

テクニカルノート

半導体マイクロチップ対応次世代シーケンサーを用いた がん関連遺伝子の解析

狩生 徹, 井上雄太郎, 水上洋一

山口大学大学研究推進機構総合科学実験センター遺伝子実験施設 宇部市南小串1丁目1-1 (〒755-8505)

Key words : 次世代シーケンサー, がん遺伝子, 体細胞変異

和文抄録

半導体マイクロチップ技術を利用した次世代DNAシーケンサーは、比較的長いリード長や迅速な解析等の利点から急速に普及しつつある。また、がんなどの疾患を標的としたアンプリコンライブラリー作製技術も進展し、両者を組み合わせた大規模変異遺伝子解析がより身近なものになった。現在我々の研究室で行っている解析を例に、臨床サンプルを用いたがん関連遺伝子体細胞変異同定法について紹介する。

はじめに

近年の大規模遺伝子解析技術の進展は目覚ましく、ベンチトップの機器を用いたギガベース規模のゲノム解析も可能となり、小規模の研究グループでも利用可能になってきた。Life Technologies社製 Ion PGMシーケンサーは半導体マイクロチップを利用してDNA伸長時に放出されるプロトンを検出する世界初の次世代シーケンサーである(図1)。塩基の蛍光標識や光学的検出方法が不要で迅速な解析(1週間以内で解析終了)が可能である。現在1回のランで最大1GB(ギガ塩基)の配列しか読めないという問題はあるが、他の大規模次世代型シーケンサーが1リード(1回の解析の長さ)75 bpに対し、Ion PGMは200 bp以上読めるため、比較的

配列ギャップが生じにくい利点がある。このため、微生物のゲノムなど比較的ゲノムサイズが小さいサンプルの解析に主に用いられている。

このベンチトップ型次世代シーケンサーの特長を利用して指定した領域の配列だけをPCRで増幅し、短時間で大量のシーケンスを行う実験方法が行われ始めている。特にがんのサンプルでは変異が生じやすい染色体上のホットスポット領域が報告されており、この領域では多種類の変異が検出されている。このホットスポット領域を短時間で解析する必要があり、私たちは、次世代シーケンサーを用いてがんのホットスポット領域のアンプリコンシーケンス法を行っている。

この方法は、46種類のがん関連遺伝子(Cancer Panel)に存在する739箇所の変異箇所を1本のチューブでPCR増幅を行い、1回のランですべての領域を数百回シーケンス解析する方法である。この方法を行うことで臨床検体に腫瘍細胞と正常細胞が混在

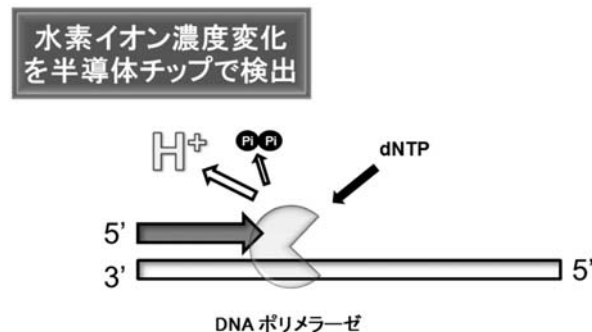


図1 DNA伸長反応と水素イオンの検出

するという複雑なサンプルでも、コホートの体細胞変異を正確に解析することが可能である。

サンプルライブラリーの作製

正常組織あるいはがん組織より市販のゲノムDNA抽出キットを用いて高分子DNAを抽出した。その後 Ion AmpliSeq Cancer Panel (Life Technologies社) 及び専用のアダプタープライマーキットのプロトコールに従い多検体処理用バーコード配列付きのライブラリーを作製した (図2)。Ion AmpliSeq Cancer Panelは190種類のプライマーミックスが含まれており、このプライマーミックスを用いてKRAS, BRAF, EGFRなど46種類のが

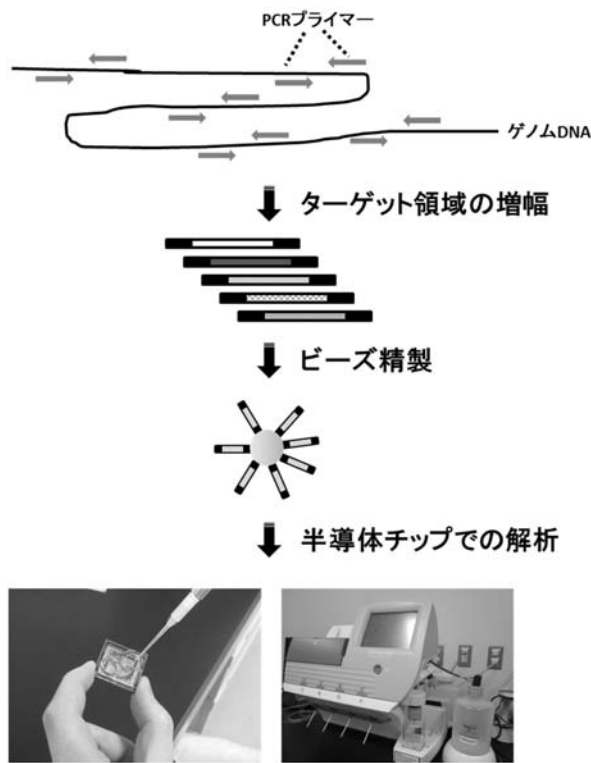


図2 Ion PGM解析の概略

Cancer Panelターゲット領域をPCRにより増幅後、DNA断片の両端にシーケンスを行うためのアダプターを結合させる。その後1ビーズに1DNA断片が取り込まれる条件でエマルジョンPCRを行い、同じ配列のDNA断片をビーズ上に増幅する。精製したビーズはピペットでマイクロチップに流し込み、半導体チップをIon PGMシーケンサーにより解析する。Ion PGMシーケンサーは、dNTPを順番にマイクロチップに送液し、DNAポリメラーゼによってDNAが伸長する際に放出される水素イオン濃度を半導体チップ上で検出する。Ion PGMシーケンサーのデータは、専用のTorrentサーバーに転送され、目的に応じた解析を行う。

ん関連遺伝子に存在する739箇所の変異を解析した (表1)。

シーケンス解析

機器に付属する Ion OneTouchシステムを用いてテンプレート調整を行った。これは1個のビーズに1DNA断片が取り込まれる条件でエマルジョンPCRを行い、同じ配列のDNA断片が増幅されたビーズを精製するものである。このサンプルを10⁶以上のリード仕様の Ion 316チップを用いて200 bpリードで解析を行った。Ion PGMシーケンサーは、各dNTP溶液の送液、DNAポリメラーゼによる伸長、半導体チップ上で放出された水素イオン濃度の検出を繰り返すことにより配列データを得るものである。データ解析は機器に付属するソフトウェア Torrent SuiteおよびCLCバイオ社のGenomics Workbenchを使用し、ヒトゲノムへのマッピング及び変異検出を行った。

解析結果

解析データ (表2) では、得られたDNA配列が60-140 bpの長さに分布しており、2.3メガリード、170 MbpがCancer Panelのアンプリコンに対するリファレンス配列にマッピングされた。私たちは1回

表1 Cancer Panelで解析可能な46種類のがん関連遺伝子

ABL1	ERBB4	KDR	PTPN11
AKT1	FBXW7	KIT	RB1
ALK	FGFR1	KRAS	RET
APC	FGFR2	MET	SMAD4
ATM	FGFR3	MLH1	SMARCB1
BRAF	FLT3	MPL	SMO
CDH1	GNAS	NOTCH1	SRC
CDKN2A	HNF1A	NPM1	STK11
CSF1R	HRAS	NRAS	TP53
CTNNA1	IDH1	PDGFRA	VHL
EGFR	JAK2	PIK3CA	
ERBB2	JAK3	PTEN	

表2 Cancer Panelを用いた解析

	Average depth	Coverage			マップされた全塩基数 (全塩基中のMapされた割合)	検出された新規バリエーション数
		Depth<1	Depth<20	Depth<100		
サンプルA	3529	100%	100%	99.1%	49.2Mbp (95.3%)	7
サンプルB	2802	100%	99.1%	98.3%	38.9Mbp (95.6%)	5
サンプルC	2754	100%	99.9%	97.0%	38.1Mbp (96.0%)	12
サンプルD	2634	100%	99.5%	96.4%	36.9Mbp (95.0%)	8

の実験で4サンプルにバーコードを付加し同時に反応を行った。その後、バーコード毎に振り分けて解析したが、各サンプル平均して40.8 Mbpがマップされた。それぞれのターゲットカバー率100%、平均のカバー数（領域を読んだ平均回数）は2930であり、このカバー数であれば癌組織に1%しか存在しないレアな変異も検出可能である。実際に体細胞変異の検出を行った結果、新規の変異が複数検出されており、これらの配列はいずれもキャピラリーシーケンサーで確認することができた。

このように次世代アンプリコンシーケンス法では理論上、数%しか存在しないがん細胞の変異も検出可能であり、今後解析例が増えていくと予想される。

Cancer Panel解析プロトコール

- ① 凍結組織からDNA抽出
- ② Cancer Panelプライマープールを用いて、抽出DNAをテンプレートにしてターゲット領域を増幅（15から19サイクル）
- ③ プライマー配列の酵素切断とバーコードアダプター付加
- ④ AMPure XP（常磁性微小ビーズ、ベックマンコーラター）とマグネティックラックを用いて、反応液からアンプリコンDNAを回収
- ⑤ Bioanalyzer（アジレント）を用いて増幅DNAが130から210 bpに分布していることを確認
- ⑥ 20 pMに希釈したライブラリー5 μ LをIon PGMに付属するIon OneTouch Duoシステムを用いてエマルジョンPCR（1ビーズに1DNA断片が取り込まれ、オイル中のエマルジョン内で増幅）とビーズ精製
- ⑦ Ion PGMシーケンサーと半導体チップによるシーケンス解析と解析ソフトによるデータ処理

解析の拡張性

半導体マイクロチップ対応シーケンサーのアプリケーション拡張は目覚ましく、様々な種類のアンプリコンシーケンスプライマーが販売され始めている。Ion AmpliSeq Comprehensive Cancer Panelでは409種のがん遺伝子およびがん抑制遺伝子内のエクソン領域をターゲットとしている。またIon

AmpliSeq Cancer Hotspot Panel v2はCOSMIC（Catalogue of Somatic Mutations in Cancer）に登録されているホットスポット領域の約2800の変異をカバーしており、幅広い変異プロファイリングに対応している。また、Ion AmpliSeq Inherited Disease Panelも既に販売されており、このプライマープールを用いることにより神経筋疾患、心疾患、発育障害、代謝異常など、700種以上の遺伝性疾患に関連する328種の遺伝子のエクソンの解析が可能である。さらに個別の研究用アンプリコンシーケシングのためのプライマーデザインも可能であり、1536までのアンプリコンを簡便かつ大規模に解析できる。プライマーデザインは、ライフテクノロジー社のホームページから遺伝子を指定することで容易に行える。但し、1本のチューブに入れるプライマーの数が増加すると増幅ミスも増える傾向にあり1000種以内に抑えた方が効率良く解析することができた。

具体的な解析例については、小児多形黄色星細胞腫のホルマリン固定サンプルより抽出したDNAよりライブラリーを作製し、MET遺伝子の点変異を同定した研究¹⁾や、胃がん組織における体細胞変異の大規模解析²⁾、大腸がん患者における生殖細胞変異を同定した報告³⁾を参照されたい。

次世代シーケンサーによるアンプリコンシーケンスの問題

解析においてサンプル調製はかなり重要である。損傷の激しいサンプルでは、必ずしも正しいデータを反映しているとは限らない。特にホルマリン固定サンプルについては、固定時に変異が入る可能性もある。また、PCR増幅中に生じる変異も無視できないため、できるだけDNA量を確保して増幅回数を減らす工夫が必要である。

次世代シーケンサーの信頼性についてはまだ十分な確認がとれているとは言えないため、得られたデータは必ず何らかの方法で精査する必要がある。キャピラリーシーケンサーは、20%以下の変異を検出できないため、次世代シーケンサーのデータと必ずしも一致するとは限らない。このためPCRなどで次世代シーケンサーのデータを検証する方法が研究されている。

また、Ion PGMに付属している解析ソフトは操作が簡便であるが、デフォルト設定では、シーケンスミスも変異として検出する可能性がある。配列を目視で確認したり、別の解析ソフトで解析し直すなどの操作が必要である。

引用文献

- 1) Yang MM, Singhal A, Rassekh SR, Yip S, Eydoux P, Dunham C. Possible differentiation of cerebral glioblastoma into pleomorphic xanthoastrocytoma : an unusual case in an infant. *J Neurosurg Pediatr* 2012 ; 9 : 517-523.
- 2) Kim JG, Takeshima H, Niwa T, Rehnberg E, Shigematsu Y, Yoda Y, Yamashita S, Kushima R, Maekita T, Ichinose M, Katai H, Park WS, Hong YS, Park CH, Ushijima T. Comprehensive DNA methylation and extensive mutation analyses reveal an association between the CpG island methylator phenotype and oncogenic mutations in gastric cancers. *Cancer Lett* 2013 ; 330 : 33-40.
- 3) Palles C, et al. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat Genet* 2012 ; 45 : 136-144.