

Validazione della diagnosi molecolare di rene policistico autosomico dominante mediante la tecnologia *Next Generation Sequencing*

Francesca Ferrari¹, Silvana Tedeschi¹, Roberta Cerutti², Piergiorgio Messa², Manuela Seia¹

¹UOS Laboratorio di Genetica Medica, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milano

²UOC Nefrologia e Dialisi, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milano

Validation of the Autosomal Dominant Polycystic Kidney Disease molecular diagnosis by Next Generation Sequencing technology

Autosomal Dominant Polycystic Kidney Disease (ADPKD) is the most common genetic disorder of the kidney and accounts for over 95% of all cystic diseases involving loss of kidney function. The diagnosis of ADPKD is confirmed by mutational analysis of PKD1 and PKD2 genes, which is complex, because of the significant allelic heterogeneity and the presence of six pseudogenes presenting highly homologous sequences with an extended region of the PKD1 gene (exons 1-33).

The aim of this study was to develop a rapid and reliable method for the molecular diagnosis of ADPKD using the technology of Next Generation Sequencing (NGS), that could be used for a diagnostic purpose. We first validated the test on two cohorts of ADPKD patients previously characterized by traditional Sanger sequencing method; then we analyzed a cohort of 58 patients not characterized and we achieved a mutation detection rate of 84.7%.

Keywords: Autosomal Dominant Polycystic Kidney Disease, Next Generation Sequencing, PKD1 gene, PKD2 gene



Manuela Seia

Introduzione

Il Rene Policistico Autosomico Dominante (ADPKD) è la più frequente malattia genetica del rene, con un'incidenza di 1 su 400-1000 nati (1, 2); è la principale causa genetica di insufficienza renale dell'adulto (rappresenta più del 95% di tutte le malattie cistiche che comportano perdita di funzionalità renale) ed è responsabile di più del 10% dei casi di Insufficienza Renale Cronica Terminale (*End Renal Stage Disease*, ESRD) (3). È caratterizzato dal progressivo sviluppo in entrambi i reni di cisti che portano a una riduzione della capacità di filtrazione renale con un decorso progressivo e irreversibile e da varie manifestazioni extrarenali, principalmente anomalie gastrointestinali e cardio-vascolari (cisti epatiche e pancreatiche, diverticolosi

intestinale, ernie addominali, aneurismi cerebrali, anomalie delle valvole cardiache, ipertrofia ventricolare sinistra) (4, 5). Le cisti si possono sviluppare in tutti i segmenti del nefrone e, sebbene solo una piccola percentuale (1-10%) dei nefroni sviluppi le cisti, l'intero organo subisce alterazioni strutturali e funzionali (2, 6).

La malattia policistica renale è trasmessa come carattere autosomico dominante ed è determinata da mutazioni nei geni PKD1 (85% dei casi) e PKD2 (15% dei casi) (7-9), che codificano per le proteine Policistina 1 e Policistina 2; queste due proteine si co-localizzano nelle cellule epiteliali dei tubuli renali. Vengono identificate mutazioni causative nell'84-91% dei pazienti con diagnosi ecografica di ADPKD (2, 10, 11).

È una malattia a penetranza completa ma con espressività variabile; nella maggioranza dei casi, la malattia si manifesta in età adulta, ma con il miglioramento delle tecniche di *imaging e*, soprattutto, con l'introduzione nella diagnostica delle tecniche di biologia molecolare è aumentato il numero di casi pediatrici di ADPKD identificati: pazienti che spesso all'esordio hanno sintomi in parte sovrapponibili ad altre patologie cistiche renali e casi con esordio grave in epoca prenatale e neonatale. L'espressività variabile è determinata dall'eterogeneità del *locus* (esiste una correlazione genotipo-fenotipo legata ai due *loci*) e dall'eterogeneità fenotipica sia inter- che intra-familiare, causata dalla combinazione di diversi fattori, quali: eterogeneità allelica (la stessa mutazione non è identificata in più del 2% delle famiglie), alleli ipomorfici (alleli che producono una ridotta quantità di proteina o proteine con ridotta attività funzionale), mosaicismi,

Accepted: July 6, 2016

Published online: July 29, 2016

Indirizzo per la corrispondenza:

Dr.ssa Manuela Seia
Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico
UOS Laboratorio di Genetica Medica
Via F. Sforza 35
20122 Milano
m.seia@policlinico.mi.it

eredità digenica e presenza di geni o alleli modificatori, la cui esistenza è supportata dalla variabilità fenotipica riscontrata tra gemelli monozigoti e tra fratelli.

La diagnosi di ADPKD è stata per lungo tempo principalmente strumentale, mediante ecografia o TAC e Risonanza Magnetica, che sono esami più sensibili e permettono di rilevare cisti fino a 2 mm di diametro rispetto ai 10 mm dell'ecografia. Tuttavia, nell'ultimo decennio, la diffusione delle tecnologie di sequenziamento genico ha permesso di supportare il dato ecografico e di diagnosticare con certezza casi nei quali l'*imaging* non era conclusivo.

Data la complessità dell'analisi, le indicazioni allo *screening* molecolare dei geni PKD1 e PKD2 sono limitate a (2, 12):

- parenti, potenziali donatori, di pazienti in attesa di trapianto che presentano dati ecografici equivoci o con un numero di cisti, valutato in rapporto all'età, inferiore al numero di cisti stabilito come limite minimo per la diagnosi certa di ADPKD o parenti giovani (≤ 25 anni), in cui l'ecografia non ha individuato cisti;
- pazienti con storia familiare negativa per ADPKD, con un dato ecografico atipico (cisti presenti in numero maggiore in un rene rispetto all'altro, presenza di un numero elevato di cisti molto piccole) o con segni extra-renali atipici o con una forma "*mild*" di malattia di cui vogliono conoscere la prognosi;
- pazienti o famiglie con esordio precoce di malattia cistica renale per identificare mutazioni che possano essere associate a una forma molto grave e precoce di ADPKD o per chiarire l'effetto di più mutazioni/varianti geniche in un paziente;
- pazienti che richiedono una diagnosi certa e definitiva sia a titolo prognostico sia per una pianificazione familiare consapevole.

Il progresso metodologico e la riduzione dei costi del sequenziamento *Sanger* a partire dalla metà degli anni 2000 hanno permesso di mettere a punto la diagnostica molecolare dei geni dell'ADPKD, che, comunque, risulta complessa, poiché non esistono *hot-spot* di mutazione, il gene PKD1 è estremamente ricco di basi GC, fattore che rende difficile ottenere efficienza nelle reazioni di PCR e sequenza, e la regione compresa tra gli esoni 1 e 33 di PKD1 (circa il 75% del gene) ha subito, durante l'evoluzione, eventi di ricombinazione intra-cromosomica che hanno prodotto 6 pseudogeni con una similarità di sequenza di circa il 98% (13-15). Un ulteriore progresso nella diagnostica genetica si è avuto in questi ultimi anni con l'introduzione della tecnologia di "*Next Generation Sequencing*", che consente di generare, mediante sequenziamento parallelo, un rilevante set di dati genetici con una tempistica significativamente ridotta e con costi ancora più contenuti. Il nostro laboratorio ha validato un pannello "*custom*" per l'amplificazione in parallelo di tutte le regioni codificanti e delle regioni introniche potenzialmente coinvolte nello *splicing* dei geni PKD1 e PKD2.

Materiali e metodi

Prelievi di sangue sono stati effettuati a pazienti adulti con diagnosi ecografica di Rene Policistico Autosomico Dominante e a pazienti pediatrici con sospetto di malattia cistica renale, dopo avere raccolto il consenso informato per l'analisi genetica e il trattamento dei dati genetici.

L'estrazione del DNA è stata eseguita con l'estrattore automatico *QiaSymphony^{SP}* (Qiagen) utilizzando 400 μ l di sangue in EDTA.

L'esperimento NGS di "*targeted sequencing*" è stato eseguito utilizzando un pannello *custom* di sonde per amplificare tutte le regioni *target* dei geni PKD1 e PKD2 (61 esoni comprensivi di 20 basi introniche ai confini esone-introne).

Nella prima fase, le regioni *target* sono amplificate mediante sette PCR *Multiplex*, le PCR *Multiplex* di ogni campione sono, poi, unite con specifici e calcolati rapporti per produrre una libreria di 63 ampliconi di adeguata concentrazione; la libreria è purificata con *AMPure beads* (Beckman Coulter, Nyon, Switzerland) e diluita. Viene, poi, eseguita la PCR "*Universal*" che consente di etichettare gli ampliconi con specifici "*molecular identifiers*" (MID) per distinguere tra loro i pazienti, e adattatori, necessari per il sequenziamento sullo strumento *MiSeq* dell'*Illumina*. Le librerie di ampliconi marcate vengono nuovamente purificate con *AMPure beads*, quantificate con il fluorimetro *Qubit[®] dsDNA BR Assay Kit* (Life Technologies, New York, USA) e viene fatto un *pool* equimolare (10 pM) delle librerie con il 5% di *PhiX* (*Illumina, San Diego, CA, USA*), usato come campione di controllo per la verifica interna del sequenziamento; infine, il *pool* è caricato sullo strumento *MiSeq* utilizzando o il *2 × 250 MiSeq Reagent Kit v2* (*Illumina, San Diego, CA, USA*) o il *2 × 250 MiSeq Reagent Kit v2 Nano*, in base al numero di campioni che costituiscono il *pool* equimolare.

Per superare il problema dell'omologia di sequenza tra PKD1 e i 6 pseudogeni, è stata identificata, per ogni amplicone, una serie di varianti proprie degli pseudogeni, "*anchor SNPs*", di modo che ogni *read* che contiene almeno due di questi "*anchor SNPs*" venga considerata derivata dagli pseudogeni e, quindi, eliminata. I risultati ottenuti dall'analisi bioinformatica, ossia le mutazioni e/o le varianti di incerto significato patologico identificate, vengono confermati mediante sequenziamento *Sanger* sul sequenziatore *ABI PRISM 3130xl* (*Applied Biosystems*).

Risultati

La validazione del pannello *custom* per i geni PKD1 e PKD2 è stata effettuata mediante due serie di reazioni successivamente caricate sul sequenziatore *MiSeq* (*Illumina*): la prima di 5 pazienti, precedentemente sequenziati con il metodo *Sanger*, con la cartuccia *2 × 250 MiSeq Reagent Kit v2 Nano* che ha un *output* massimo di 500 Megabasi (1 milione di *read*) e abbiamo potuto constatare che, dopo aver eliminato le letture degli pseudogeni, per avere una profondità di lettura

adeguata a identificare con sicurezza le varianti/mutazioni, si poteva caricare un massimo di 4 campioni. Il secondo esperimento è stato effettuato con la cartuccia 2 × 250 *MiSeq Reagent Kit v2*, che ha un *output* massimo di 15 Gigabasi (25 milioni di *read*) su cui sono stati caricati 28 campioni con varianti/mutazioni già note e abbiamo verificato che lo spazio permetteva di caricare un massimo di 70 campioni. Tutte le varianti identificate con il metodo Sanger nei 33 pazienti dello studio retrospettivo sono state rilevate anche dalla tecnica NGS.

Un requisito importante affinché i dati siano affidabili è la copertura delle regioni *target*, ossia il numero di volte che una base viene letta; la copertura di un amplicone è definita come la sua copertura media di base. Nonostante la copertura media degli ampliconi sia buona, dall'analisi dei file *FastQ* su *IGV (Integrative Genomics Viewer: <https://www.broadinstitute.org/igv/>)* abbiamo evidenziato sei regioni di poche basi, tutte in siti di confine tra esone e introne, quattro in PKD1 e due in PKD2, che non avevano una copertura sufficiente. Quindi, per il momento, per quei pazienti in cui non viene identificata una mutazione certamente patologica, sequenziamo queste regioni con il metodo *Sanger*.

Per il primo esperimento NGS con campioni non caratterizzati, abbiamo selezionato 46 pazienti adulti con diagnosi ecografica di ADPKD e 12 pazienti di età pediatrica con sospetto di malattia cistica renale; abbiamo utilizzato la cartuccia 2 × 250 *MiSeq Reagent Kit v2*. La corsa ha prodotto una densità media di *read* di 1020 K/mm², una percentuale di *read* identificate di 86.72% e % Q30 media (Q30: la probabilità che una base sia chiamata correttamente è del 99.9%) del 75.85%.

Dei 46 pazienti adulti, abbiamo individuato una mutazione patogenetica o probabilmente patogenetica in 39 casi (84.7%), mentre, in 2 casi (4.34%), è stata riscontrata una variante missenso non riportata in letteratura e nel *database* ADPKD (<http://pkdb.mayo.edu/>) che abbiamo definito come *Unclassified Sequence Variants (UCV)* (18); siamo in attesa dei campioni della famiglia per procedere a uno studio familiare della segregazione allelica di tali varianti. Cinque casi (10.86%) sono risultati negativi allo screening mutazionale. Le nostre percentuali sono coerenti con quelle riportate in letteratura (13, 16-18). Dei 12 pazienti di età pediatrica, 4 (33.3%) sono risultati portatori di una mutazione sicuramente patogenetica; in 2 (16.6%), abbiamo individuato una variante di incerto significato patogenetico, mentre 6 pazienti (50%) sono risultati negativi allo *screening*. Il 54.3% delle mutazioni patogenetiche identificate è rappresentato da nuove mutazioni, non riportate in letteratura.

Per tutti i pazienti, con o senza mutazione individuata dall'analisi NGS, è stata eseguita la tecnica MLPA per rilevare ampie delezioni o duplicazioni. Abbiamo identificato una delezione in eterozigosi dell'esone 15 del gene PKD1 in un paziente adulto negativo all'analisi NGS; è in corso uno studio per la caratterizzazione del *breakpoint* necessaria per definire l'estensione della delezione.

Discussione

Nei pazienti ADPKD, la diagnosi genetica non è eseguita di *routine*, in quanto è difficile sequenziare con il metodo *Sanger* i geni PKD, in particolare PKD1, responsabile della patologia nell'85% dei casi (8, 18). In questo studio, abbiamo dimostrato che la *Next Generation Sequencing* è una tecnica affidabile, riproducibile e meno laboriosa con cui superare la sfida dell'omologia degli pseudogeni. Utilizzando questa tecnica, siamo riusciti a dare una diagnosi molecolare a circa l'85% (84.7%) dei pazienti.

Il sospetto dell'esistenza di un terzo *locus* genico PKD3 (19), che sembrava giustificare il 10-15% dei pazienti con caratterizzazione negativa per PKD1 e PKD2, è stato smentito nel 2014 (20) da una più rigorosa analisi familiare; tuttavia, nel Giugno del 2016, grazie a uno studio multicentrico di 3600 famiglie affette, il sequenziamento dell'intero esoma di pazienti negativi all'analisi mutazionale dei geni PKD1 e PKD2 ha permesso di identificare un nuovo gene, GANAB, che codifica per la subunità catalitica della Glucosidasi II (GIIa), una proteina necessaria per la maturazione e la localizzazione delle proteine Policistina 1 e 2, che è risultata responsabile di circa il 3% dei casi negativi per mutazioni in PKD1 e PKD2 (0.3% dei casi totali) (21).

L'applicabilità clinica della *Next Generation Sequencing* è motivata dalla diminuzione dei costi, dalla riduzione dei tempi di refertazione e dallo sviluppo di flussi di lavoro più efficienti per la preparazione di *pool* di campioni, per il sequenziamento parallelo e per l'analisi dei dati (22, 23). L'affidabilità dell'analisi, invece, è legata principalmente al livello di copertura, ossia al numero di campioni che si caricano, alla dimensione del pannello di geni e all'*output* della cartuccia che si utilizza; infatti, si può stabilire di utilizzare un livello di *capture* così elevato da rendere l'analisi NGS molto affidabile. Nel nostro esperimento, abbiamo un numero di *read* per base che varia tra 100 e 800, valore che è sicuramente affidabile per un referto diagnostico.

Ampie delezioni e duplicazioni in eterozigosi, che nel ADPKD costituiscono circa il 4% delle mutazioni, sono difficilmente identificabili mediante NGS. Ogni tentativo di analisi comparativa mediante *software* ha mostrato alte percentuali di insuccesso, pertanto, per lo *screening* mutazionale dei pazienti con diagnosi ecografica o sospetto di ADPKD, è sempre necessario integrare l'analisi di sequenza con la tecnica MLPA, per la ricerca dei riarrangiamenti genici (delezioni/duplicazioni), poiché è alquanto frequente che i pazienti ADPKD siano portatori di più mutazioni.

Per i 2 pazienti adulti in cui abbiamo trovato una variante di incerto significato, non riportata nei *database* e in letteratura (*Unclassified Sequence Variants, UCV*), sarà necessario un accurato studio familiare; cinque pazienti non risultano portatori di alcuna mutazione/variante. Per questi ultimi casi, possono esserci diverse spiegazioni: è possibile che la mutazione possa essere localizzata in regioni introniche profonde, nel promotore o in 5' e 3' UTR, regioni che non sono analizzate con le metodiche molecolari attualmente in uso. Bisogna, inoltre, considerare il mosaicism, evento

che può avvenire in uno stadio precoce della vita embrionale e che è stato descritto in due famiglie, ma che è probabilmente più rilevante, soprattutto nei casi *de novo* (24, 25), ossia mutazioni non identificate in alcuno dei genitori, e il ruolo degli alleli ipomorfici (26), fattori che possono spiegare anche la variabilità fenotipica intra-familiare spesso riscontrata e che possono essere molto importanti nella prognosi del paziente.

Questo studio rappresenta un significativo passo in avanti per la diagnostica molecolare del Rene Policistico Autosomico Dominante nei pazienti italiani, perché abbiamo individuato molte mutazioni, anche non riportate precedentemente, e abbiamo validato un metodo NGS con *detection rate* superiore al sequenziamento *Sanger*, riducendo in modo significativo i costi e i tempi tradizionalmente necessari per la diagnosi dovuti alla complessità dei geni coinvolti.

La possibilità di predire l'andamento della malattia e la progressione dell'insufficienza renale in base alla tipologia di mutazione identificata (27, 28), così come, in un futuro prossimo, di classificare i pazienti in categorie di suscettibilità o "non-responder" a terapie mirate renderà l'analisi molecolare sempre più richiesta.

Ringraziamenti

Si ringrazia l'Associazione Italiana Rene Policistico (AIRP) per il sostegno allo svolgimento di questo progetto.

Disclosures

Financial support: No financial support was received for this submission.
Conflict of interest: The authors have no conflict of interest.

Bibliografia

1. Tan YC, Blumenfeld J, Rennert H. Autosomal dominant polycystic kidney disease: genetics, mutations and microRNAs. *Biochim Biophys Acta*. 2011;1812(10):1202-12.
2. Harris PC. What is the role of somatic mutation in Autosomal Polycystic Kidney Disease? *JASN*. 2010;21:1073-6.
3. Spithoven EM, Kramer A, Meijer E, et al. Renal replacement therapy for autosomal dominant polycystic kidney disease (ADPKD) in Europe: prevalence and survival—an analysis of data from the ERA-EDTA Registry. *Nephrol Dial Transplant*. 2014;29 (Suppl. 4):iv15-25.
4. Perrone RD. Extrarenal manifestations of ADPKD. *Kidney Int*. 1997;51:2022-36.
5. Grantham JJ, Torres VE, Chapman AB, et al. Volume progression in polycystic kidney disease. *N Engl J Med*. 2006;354(20):2122-30.
6. Pei Y. A "two-hit" model of cystogenesis in autosomal dominant polycystic kidney disease? *Trends Mol Med*. 2001;7(4):151-6.
7. Veldhuisen B, Saris JJ, de Haij S, et al. A spectrum of mutations in the second gene for autosomal dominant polycystic kidney disease (PKD2). *Am J Hum Genet*. 1997;61(3):547-55.
8. Rossetti S, Strmecki L, Gamble V, et al. Mutation analysis of the entire PKD1 gene: genetic and diagnostic implications. *Am J Hum Genet*. 2001;68(1):46-63.
9. Rossetti S, Chauveau D, Walker D, et al. A complete mutation screen of the ADPKD genes by DHPLC. *Kidney Int*. 2002;61(5): 1588-99.
10. Torra Balcells R, Ars Criach E. Molecular diagnosis of autosomal dominant polycystic kidney disease. *Nefrologia*. 2011;31(1):35-43.
11. Cornec-Le Gall E, Audrézet MP, Le Meur Y, Chen JM, Férec C. Genetics and pathogenesis of autosomal dominant polycystic kidney disease: 20 years on. *Hum Mutat*. 2014;35(12): 1393-406.
12. Huang E, Samaniego-Picota M, McCune T, et al. DNA testing for live kidney donors at risk for autosomal dominant polycystic kidney disease. *Transplantation*. 2009;87(1):133.
13. Audrézet MP, Cornec-Le Gall E, Chen JM, et al. Autosomal dominant polycystic kidney disease: Comprehensive mutation analysis of PKD1 and PKD2 in 700 unrelated patients. *Human Mutat*. 2012;33(8):1239-50.
14. Bogdanova N, Markoff A, Gerke V, McCluskey M, Horst J, Dworniczak B. Homologues to the first gene for autosomal dominant polycystic kidney disease are pseudogenes. *Genomics*. 2001;74(3):333-41.
15. Eisenberger T, Decker C, Hiersche M, et al. An efficient and comprehensive strategy for genetic diagnostics of polycystic kidney disease. *PLoS One*. 2015;10 (2):e0116680.
16. Trujillano D, Bullich G, Ossowski S, et al. Diagnosis of autosomal dominant polycystic kidney disease using efficient PKD1 and PKD2 targeted next-generation sequencing. *Mol Genet Genomic Med*. 2014;2(5):412-21.
17. Rossetti S, Consugar MB, Chapman AB, et al. Comprehensive molecular diagnostics in autosomal dominant polycystic kidney disease. *J Am Soc Nephrol*. 2007;18(7):2143-60.
18. Rossetti S, Hopp K, Sikkink RA, et al. Identification of gene mutations in autosomal dominant polycystic kidney disease through targeted resequencing. *J Am Soc Nephrol*. 2012;23:915-33.
19. Ariza M, Alvarez V, Marin R, et al. A family with a milder form of adult dominant polycystic kidney disease not linked to the PKD1 (16p) or PKD2 (4q) genes. *J Med Genet*. 1997;34:587-9.
20. Paul BM, Consugar MB, Ryan Lee M, et al. Evidence of a third ADPKD locus is not supported by re-analysis of designated PKD3 families. *Kidney Int*. 2014;85:383-92.
21. Porath B, Gainullin VG, Cornec-Le Gall E, et al. Mutations in GANAB, Encoding the Glucosidase II α Subunit, Cause Autosomal-Dominant Polycystic Kidney and Liver Disease. *Am J Hum Genet*. 2016;98(6):1193-207.
22. Tan AY, Michael A, Liu G, et al. Molecular Diagnosis of autosomal dominant polycystic kidney disease using next-generation sequencing. *J Mol Diagn*. 2014;16:216-8.
23. Mallawaarachchi AC, Hort Y, Cowley MJ, et al. Whole-genome sequencing overcomes pseudogene homology to diagnose autosomal dominant polycystic kidney disease. *Eur J Hum Genet*. 2016 May 11 (doi: 10.1038/ejhg.2016.48) (in press).
24. Consugar MB, Wong WC, Lundquist PA, et al. Characterization of large rearrangements in autosomal dominant polycystic kidney disease and the PKD1/TSC2 contiguous gene syndrome. *Kidney Int*. 2008;74(11):1468-79.
25. Connor A, Lunt PW, Dolling C, et al. Mosaicism in autosomal dominant polycystic kidney disease revealed by genetic testing to enable living related renal transplantation. *Am J Transplant*. 2008;8(1):232-7.
26. Rossetti S, Kubly VJ, Consugar MB, et al. Incompletely penetrant PKD1 alleles suggest a role for gene dosage in cyst initiation in polycystic kidney disease. *Kidney Int*. 2009;75:848-55.
27. Hwang YH, Conklin J, Chan W, et al. Refining Genotype-Phenotype Correlation in Autosomal Dominant Polycystic Kidney Disease. *J Am Soc Nephrol*. 2016;27(6):1861-8.
28. Cornec-Le Gall E, Audrézet MP, Chen JM, et al. Type of PKD1 mutation influences renal outcome in ADPKD. *J Am Soc Nephrol*. 2013;24(6):1006-13.