIZIRANE (OSEBNE) MEDICINE 2010



Informacija o genomu naj bi v perspektivi imela poseben status v klinični medicini. Zato so se že zelo zgodaj razmahnile komercialne ponube genomskih testov, ki bazirajo na omejenih študijah klinične veljave, pogosto brez analize dobrobiti za zdravje.

V principu naj bi osebna genomika zmanjšala stroške zdravljenja zaradi boljše diagnostike, bolj efektivnega zdravljenja, vendar je zaenkrat zelo malo z dokazi podprtih tovrstnih dejstev.

Dobre znanstvene ideje same po sebi ne morejo spremeniti medicinske prakse. Zato mora osebna genomika prestati prav takšne teste in standarde, kot katerokoli drugo polje medicine.

ZADRŽKI PRED OSEBNO GENOMIKO?

Poleg realnih obetov za izboljšanje diagnoze in zdravljenja, osebna genomika posamezniku lahko tudi škodi, predvsem v etičnem in psihološkem smislu, v primeru lažnih pozitivnih/negativnih podatkov ali nezmožnosti razumevanja podatkov.

Genomika je velik izziv za medicino. Nikdar do sedaj še ni bilo tako velikega razkoraka med količino podatkov in našo zmožnostjo, da jih razumemo in koristno uporabimo. **Zato je nujno, da nadalje razvijamo orodja (tudi bioinformatična!) in čim hitreje izobrazimo nove generacije zdravnikov in raziskovalcev.** Tako bomo ohranili dobrobit genomike in minimizirali potencialne negativne učinke.

http://konradjkarczewski.com/2012/09/26/personal-genomics-in-wordle-form/